Evaluation of Values Based Recruitment (VBR) in the NHS

Literature Review and Evaluation Criteria

February 2014

Acknowledgements

Contents

Figures & Tables

1. Introduction and Objectives

Introduction

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) and is recognised as a key priority for HEE and the Local Education and Training Boards (LETBs).

1.2 The report of the Mid Staffordshire NHS Foundation Trust Public Inquiry (Francis, 2013) highlighted the critical role that the workforce plays in ensuring the provision of high quality and safe healthcare services and, in particular, the significance of staff values and behaviours on the level of care and patient experience.

1.3 The values based recruitment (VBR) programme is a priority for HEE aiming 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. There are three projects in the programme. Project 1, recruiting for values in NHS funded training posts; Project 2, recruiting for values in NHS employment; and Project 3, an evaluation of the VBR programme to measure the impact of recruiting for values in higher education institutions and NHS employment. The literature review in this document forms the first key deliverable for Project 3.

1.4 Several definitions for VBR have been considered by the VBR programme board and the agreed definition differentiates between the definition, purpose and delivery as follows:

Definition: Values Based Recruitment 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.

Purpose: The purpose of Values Based Recruitment is to ensure that the future and current NHS Workforce is selected against the values of the NHS Constitution so that we recruit for the right 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.

Delivery: Values Based Recruitment can be delivered in a number of ways: through pre-screening assessments, to values based interviewing techniques, role play, written responses to scenarios, and assessment centre approaches amongst others.

1.5 The definition reflects the scope of the VBR programme and does not attempt to define the broader definition of values based employment. Figure 1 overleaf illustrates this distinction.

Figure 1 The Health Education England Model for Values Based Recruitment
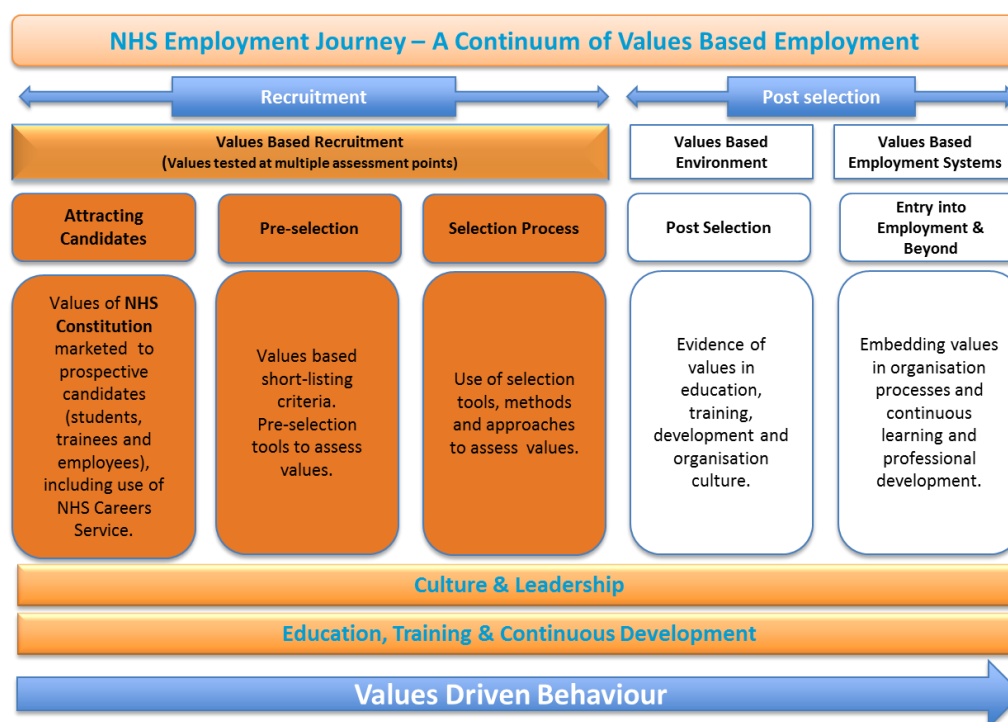
The key objectives of the evaluation of VBR include:

- Ensuring that the VBR programme is supported using an evidence-based approach. The literature review forms a core part of providing an appropriate evidence base.
- Providing the evidence base on the selection and recruitment methods available to assess values.
- Supporting and guiding the work of Projects 1 and 2 in developing a framework and guidance for VBR and understanding of evidence based tools and resources to support organisations to implement VBR locally, and the national direction of the VBR programme.
- Ensuring that the impact of recruiting for values for NHS funded training programmes and employment is evaluated through design of longitudinal tracking.

Structure of this Report

1.6   This literature review explores a number of key questions and provides the evidence base to support next steps. This report is structured into the following sections:

Section 2: Literature Review of Values Based Recruitment
Section 3: Key Concepts and Evaluation Criteria for Selection
Section 4: A Review of Selection Methods for Values Based Recruitment
Section 5: Summary, Implications and Recommended Next Steps
Section 6: References and Bibliography
Appendix: Glossary of Terms

2. A Literature Review of Values Based Recruitment

Values: what are they, and how do they link to personality, motivation and behaviours?

2.1 In reviewing the evidence base on values based recruitment (VBR), it is important to consider how values are defined in the research literature and explore how they link to other concepts that are often discussed simultaneously (for example, personality, motivation and behaviours). This is particularly important in a recruitment context as there are implications for the measurement tools available to assess each of these constructs.

2.2 Values are a set of enduring beliefs which a person holds about what is good or desirable in life. Each individual holds numerous values and a particular value may be very important to one person but unimportant to another (Schwartz, 2012). Whilst values are relatively stable over time, a person's values can change or adapt based on their experiences or environment (Rokeach, 1973).

2.3 Values influence behaviour, however this relationship is complex and simply holding a value does not necessarily mean that the individual will always behave in a way which is consistent with that value. There are other factors which are likely to influence a person's behaviour including knowledge, skills, experience, personality and motivation. The extent to which each of these factors will influence behaviour will vary depending on the given situation.

2.4 Until recently there has been limited understanding of how personality and values are related to one another, much less how they might jointly impact behaviour. Parks & Guay (2009; 2012) provide a detailed review of the personality and values literature in terms of how the constructs are distinct in order to clarify how each relates to motivation and behaviour. In summary, personality relates to enduring dispositions, whereas values relate to enduring goals.

Research on values

2.5 Schwartz and colleagues have been the dominant researchers in the values domain for much of the last two decades (Schwartz, 1992, 1994; Schwartz & Bilsky, 1987, 1990).

2.6 In general, values research has been ascribed to one of two basic models (Ravlin & Meglino, 1987a), "values as preferences" and "values as principles". Values as preferences (work values) are seen as essentially attitudes. They indicate the preferences that individuals have for various environments (Ravlin & Meglino, 1987a). Values as principles, often termed personal values, are guiding principles regarding how individuals ought to behave. For example, an individual who values honesty believes that all people ought to be honest, while an individual who values achievement believes that people ought to have many accomplishments that will be socially recognised. Research suggests that values as principles (personal values), should more directly impact motivation, because they are general beliefs that one ought to behave a certain way. Parks & Guay (2009) define values as "learned beliefs that serve as guiding principles about how individuals ought to behave". While work

values (preferences) predict vocational choice and job satisfaction, they are narrower in focus than personal values, and thus relate to a narrower range of outcomes. Personal values, however, are predictive of a broad range of behaviours across various life domains (Locke, 1997; Rokeach, 1973; Schwartz, 1992). Because personal values relate to how individuals feel they ought to behave, they have a motivational impact on behaviour in general. In the context of the NHS and VBR it is suggested that personal values, are as important to consider as work values.

2.7   Values are evaluative; they guide individuals' judgments about appropriate behaviour both for oneself and for others. Values are also general; they transcend specific situations, and values are relatively stable. Additionally, values are ordered by importance, such that one will tend to act according to the more important value when two values are in conflict. For example, consider a person who values hedonism (pursuit of pleasure) more than benevolence (concern for relationships). If forced to choose between golfing and helping his brother move, he would be more likely to golf, because he places greater importance on fulfilling personal desire than on relationship with others.

2.8   Research suggests that values develop initially through social interactions with role models such as parents and teachers. Because values are learned, there tend to be similarities in values patterns within cultures, as shared values are passed from generation to generation (Meglino & Ravlin, 1998; Oishi, Schimmack, Diener, & Suh 1998). Values are shaped during adolescence however they are generally quite stable in adulthood (Kapes & Strickler, 1975; Rokeach, 1972). Nonetheless, because values are learned initially through social interactions, being exposed to a new social environment can facilitate changes in one's values structure, which is why socialisation efforts can sometimes change the values of newcomers to become more like those of the organisation (Cable & Parsons, 2001). Not all employees respond equally to socialisation however, suggesting that some individuals are less willing to make changes in their values structures than others (Weiss, 1978).

Personality and values

2.9   Psychologists agree that the five factor model (FFM) is now a universal template to describe personality (Mount & Barrick 2002; McCrae & Costa 1997). The FFM's five factors (and examples of traits) are Conscientiousness (responsible, organized, efficient), Emotional Stability (self-confident, resilient, well-adjusted), Extraversion (talkative, ambitious, assertive), Agreeableness (friendly, cooperative, loyal), and Openness to Experience (curious, imaginative, open-minded; Goldberg, 1992; Mount & Barrick, 2002).

2.10 Personality has been shown to relate to performance, motivation, job satisfaction, leadership, and other work outcomes (personality and personality testing is discussed further in the section 4 on selection methods for VBR). Values and personality both describe components within each individual, and both are believed to impact behaviour, decision-making, motivation, attitudes, and interpersonal relations. Yet, there are also important differences. Roccas and colleagues (Roccas et al., 2002) refer to personality as enduring dispositions, and to values as enduring goals. Whilst personality generally represents the behaviours that come most naturally, values reflect effort (a choice) to behave a certain way. We do not generally think about or choose to be extraverted or introverted. However, there is an element of

personal choice involved when we behave consistently with our values. So in a given situation we may choose to be, for example, to be more extraverted than usual in order to obtain an outcome consistent with our values

2.11 There are several differences between personality and values. Values include an evaluative component lacking from personality. Values relate to what we believe we ought to do, while personality relates to what we naturally tend to do. Personality traits do not conflict with one another (i.e., one can simultaneously express the personality traits of Extraversion and Conscientiousness), yet values do conflict, as some are pursued at the expense of others.

2.12 In spite of theoretical distinctions, separating out behaviour that is caused by personality as opposed to values is difficult in practice. However, personality and values do exhibit different patterns of relationships with different constructs (Roccas et al., 2002), and a recent meta-analysis (Parks, 2007, as reported by Parks & Guay, 2009) demonstrates that personality and values are separate constructs.

2.13 Parks & Guay (2009) proposed that values and personality have different influences on different motivational processes, where values primarily impact the goals that individuals choose to pursue (goal content), while personality traits (especially Conscientiousness and Emotional Stability) primarily impact the amount of effort and persistence that individuals exhibit in pursuit of those goals (goal striving). So, values are expected to influence which goals one chooses to pursue, which is then expected to lead to effort and persistence in goal pursuit (goal striving). Goal striving is, in turn, expected to lead to goal accomplishment. Goal striving is proposed to also be impacted by the individual's personality traits, such that individuals who are more conscientious and more emotionally stable engage more effort and persistence. Parks & Guay's (2009; 2012) model provides a rationale for values contributing incrementally to behaviour (beyond personality) as depicted in Figure 2.

Figure 2. Model of relationship between personality and values on outcomes (Parks & Guay 2009)
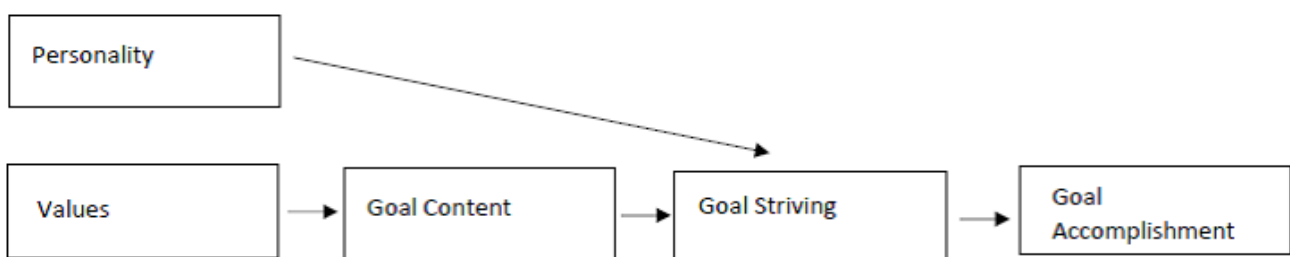
Key message

A review of the personality and values literature illustrates how the constructs are distinct in order to clarify how each relates to motivation and behaviour. In summary, personality relates to enduring dispositions, whereas values relate to enduring goals. Personality generally represents the behaviours that come most naturally, values reflect effort (a choice) to behave a certain way. This is an important distinction when considering assessment tools and measures.

Literature review methodology

2.14 A literature review search strategy was outlined in advance of the search. In summary, the search strategy for VBR set out to cover the following research questions:

- What is VBR? How does it link to existing theoretical contexts?
- What evidence is there of the impact of VBR on employee and employer outcomes?
- What are the implications for measuring and recruiting for values?

Sources, dates and search terms

2.15 'Values Based Recruitment' (VBR) is a relatively new concept within healthcare, and so a search of the published literature using this terminology was unlikely to yield a large volume of appropriate evidence. Consequently, in order to explore the evidence-base for the effectiveness of VBR the search was widened to other contexts and similar practice using a realist review methodology (Pawson et al., 2005; Wong et al., 2012). A range of sources were used to assist in the reviews. These included databases, journals, government reports, web searches and expert contacts. The search date was limited to 15 years and covered the years 1998 to 2013. It was expected that a search of the academic literature may result in a limited number of hits on values based recruitment alone. As such, other approaches and their terminology that are similar to VBR were explored, such as 'Person-Organisation Fit' (see Appendix 1 for full list). Many of the terms are broad (and identified by a *). These were included as additional filters rather than primary search terms (using advanced options), although the initial search assisted with this filtering. Search terms have been identified by their relation to the aims. These are not an exhaustive list of search terms, but were a starting point given the nature of this rapid review.

Search methodology

2.16 The methodology involved three main stages; a) a search using the database search strategies outlined above; b) filtering by abstract, where all documents/articles were scanned for relevance on the basis of title and abstract; non-relevant papers were rejected at this stage according to the inclusion criteria; c) detailed review whereby the obtained papers were read and reviewed.

Literature review results

2.17 The review of relevant databases resulted in 50 documents being identified. Following scrutiny of these papers for relevance and quality of evidence, 17 papers were retained to contribute to this report. A further three papers were identified following a review of the references within selected articles.

2.18 The majority of the papers addressing values-based recruitment were identified within the work and organisational psychology research literature. Specifically, the notion and impact of 'value congruence' between an employee and the organisation they work within (also known as 'person-organisation fit'), has been the focus of research for a number of years. It is

notable that there is an absence of a significant body of evidence from within the healthcare context specifically.

2.19 The research literature identified in this review provides several important insights regarding the impact of value congruence between employees and organisations that are highly relevant to the implementation of VBR in healthcare. However, the organisational (contextual) differences in this literature indicate that results should be interpreted with some caution as outcomes may not be immediately generalisable to a healthcare context and in particular to the NHS. Whilst the drivers for implementing VBR into the NHS are focused around the need to ensure the best possible care for patients, consistently, across professional, institutional and geographical boundaries, the literature on employees' values in other contexts and occupations may have different drivers, for example, to improve job satisfaction and productivity, or reduce staff turnover. Such drivers are often used as constructs against which the impact of VBR is measured, and these clearly differ from the aim of ensuring a caring, compassionate workforce. In addition, the NHS is an exceptionally large and highly complex organisation, with sub-cultures and significant differences across professional groups, job roles and locations. Whilst some of the literature retrieved involves studies involving large organisations, it is unlikely that they are as large, complex and heterogeneous as the NHS.

2.20 Rapping (2009) describes one example that has similar drivers for VBR to those for healthcare in the UK. He describes the need for reform of the defence workforce in the USA. Describing the broken culture of lawyers in Georgia as providing 'an example of a national crisis', and driven by a culture of incompetent practice, the authors claim that legislative approaches to address the problems through structural and financial reform would be insufficient, and a change in the values inherent within the workforce is needed. Acknowledging the significant challenges this presents and the likely resistance from existing lawyers embedded within the existing culture, the author calls for a strategy involving three elements: values-based recruitment, values-based training and values-based mentoring. At the heart of the problems within the defence system is reported a lack of 'loyalty to the client' during their practice. Such issues are resonant with the problems identified within the NHS around poor standards of patient care and putting patient needs at the centre of practice. In terms of the VBR element proposed by Rapping, he describes two steps as being necessary: (1) attracting the right pool of candidates, and (2) recruitment and selection on the basis of their values in terms of loyalty to the client. Whilst improved opportunities are called for in addressing the need to attract the best candidates (step 1), unfortunately the authors provide no further details regarding effective methods and mechanisms for VBR, other than to state that 'the leader has to devise strategies for identifying these attitudes [values] in candidates'.

Key message

The research provides several important insights regarding the impact of value congruence between employees and organisations that are highly relevant to the NHS. Whilst the drivers for implementing VBR in the NHS are focused around the need to ensure the best possible care for patients, consistently, across professional, institutional and geographical boundaries, the literature on employees' values in other contexts may have different drivers, for example, to improve job satisfaction and productivity, or reduce staff turnover. As anticipated, there was limited published research evidence relating to VBR directly.

Theoretical contexts of employee values

2.21 In terms of the context of employees' values and their impact on the organisation and workforce, a number of theories are discussed in the literature. Whilst many studies report different types of 'fit' theories, representing value congruence between employees and the organisational culture and or workforce (discussed below), two other theories that contribute to this area of research are Schneider's Attraction-Selection-Attrition (ASA) theory (Schneider, 1987; Schneider et al., 1995), and Socialisation Theories (Chao et al., 1994; Cable & Parsons, 2001; Bauer et al., 1998).

2.22 ASA theory is based on the notion that 'the people make the place' within organisations (Schneider, 1987). The theory proposes that over a period of time, the values and personalities of individuals working within a particular organisation will become increasingly homogeneous, as individuals are 'attracted' to an organisation with values that they recognise and identify with, are 'selected' by an organisation as a result of value congruence, and where shared values exist are likely to remain within the organisation (conversely, where value congruence is low, 'attrition' will occur). Consequently, it is suggested that it is the people working within the organisation that make the culture (Ostroff et al., 2005).

2.23 There have been challenges to the conditions under which some aspects of Schneider's ASA theory exist. An empirical study by Bilsberry (2007) tested the attraction aspect of ASA theory within eight utility organisations in the UK, concluding that only applicants with familiarity, proximity and previous exposure to an organisation are attracted to it which is consistent with Schneider's proposal. However, candidates appeared to select jobs to apply for on the basis of the type of role or job, rather than being attracted to the organisation as such. Considering these findings within the context of healthcare, individuals with family members working in the NHS, or those having completed placements as a student may form judgements of organisational values prior to employment and may be attracted to the organisation (or not) accordingly. However, it is also clear that the motivations for individuals' choices regarding which jobs to apply for can be affected by many complex social and personal circumstances, and individuals may be attracted to certain roles and/or organisations for extrinsic rewards such as job security, social status and income (Hollup, 2012). Within the UK, Arnold et al. (2003) explored the images of the NHS amongst 231 potential recruits and concluded that there was relatively little sign that they personally identified with the NHS.

2.24 Socialisation theories have also been linked to the development of values within an organisation following recruitment (Chao et al., 1994; Bauer et al., 1998; Cable & Parsons, 2001; de Cooman, 2009). Such theories suggest that although values are relatively stable constructs, upon entering an organisation an individual's values may change in accordance with those of their immediate work colleagues, and therefore there may be a need to continue to focus on individuals' values throughout employment (De Cooman, 2009).

2.25 Anderson (2001) explored the role of selection and recruitment methods on 'pre-entry socialisation'. He suggests that selection methods have traditionally been conceived of as neutral predictors of applicant suitability and subsequent role performance (Schmitt, Ones & Hunter 1992, Guion 1998). However, candidates form impressions of the organisation directly from their experiences in the selection and recruitment process. Anderson & Ostroff (1997) propose that selection methods initiate the pre-entry socialisation process and they outline

how selection methods vary in their degree of socialisation impact (for example, information provision). Table 1 sets out the degree of socialisation impact from a range of pre-screening and candidate assessment methods. Information Provision refers to the fact that selection methods convey information to the candidates whether intended to or not by the recruiting organisation. Preference Impact refers to the evidence that candidates prefer some selection methods over others (for example, those that are job relevant, less intrusive in areas of a personal nature). Expectational Impact refers to the propensity of candidates to generate expectations about the organisation even based on limited information during selection. These are important considerations for the NHS VBR context in terms of understanding the likely impact of chosen selection and recruitment methods on subsequent behaviour in the job.

Table 1 Socialisation impact of selection methods (Adapted from Anderson & Ostroff 1997)

Selection Method	Information Provision	Preference Impact	Expectational Impact
Pre-screening methods			
Application form	Low	Low	Low
Recruitment brochure	Medium	High	Medium/High
Realistic job previews (RJPs)	High	High	High
Assessment methods			
Unstructured interviews	Low/Medium	Medium	Medium
Behavioural interviews	Low	Medium	Low
Situational interviews	High	Medium	Medium
Cognitive ability tests	Low	Medium	Low
Personality tests	Low	Low/Medium	Low
Situational Judgement Tests	High	Medium	Medium
Work samples	High	High	High
Group discussions	Medium	Medium	Medium
Presentation exercises	Medium	Medium	Medium

2.26 The effects of both ASA and socialisation theories have been shown to operate simultaneously within organisations (De Cooman, 2009), and in the context of wishing to influence culture/values within an organisation, the evidence clearly shows the need for a multifaceted approach to organisational values beyond recruitment issues alone (Rapping, 2009). Small changes in individuals' values have been demonstrated over a period of time following initial employment (Meglino & Ravlin, 1998; De Cooman, 2009), under certain conditions.

2.27 Similarly, 'value internalisation' and 'behavioural modelling' are reported as having an impact on an individual's values following employment within an organisation (Maierhofer et al., 2000; Ostroff et al., 2005; de Cooman, 2009). Value internalisation explains the subtle change in an individual's values over time following employment as a result of their experiences, and both managers and work colleagues may have a significant impact on new recruits, being powerful role models (Maierhofer et al., 2000). Furthermore, managers also influence employees' behaviours through their own actions.

Key message

In the context of influencing culture and values within an organisation, the evidence clearly shows the need for a multifaceted approach to organisational values beyond recruitment issues alone. In terms of VBR within the NHS, individuals recruited into the organisation with optimal values for the delivery of high quality, competent and compassionate care, may be at risk of changing practice through either socialisation (value internalisation) or attrition, if placed within teams where suboptimal values are evident. Therefore VBR is only one part of embedding values in the NHS. The evidence suggests there are important considerations for VBR in terms of the likely impact of chosen selection and recruitment methods at pre entry on subsequent behaviour in the job.

Value congruence and 'fit' theories

2.28   The concept that the alignment of optimal values between employees within an organisation leads to that organisation operating in a more effective way, has been the driver for several large companies to embark on programmes promoting values, such as Disney, Hewlett-Packard, Boeing and charity organisations such as the NSPCC. However, robust evidence that such programmes have had an impact are less prominent, although research shows that individuals are more comfortable working in an environment that is consistent with their own values (Finegan, 2000).

2.29   Value congruence is often the measurable construct representing the extent to which an individual's values are similar to those of the organisation in which they work. This measure is used to define the level of 'fit' an individual has with the organisation, its culture (values) or that of the other employees within it. In an attempt to investigate the dynamics of value congruence, a number of types of 'fit' have been described in the research literature.

2.30   Value congruence is described in terms of 'Person-Organisation' (P-O) fit, 'Person-Environment' (P-E) fit, 'Person-Culture' (P-C) fit, representing the congruence between an employee's values and those of the organisation. Further, value congruence with colleagues or co-workers within the organisation have been described as 'Person-Group' (P-G) fit, or 'Person-Person' (P-P) fit, and in terms of an individual's knowledge, skills and attitudes required for the job the term 'Person-Job' (P-J) fit is often used (Ostroff & Zhan, 2012). Whilst a detailed review of the relationships between different types of fit (both in terms of 'direct' i.e. perceived fit and 'indirect' fit) can be found elsewhere (Kristof-Brown, 2002; Ostroff et al., 2005), this report provides an overview of these constructs representing value congruence to help interpret the outcomes reported within the research literature.

2.31   It is highly likely that the different types of fit are strongly related. Some argue that organisations do not have 'values' as such, but rather the organisational values are actually represented (and measured) by the values of the key members of the organisation i.e. the workforce (Meglino & Ravlin, 1998; Van Vianen, 2000; Ostroff et al., 2005). This also supports the need for caution when interpreting the findings from studies involving organisations from non-healthcare contexts, as the results may not be generalisable to an organisation that is as large and complex as the NHS.

2.32   De Cooman et al. (2009) demonstrated that individuals did not distinguish between the values of the organisation, and those of its members, upon entry. However, after two years within the organisation they were able to distinguish between these entities. Furthermore, employees

have been shown to be operating between two different subcultures (concerning their managers and their co-workers) in some instances (Van Vianen, 2000; Ostroff et al., 2005).

Key message

Organisational values are actually represented (and measured) by the values of the key members of the organisation i.e. the workforce. This supports the need for caution when interpreting research involving organisations from non-healthcare contexts, as the results may not be directly relevant to an organisation that is as large and complex as the NHS.

The NHS Constitution is frequently used to describe the values of the organisation, and the content of this document was developed by healthcare leaders and policymakers. However, the concerns raised following recent inquiries into poor standards of care suggest that variation exists across the NHS with regard to how closely individuals, teams and institutions operate. The NHS Constitution may provide an accurate representation of the values in some areas of the NHS, but represent more of an aspirational standpoint in other areas.

Impact of value congruence on job satisfaction, organisational commitment and turnover

2.33 A key objective of this literature review is to explore the evidence-base underpinning the effectiveness of VBR and how this relates to important outcomes, especially demonstrating care and compassion towards patients. However, the majority of the literature retrieved describes the impact of value congruence on other outcomes (largely from the employee perspective) such as job satisfaction and employee turnover, with very little research focusing on job performance or specific behavioural outcomes. For example, the majority of studies explore the impact of value congruence between employees' organisational commitment (affective, normative and continuance) and intended turnover (i.e. intentions to quit).

2.34 When an individual's values closely match those of the organisation (as defined by co-workers or supervisors) they report a significant increase in job satisfaction, and satisfaction with the organisation (Meglino & Ravlin, 1998; Kristof-Brown, 2002; Verquer et al., 2003; Amos & Weathington, 2008). Perceived congruence between employee and organisational values are also positively associated with satisfaction (Amos & Weathington, 2008).

2.35 In terms of an individual's commitment to an organisation, their perception of the degree of similarity between the organisational values and their own values is key (Hoffman & Woehr, 2005; Finegan, 2000). Affective (where a person is emotionally attached to an organisation), normative (where an individual has feelings of obligation towards an organisation) and continuance commitment (where an individual is committed as a result of accumulating investments in the organisation) are each predicted by different clusters of values, with 'humanity' values (defined in this study as courtesy, consideration, co-operation, fairness, forgiveness and integrity) being most associated with affective commitment to an organisation (Finegan, 2000).

2.36 De Cooman et al. (2009) demonstrated that where perceived value congruence between an individual and the organisation was low, the individual was more likely to leave that organisation over time. The negative relationship between value congruence and intended

turnover or attrition is also evident in other studies (Meglino & Ravlin, 1998; Verquer et al., 2003; Amos & Weathington, 2008).

Impact of value congruence on organisational performance

2.37 Few studies explored the impact of value congruence on organisational performance outcomes. Ostroff et al. (2005) reported that value incongruence was likely to lead to frustration, difficulty in working effectively with others and a lack of role clarity from the perspective of the employee. In addition, values have little impact on actual performance or work behaviours if task or situational variables are in place that restricts the behaviour from taking place (Meglino & Ravlin, 1998).

Measurement tools for values

2.38 Some researchers believe values should be measured using a normative approach (measuring values independently of each other, with individuals rating their agreement or otherwise with statements representing a particular value for example), whereas others favour an ipsative approach (where values are ranked against each other; Meglino & Ravlin, 1998). An ipsative approach may be more appropriate when wishing to understand an individual's preference or choice if presented against alternative courses of action (or as a ranking priority in a Situational Judgement Test, for example). Where the evidence-base supporting the impact of VBR on performance is limited (at best), the consideration of approaches to measurement is important.

2.39 The diversity of measurement tools claiming to be of use for values-based recruitment is noteworthy, and hence generalisation of findings is problematic (Meglino & Ravlin, 1998). Some studies have been criticised for using a unilateral measure for values, since values are known to be complex constructs. Some researchers suggest that a more appropriate approach would be for measurement tools to focus on 'clusters' of related values rather than using an overall value hierarchy (Finegan, 2000).

2.40 van Vianen (2000) questions whether measurement tools for values should be included during recruitment, and that in a selection process values may be 'fakeable' (van Vianen, 2000). Sections 3 and 4 in this literature review explore further the availability of measurement tools and the criteria to evaluate their effectiveness in this context.

Key Message

The diversity of measurement tools claiming to be of use for values based recruitment is noteworthy. A more detailed evaluation of the measures is required to determine the relative effectiveness and efficiency of each approach.

The role of recruiters in VBR

2.41 Several studies describe the impact of value congruence and VBR from the perspective of the recruiter, who is usually a manager from within the organisation. Although a candidate's knowledge, skills and attitudes (Person-Job fit) were found in one study to have a stronger

relationship with recruiter decisions (on whether to appoint the individual) than value congruence (Person-Organisation fit) (Kristof-Brown, 2000), a number of studies reported the importance of the recruiter's own values in the selection decision-making process.

2.42 Kristof-Brown (2000) highlighted that although recruiters were able to distinguish between Person-Job fit (the knowledge, skills and attitudes required for the job) and Person-Organisation fit (personal value congruence with organisational values), the two constructs were in fact highly related. However, recruiters seldom agreed on the specific indicators underpinning a good Person-Job or Person-Organisation fit, often relying on idiosyncratic criteria instead (Kristof-Brown, 2000). The author speculates as to whether interviewer training might reduce this potential source of error and bias.

2.43 Importantly, evidence suggests that recruiters tend to use their own personal values as a benchmark against which to measure person-organisation fit during the recruitment process (i.e. value congruence) (Judge & Ferris, 1992; Kristof-Brown, 2000). This is not perhaps surprising given that research shows that organisations do not have values per se, but rather these are the values of key members within the organisation (Meglino & Ravlin, 1998; Van Vianen, 2000; Ostroff et al., 2005). For many organisations, it is possible that these key individuals may also be the managers responsible for recruitment.

2.44 The challenges of VBR across a large organisation are perhaps highlighted by the outcomes from a study that investigated Person-Organisation fit (value congruence) amongst newly appointed employees placed in different departments within large companies in the Netherlands (van Vianen, 2000). This study reveals that recruiters from different branches within the organisation differed in their perceptions of the organisational culture, particularly with regard to values associated with human relations and innovation. As such, when measuring values, if the methods and tools used are based on criteria that require agreement between members of the organisation, this may be a challenge in complex organisations where multiple structures and hierarchies exist (van Vianen, 2000).

Key message

A single VBR tool used for attraction and marketing purposes may be appropriate in conveying the values and behaviours expected across the NHS. However, if there is a need to assess values in order to make selection decisions (either at pre-screen or interview stage) then a single tool is unlikely to be appropriate. The research evidence suggests that a tailored approach is more likely to accurately assess the diverse requirements of different job roles across the NHS. Recruiters responsible for making decisions during VBR will have a significant impact on the outcomes. Where variation exists amongst employees with regard to their own value congruence with optimal values for high quality, compassionate care, it is important to ensure that those responsible for recruitment (particularly when interviewing) represent the values the organisation is seeking to attract.

3. Key Concepts and Evaluation Criteria for Selection

The selection and recruitment process

3.1 Figure 3 outlines the main elements involved in designing and implementing a selection system. The process starts by conducting a thorough analysis of the relevant knowledge, skills, abilities and attributes associated with performance in the target role. This information is used to construct a person specification (and job description where appropriate mapped to the organisational values). This is used to decide which selection instruments are best used to elicit applicant behaviour related to the selection criteria. In deciding to apply for a post (or a place in training), applicants will engage in self-selection where they can make an informed judgment about whether the particular role suits their skills, abilities and values.

Figure 3. Designing and implementing a best practice selection system
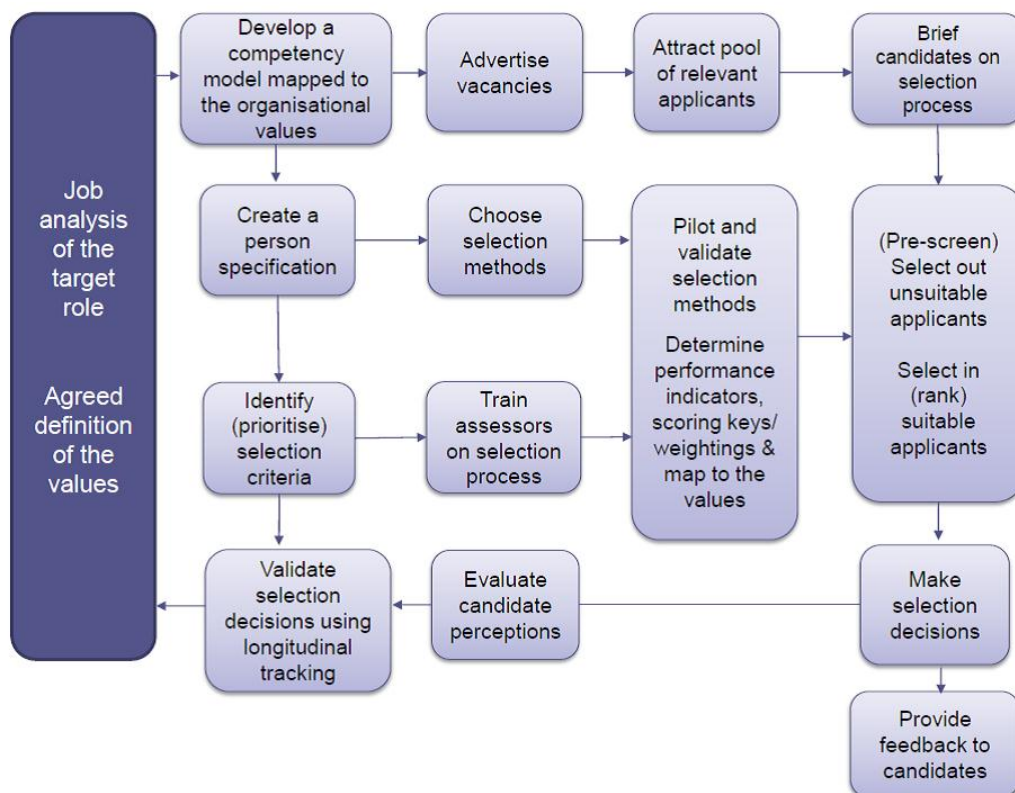
3.2 An in-depth job analysis is the cornerstone of producing an effective selection process, as the aim is to accurately identify appropriate selection criteria. In job analysis studies, researchers use various methods such as direct observation and interviews with job incumbents (Patterson et al., 2000). Having defined these criteria at a level appropriate for the career stage (for example, entry to specialty training), this information is used to guide choice of selection methods. Outputs from this analysis should detail the responsibilities in the target job and provide information about the particular competencies and characteristics required of the jobholder. The outputs from a job analysis can then be mapped to the agreed organisational values so that these are clearly represented alongside key competencies and aptitudes.

3.3 The importance of conducting a thorough job analysis cannot be understated; job analysis has been described by some as the basis for virtually all human resource activities necessary for

the successful functioning of organisations (Mirabile, 1990; Morgeson & Campion, 1997; Oswald, 2003; Siddique, 2004; Patterson et al., 2011).

3.4   In a healthcare context, job analysis studies conducted in the UK have identified a wide range of attributes beyond clinical knowledge and academic achievement that need to be considered to ensure that clinicians train and work within a profession for which they have a particular aptitude (Patterson et al., 2008). These findings support the notion that generic core values common to all roles in the healthcare environment can be identified. Job-specific competencies that differentiate between different roles and professions in a healthcare context also inform the development of robust selection criteria and provide the basis of a reliable, valid and legally defensible selection system (Patterson et al., 2008).

3.5   Once selection decisions are made, and the accepted applicants enter education, training or employment, information on the performance of students, trainees or employees related to the original selection criteria should be used to examine the predictive validity of the selection instruments (i.e. the extent to which selection scores are correlated with subsequent assessments and job performance).

3.6   Figure 3 emphasises that best practice selection is a two-way selection process. In order to attract the best students, trainees, or employees, Universities and employers have become increasingly aware that evaluating candidates' reactions to the selection process is essential, particularly in relation to perceptions of fairness. With regard to VBR, there needs to be a good 'fit' between an individual's personal values and that of the organisation, although often this is not made explicit within recruitment processes. Selection systems can be designed to make this more explicit both in attraction and also in the assessment process.

3.7   Given the size and scale of the NHS, since large resources are often spent on selection, the utility should be evaluated. This is an important consideration when making judgement about scalability of selection. In addition, information collected at selection can be used to review organisational fit within the induction and 'on-boarding' process.

3.8   The core elements of best practice selection are clear and yet research shows that there are two elements in the process that are often not conducted effectively. First, many organisations fail to conduct a thorough job analysis to identify precisely the key knowledge, skills, values and behaviours associated with competent performance and organisational fit within the target role.

3.9   Second, longitudinal validation studies are rarely conducted in organisations, as they are time consuming and difficult to administer. These studies often require tracking the performance of trainees or employees over several years – from point of selection to through to their career pathway. For example, although there is a relatively large body of research literature exploring selection issues in medical education and training, significantly more validation research has occurred within undergraduate selection (as opposed to exploring selection for employment). Furthermore, research has tended to focus on the predictive validity of various cognitive factors (for example, prior academic performance or knowledge tests) with respect to subsequent exam performance (Ferguson et al., 2002). The examinations criteria used to judge subsequent performance at medical school are potentially more readily observed as there are standardised assessments involved.

3.10   Research demonstrates that best practice selection is an iterative process. Results from evaluation studies should be used to review the original selection criteria and choice of selection methods. Thus, feedback can be used to continually improve selection systems to enhance accuracy and fairness.

Key message

Best practice personnel selection involves thorough job analysis to identify precisely the key knowledge, skills, values and behaviours associated with competent performance and organisational fit within the target role. Using feedback from validation studies is important to continually improve accuracy and fairness.

Evaluation criteria for selection and recruitment

3.11   Evaluation criteria with which to judge the effectiveness and efficiency of various selection methods have been reviewed consistently in the research literature over several decades (Arnold et al., 2010; Smith & Robertson, 1993; Schmitt, 2012; Patterson, 2012; Cleland et al., 2012). These evaluation criteria can be categorised into four broad domains, including: accuracy and effectiveness; cost and efficiency; practicalities associated with the implementation; and, stakeholder acceptance and feedback. Table 2 details sixteen key evaluation criteria in each of these domains relevant to the VBR agenda.

3.12   When reviewing the evaluation criteria, it is important to note that each criterion is not mutually exclusive. Furthermore, it is also possible for some criteria may be at odds with one another. For example, a highly reliable and valid selection tool may be very expensive. Similarly, a highly valid tool may not be acceptable to key stakeholders and may generate negative candidate reactions. As a result, it is important that the consideration of these criteria should be tailored (or weighted) depending on the recruitment context. Each criterion can apply to either a single selection tool or method (for example, an interview) or to a selection process in its entirety.

3.13   In Table 2, we have also provided some illustrations of how each criterion can be evidenced. For example, in order to demonstrate fairness of a selection tool, evidence can be collected by asking candidates to complete an appropriately designed evaluation questionnaire after the (recruitment process has taken place) to measure their perceptions of fairness. Similarly, a researcher could examine the potential for sub-group differences in scoring patterns for the selection methods used to judge the fairness of the selection methodology.

Key message

Evaluation criteria with which to judge the effectiveness and efficiency of various selection methods in general have been established in the research literature. Sixteen key evaluation criteria relevant to the VBR agenda are identified with a description of how each criterion could be evidenced. Each criterion is not mutually exclusive and it is also possible for some criteria to be at odds with one another, for example a highly reliable and valid selection tool may be very expensive.

Table 2 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. 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*

	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 or tailored for use in another context or role*	• *Judgement by recruiters*
Practicalities and implementation	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.*	• *Judgements by recruiters and administrators*
	12. Expertise required for analysis & interpretation of information generated by the tool	*Some selection tools (e.g. 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 (eg. 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 feedback	14. Positive employee/ trainee/ student perceptions	*The extent to which employees, trainees or students react positively to the selection process and each selection method within that process. Positive perceptions will result in the candidate being more attracted to joining the organisation*	• *Evaluation questionnaires of candidate perceptions*
	15. Generates appropriate feedback	*When using selection tools, e.g. personality assessments, it is good practice to ensure that candidates receive appropriate and useful feedback.*	• *Evaluation questionnaires of candidate perceptions* • *Judgements by recruiters*
	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*

Practical issues in conducting longitudinal validation studies in selection and the criterion problem

3.14   Executing validation studies is complex in practical terms since researchers would rarely use one single predictor to make selection decisions. Applicants are judged on multiple selection criteria (depending on the specific role) which may include indicators of aptitude, in addition to values. For example, given the multifaceted nature of the role of a nurse, recruiters are likely to design multiple selection tools to assess the selection criteria. Therefore, recruiters must decide whether an applicant must score highly on all selection criteria (non-compensatory) or whether high scores on some criteria can make up for low scores on another (compensatory). In practice, recruiters assign different weightings to various selection criteria, depending on the nature of the job role. If values are the most important criterion and an applicant doesn't demonstrate these, they may not be considered further.

3.15   In theory, the way to collect criterion-related validity data (i.e. how well scores on a selection measure predict some future outcome) is to use a predictive (or follow-up) design. This design involves collecting predictor information for candidates (for example, interview ratings, test scores) and then following-up to gather data on their performance (for example, during their first year of employment or exams at University). Predictive validity is assessed by examining the correlation between scores at selection and subsequent criterion data (for example, through relevant in-training or work-based assessments).

3.16   Research shows it is unusual in field studies to obtain validity coefficients in excess of $r=0.5$ (Salgado et al., 2001). Cohen (1988) defined validity coefficients below $r=0.1$ as being weak, below $r=0.3$ as being moderate, and above $r=0.5$ as being strong (Anastasia & Urbina, 1997; Arnold et al., 2010). This means that relatively low correlations between selection and outcome data still provides useful information to drive improvements in selection.

3.17   Conducting validation studies in practice presents a variety of problems. One major problem is in accessing the appropriate outcome data to validate the selection process. Often the criteria used to measure performance in the job role do not match the criteria used for selection. Conversely, sometimes the criterion and predictor are similar (for example, using knowledge-based tests to predict exam performance at University) which may lead to problems of common method variance and content overlap. Ideally, predictor scores should not be used to make selection decisions until after a predictive validation study has been conducted. Practically, this is difficult to achieve and so piloting is essential to conduct an appropriate validation study. Table 3 overleaf presents three sources of error that are important to consider when conducting validation studies in selection - including sampling, measurement precision and restriction of range issues.

Table 3 Sources of error in validation studies

Sampling error

If relatively small samples are used in validation studies the results obtained may be unduly influenced by the effects of small numbers of people within the sample whose results may be unusual. As sample size increases, more reliable results are obtained. With an organisation as large as the NHS, appropriate sampling is likely to be a significant challenge in validation studies.

Poor measurement precision

The measurement of attributes at both the predictor (i.e. selection method) and criterion (i.e. job performance) stages of the validation process is subject to unsystematic error. This error (unreliability) in the scores obtained will reduce the ceiling for the observed correlation between predictor and criterion: the error is unsystematic and random, thus this element of the predictor or criterion score will not correlate systematically with anything. This means that as reliability decreases, the maximum possible correlation between predictor and criterion will decrease.

Restricted range of scores

The sample used in a validation study may not provide the full range of theoretically possible scores on the predictor or criterion measures, or both. A restricted range of scores has a straightforward statistical effect on limiting the size of the correlation between two variables. So, range restriction in a sample serves to reduce the magnitude of the observed correlation coefficient.

Validity of selection procedures

3.18   Any single validation study is unlikely to provide a definitive answer regarding the validity of a selection method. This is because a particular study can only be conducted on a sample of relevant people and has to be conducted at a particular time, using specific measures. There are likely to be factors, such as sampling, measures, and timing of the study that will influence the results in some way. To estimate the validity of a particular selection procedure, more than one study design is needed, so that error is minimised. Most selection systems combine several predictors (selection tools), such as an applicant's score on an interview and academic achievements.

3.19   In validation studies, a key issue is the value of adding another predictor in increasing the predictive power of the selection process. This is known as incremental validity. For example, recruiters may want to know how validity is improved as a result of using a personality test (rather than relying solely on interview scores). Information on the incremental validity of a specific selection tool is valuable as it allows organisations to conduct a cost-benefit analysis of using additional tools.

Key message

Executing validation studies is complex in practical terms since researchers would rarely use one single predictor to make selection decisions. Applicants are judged on multiple selection criteria (depending on the specific role) which may include indicators of aptitude, attainment, in addition to values.

Candidate reactions

3.20 Of crucial importance are candidates' reactions to different recruitment methods (Hausknecht et al., 2004), which is especially important in attraction and pre-screening when considering a VBR agenda. Considerable research has attempted to determine applicants' views on selection methods. Research tends to explain the different factors that affect applicant reactions using theories of organisational justice. Distributive justice focuses on perceived fairness regarding equity (i.e. where the selection outcome is consistent with the applicant's expectation) and equality (i.e. the extent to which applicants have the same opportunities in the selection process). Procedural justice refers to the formal characteristics of the selection process such as information and feedback offered, job-relatedness of the selection procedures, and the level of ability of the staff involved in the selection process (i.e. recruiter effectiveness; Anderson et al., 2001).

3.21 Four main factors account for applicant reactions, relating to the extent to which selection methods:

- Are based on a thorough job analysis and appear more job-relevant;
- Are not personally intrusive;
- Do not contravene procedural or distributive justice expectations; and
- Allow applicants to meet in person with the recruiters.

3.22 The research literature also shows that applicants prefer multiple opportunities to demonstrate their skills (as in selection centres) and prefer selection systems that are administered consistently for all applicants. In particular, when competition ratios are high, applicant reactions and candidate expectations of fair play are crucial (Patterson et al., 2011).

Key message

When considering a VBR agenda, the evidence suggests that candidate perceptions towards different recruitment methods are especially important in attraction and pre-screening.

---

4. A Review of Selection Methods for Values Based Recruitment

Introduction

4.1 There is a large volume of research examining the effectiveness of different selection methods as predictors of job performance across all occupational groups. This section outlines the relative accuracy of different selection methods for healthcare specifically (both for entry to education and training, and for employment). We review the latest research evidence surrounding specific selection methods that are directly relevant to VBR. A total of seven selection methods are reviewed including; (i) interviews, (ii) references/referee's reports, (iii) CVs, application forms, personal statements, and autobiographical submissions, (iv) personality tests (v) emotional intelligence tests, (vi) situational judgment tests (SJTs), and (vii) selection/assessment centres. Following this review, the implications are summarised for future practice in relation to VBR in the NHS.

4.2 In assessment for selection the intention is to predict who will be a competent clinician, for example. In other words, to identify those individuals who will be competent in the role and/or successfully complete training, before training commences. Undergraduate (pre-entry) selection may have different goals to selection practices for those entering employment. For example, historically undergraduate courses in healthcare have tended to select primarily on academic ability, whereas selection for employment focuses more on job-fit. It cannot be assumed that those with high academic ability alone can be turned into effective clinicians via education and training - other attributes and qualities, and the right values need to be present from the start. As a result, an individual's values are important to assess at each level of entry.

4.3 Recently there has been much debate about University admission processes within healthcare internationally (Prideaux et al., 2011). Faced with limited student places and large numbers of applicants, traditionally most Universities have relied upon academic criteria in admission procedures. This assumes that with a good level of academic ability, the other skills required to be an effective clinician are trainable. However, researchers recognise that future clinicians should not be selected solely on academic performance criteria (Greengross, 1997; Hughes, 2002; Reede, 1999).

4.4 Conceptually, a key issue is whether Universities are aiming to select individuals who will make successful students or those who will make competent clinicians (McManus, 2003). Clearly, success as a student and competence as a clinician are not mutually exclusive, and the former is not necessarily a precursor of the latter.

> **Key message**
>
> Historically undergraduate courses in healthcare have tended to select primarily on academic ability, whereas selection for employment focuses more on job-fit. Furthermore, it cannot be assumed that those with high academic ability alone can be turned into effective clinicians via education and training - other attributes and qualities, and the right values need to be present from the start.

4.5 The vast majority of research literature available focuses on selection for medicine and dentistry and here the focus has largely been on pre-entry recruitment for medical and dental school admissions. Many of these studies are cited in this section of the report. It is notable that significantly less literature is available for postgraduate selection and for employment in the healthcare context for many other job roles. It is widely recognised that an effective nurse requires a complex mix of skills relating both to clinical competence as well as values, attitudes and behaviours (Nelson, 2004) and that this has implications for selection (Land, 1994). However, research into which selection methods are best able to predict success as a nurse student and practitioner, or indeed other healthcare roles, is severely lacking. Some studies report on perceptions of particular selection techniques (Rhodes & Nyawata, 2011), but there is an absence of research exploring the extent to which performance at selection predicts performance, both in terms of clinical competence and behavioural outcomes.

4.6 In practice, the selection criteria used not only vary within and between healthcare professions, but also within and between locations. An important consideration in designing selection systems is the reactions of key stakeholders to the methods used (for example, reactions towards the use of interviews versus personality tests). Stakeholder acceptance of the methods used - beyond the reactions of applicants and recruiters, to include stakeholders such as the regulators and the general public - are also an important aspect of selection system design (Patterson et al., 2012a).

Selection methods

4.7 Over several decades, many different selection methods have been reviewed by researchers across many different occupations (Campion et al., 1997; Lievens & Thornton, 2005; Ryan & Ployhart, 2013; Salgado & Anderson, 2002). Historically, there has been a great emphasis on identifying which methods are the most reliable and valid. However, current research also advocates including analyses of selection system design, in how best to combine selection methods for pre-screening as well as at final stage selection.

(1) Interviews

4.8 Interviews are ubiquitous in the selection processes across all occupational groups internationally (Campion et al., 1997; Ryan et al., 1999; Wilk & Cappelli, 2003; Macan, 2009). Interviews are used at different stages of the selection process, either as the sole method of selection, or in conjunction with other methods. Interviews can vary in terms of; (i) purpose, (ii) duration, (iii) mode of administration (e.g. telephone, face-to-face, video conference), (iv) number of interviewers (i.e. one to one or panel) and (v) degree of structure (i.e. unstructured, semi-structured or structured, or multiple mini-interviews).

4.9 Among organisational leaders, interviews have been found to be the selection method most preferred by supervisors (Lievens, Highhouse & De Corte, 2005) and human resources (HR) practitioners (Topor et al., 2007). Research consistently shows that applicants perceive interviews as fair as compared to other selection procedures, such as cognitive ability tests (Hausknecht, Day & Thomas, 2004) and applicants expect interviews as part of a selection process (Lievens, De Corte & Brysse, 2003).

4.10   As a selection method, the interview has been a popular topic among researchers for over 100 years and several large-scale meta-analytic studies show that "structured" interviews can display relatively high levels of validity without the adverse impact typically found with cognitive ability tests (Conway, Jako & Goodman, 1995; Huffcutt & Arthur, 1994; Huffcutt & Roth, 1998; McDaniel et al., 1994; Wiesner & Cronshaw, 1988). However, it is important to note that there is substantial variability in what researchers define as a structured interview. Often researchers classify interviews as either 'structured' or 'unstructured' (traditional). However, it is apparent that under the heading 'structured interview' there are at least two types of methodology including, "situational interviews" (scenario-based), and "behavioural interviews" (competency-based).

Situational interviews

4.11   In situational interviews, interviewees are asked to imagine a set of circumstances and then indicate how they would respond in that situation. Questions are future oriented (rather than past-oriented as in behavioural or competency-based interviews). One advantage of situational questions is that all interviewees respond to the same hypothetical situations rather than describe experiences unique to them from their past. In this way, an advantage is that situational questions allow respondents who have had no specific job experience relevant to a particular question to provide a response. Two core aspects of the situational interviews are the development of dilemmas that employees are likely to encounter in the role, and scoring guides should be determined in advance for trained assessors to evaluate responses to each situation.

4.12   Situational interviews are different to SJTs as interviewers are able to probe responses whereas SJTs have pre-determined fixed response choices and are machine-markable. This makes situational interviews more labour intensive compared to SJTs and in SJTs, a broader number of standardised scenarios can be presented in a short space of time when compared to interviews. However, both methods are complementary, where SJTs tend to be used more for short-listing purposes as they can be machine-marked or delivered on-line to large numbers of candidates. SJTs can also target a broad range of competencies, whereas situational interviews provide a more in-depth analysis (although more labour intensive) using probing interview questions. An example situational interview question is: "You are in a meeting. Your manager blames you for not doing well on a task, in front of all your peers and managers from other departments. You believe that your manager is wrong in her criticism, and that she might have come to this conclusion hastily without knowing all the information. You feel you are being treated unfairly in front of your peers. You feel that your reputation may be affected by this criticism. What would you do in this situation and why?"

Behavioural (competency-based) interviews

4.13   These interviews use past-oriented questions in that interviewees are asked to relate what they did in past jobs or life situations that are relevant to the particular job relevant knowledge, skills, and abilities required for success. This style of interview is based on the premise that past behaviour is the best predictor of future performance in similar situations. By asking questions about how applicants may have approached situations in the past that are similar to those they will face on the job, employers can infer how they might perform in future

situations. A behavioural interview question example is "Describe a situation in which you were able to use persuasion to successfully convince someone to see things your way. What did you do and what was the outcome?".

Validity of interviews in general

4.14 Research consistently shows that criterion-related validity is highest for interviews that are structured, ask relevant and standardised questions based on thorough role analysis, and utilise a panel of interviewers trained in best practice interview techniques and use validated scoring criteria (Campion et al., 1988; Goho & Blackman, 2006; McDaniel et al., 1994). Meta-analytic studies have found that structured interviews are valid predictors of job performance (McDaniel et al., 1994; Wiesner & Cronshaw, 1988).

4.15 Research evidence also suggests that structured interviews have incremental validity over cognitive ability tests (Cortina et al., 2000; Schmidt & Hunter, 1998) and result in relatively small differences between ethnic groups (Bobko et al., 1999). Adding structure to an interview may also increase the chances of an organisation successfully defending a lawsuit (Posthuma et al., 2002).

4.16 Unstructured interviews are still widely used for selection in a variety of occupations, despite their low reliability, low predictive validity, and thus poor legal defensibility (Klehe, 2004; Terpstra et al., 1999; Williamson et al., 1997). Unstructured interviews are prone to potential biases and errors, including: (i) stereotyping, (ii) first impressions (for example, making a judgment solely on first impressions rather than allowing the candidate a chance to demonstrate their skills (i.e. "I know if they are the right person immediately"), (iii) halo and horns effects (for example, selectors being unduly influenced by one positive or negative characteristic of the applicant) and, (iv) leniency. All of these aspects are likely to distort interviewer ratings of candidates (Edwards et al., 1990).

4.17 Researchers have examined the constructs that interviews actually measure. Several meta-analytic studies (Berry et al., 2007; Roth et al., 2005; Salgado & Moscoso, 2002) have examined a wide variety of constructs from cognitive ability to personality, although evidence generally shows low construct-related validity evidence. However, interviews that are better designed and developed specifically to assess particular constructs show greater evidence of construct-related validity. In this respect, in principle, appropriately designed structured interviews are likely to be capable of assessing values.

4.18 Despite the evidence supporting interviews as good predictors of general job performance, there is limited research on the reliability and validity of interviews for healthcare settings in general (and even less focusing on assessing values). There is however, a relatively large body of literature exploring interviews used for medical school admissions. Here, meta-analyses have shown that over a fourteen-year period, interviewer reliability of structured medical school selection interviews (i.e. interviewers giving similar scores to similar interview performances) ranged from r=.27 to r=.38 (Kreiter et al., 2004). Furthermore, a high degree of variability has been identified between interview formats and the characteristics they purport to measure, meaning different types of interviews may assess different interviewee characteristics (Albanese et al., 2003).

4.19 Postgraduate interviews for medicine have been shown to be susceptible to interviewer bias, whereby candidates are awarded preferential ratings if their personality inventory scores are similar to those of the interviewers (Quintero et al., 2009). Within dentistry, a five-year longitudinal study into the development and implementation of the Canadian Dental Association structured interview showed that a good level of inter-rater reliability can be achieved by; (i) conducting a thorough job analysis to identify key competencies, (ii) using a critical incident technique to develop relevant structured interview questions, and (iii) training the interviewers appropriately (Meredith et al., 1982).

Key message

When designed appropriately, structured interviews could be used for values-based recruitment, although they are a relatively labour intensive (and so expensive) selection method. The reliability and validity of an interview is dependent on developing interview questions following best practice.

Multiple-Mini Interviews

4.20 The multiple-mini interview (MMI) is a method to structure the delivery of interviews and is used widely for medical school admissions. MMIs comprise a series of short interview stations (for example, five minutes), each of which employs a single standardised short interview scenario and a single rater (interviewer). The primary purpose of the MMI is to overcome problems with the test-retest reliability of traditional panel interview techniques (Eva & Reiter, 2004).

4.21 The MMI has been investigated psychometrically at a number of centres internationally and has been shown to have good reliability and validity when designed appropriately (Eva et al., 2004a, 2004b; Eva et al., 2009; Roberts et al., 2008). Moreover, candidate and interviewer reactions have been positive (Kumar et al., 2009; Razack et al., 2009). MMIs have been successfully implemented in several schools worldwide (Harris & Owen, 2007; Lemay et al., 2007; Roberts et al., 2008). MMIs have also been shown to be reliable and valid in selecting suitable candidates for postgraduate positions (Hofmeister et al., 2009). Investigation into the financial utility of the MMI for selecting students into medical and dental programmes is ongoing (van der Vleuten, 1996). However, initial findings suggest that although MMIs require more physical space than traditional interviews, they can be cost-effective (Rosenfeld et al., 2008). Practically, there is a trade-off here between cost and reliability (i.e. the more interview stations the higher reliability but increased interviewer time and cost).

4.22 There is emerging consensus on the credibility, feasibility and acceptability of MMIs as an undergraduate and postgraduate medical selection tool (Dore et al., 2010; Dowell et al., 2012; Prideaux et al., 2011). However, it has been argued that for the MMI method to be valid, the design of MMI stations should be closely mapped to outputs from a thorough job analysis study and selection criteria (Patterson & Ferguson, 2012). Furthermore, there is limited research evidence exploring the construct validity of MMIs and this is a significant gap in the research literature.

Group interviews

4.23 A group interview is a selection tool which involves an assessor asking questions of two or more candidates simultaneously. There are several reasons why this form of assessment may be appealing to employers and Universities, especially for those dealing with high volume recruitment. First, group interviews require less interviewer resource than traditional one-to-one or panel interviews and may therefore be considered more cost effective. Employers may also be attracted to the idea that the nature of the group interview allows for an assessment of how well a candidate manages the interaction with others during the interview itself.

4.24 There are however, important limitations of group interviews which should be weighed against these perceived benefits. Whilst there is evidence to suggest that group interviews offer better predictors of performance than academic criteria (Byrnes, Kiger & Shechtman, 2003) they appear to be significantly worse predictors when compared to traditional one-to-one interviews (Tran & Blackman, 2006).

4.25 There are several possible explanations to explain why group interviews are less effective selection tools. First, whilst group interviews may rely on fewer assessors, the demands placed on the assessor(s) responsible for administering the interview is likely to be higher than in other selection methods. In facilitating a group interview, the assessor(s) will be responsible for delivering questions and recording responses for multiple candidates. This requirement to multitask and the potential for information overload may compromise the quality and validity of the interviewer's selection decisions.

4.26 Whilst the way in which the questions asked in a group interview will vary across organisations, a common characteristic of this method is that candidates will have the opportunity to hear one another provide responses. This has two implications; firstly this lack of privacy may cause a candidate to feel more apprehensive and self-conscious and therefore to disclose less information. Second, candidates are likely to be influenced by responses from others and the quality of their own response may therefore be dependent on the order in which the questions are asked (Tran & Blackman, 2006). Both of these factors are likely to negatively impact on the accuracy of the selection decisions made.

4.27 There is evidence to suggest that candidates are significantly more likely to perceive the group interview format as unfair and inappropriate than a traditional one-to-one interview (Tran & Blackman, 2006). This has implications for how a candidate views the recruiting organisation; increased perceptions of unfairness are more likely to lead to a candidate turning down a job offer or appealing the selection decision.

Key message

Although group interviews are more cost efficient in terms of assessor time, evidence for reliability and validity is lacking. There is an increased likelihood that candidates will perceive group interviews as unfair compared to other interview techniques.

---

(2)   References/referee's reports

4.28   Large-scale empirical studies consistently show that references tend to be unreliable and ineffective at predicting job performance (McCarthy & Goffin, 2001; Ferguson et al., 2003; Muchinsky, 1979). Despite these findings, references are widely used in selection in a variety of occupations, including healthcare, and it is likely that they will continue to be used as an additional guide in the selection process (Muchinsky, 1979; IRS Employment Review, 2002). In practice, employers tend to value references, even though references tend to be poor at differentiating between candidates. In a healthcare selection context, there is limited evidence that references are reliable, and there is no evidence that references measure anything different from interviews (Prideaux et al., 2011). Anecdotally, low scores on reference reports can be informative. Practically, recruiters tend to favour references as an employment record rather than for use in ranking candidates (Patterson et al., 2012).

4.29   In the UK, references for undergraduate applicants are used. However, the reliability is often questionable given recent changes in data legislation, which remove the confidentiality that existed previously (Hughes, 2002). In studying predictive validity, Ferguson and colleagues 2003) showed that references obtained though the Universities and Colleges Admissions Service (UCAS) did not predict pre-clinical or clinical performance. However, medical schools differ in terms of the weight placed on references obtained through the UCAS application (Parry et al., 2006).

4.30   Stedman et al. (2009) examined the content of referees' reports, and found that the writers of these reports typically apply positive and negative attributions homogeneously across applicants, thus making it impossible for admissions committees to differentiate between applicants on the this basis of these data. Stedman and colleagues concluded, therefore, that the utility of referees' reports in student selection is questionable at best. Parry and colleagues (2006) conducted a review of admissions processes for medical courses within English universities, and found that the vast majority of schools include referees' reports as part of their selection process.

4.31   Overall, there is clear consensus among researchers that referees reports were of limited use in predicting subsequent job performance. Researchers are generally critical of the inclusion of referees' reports in selection, and remarked that the information they contain may unduly bias admissions committees. A recent study based in the context of veterinary medicine, stated explicitly that referees' reports have been shown to be biased, too flattering to their subjects, and not good predictors of performance (McConnell et al., 2011).

Key message

The use of referees' reports remains widespread in selection although there is little research evidence to support the validity. There is a substantial discrepancy between evidence and practice in selection with respect to referees' reports.

---

(3)   CVs, application forms, personal statements and autobiographical submissions

4.32   The curriculum vitae (CV) is often the first form of contact between an applicant and an organisation and can be influential on the outcomes of the subsequent selection process. CVs usually comprise hard verifiable items such as education and work experience, and soft items such as candidates' interests. However, despite their wide use, the unstandardised nature of CVs leaves their predictive validity questionable at best (Highhouse, 2008).

4.33   Application forms are often used as an alternative to CVs because they are a more structured method for short-listing candidates. The information obtained through application forms is collected in a systematic way, making it easier for employers to assess objectively the candidate's suitability for a given post, and to compare applicants. Application forms may include questions on biographical information, educational background, previous work experience and competencies identified through a job analysis. Application forms are a crucial part of the selection process and the quality of information obtained varies according to the design of the form. Research shows that structured application forms can provide valid information about a candidate, and demonstrate incremental validity as a predictor of future performance beyond the contribution of clinical problem solving tests, as long as they are based on appropriate selection criteria obtained through a job analysis (Patterson et al., 2009). However, the validity of application forms is threatened by the developing industry of online resources and organisations that provide model answers to questions (Plint & Patterson, 2010).

4.34   Overall, the quality of research evidence on personal statements and autobiographical submissions for use in selection is limited. Several cross-sectional studies can be identified within the healthcare context, designed to assess the effectiveness of autobiographical submissions in predicting subsequent performance in University and the evidence is mixed, at best. For example, Oosterveld & ten Cate (2004) report that autobiographical submissions have low reliability compared to other common selection instruments. They conclude that the contents of written testimonies including autobiographical submissions are not likely to reflect the genuine nature of candidates as well as selection methods like interviews or observations. Similarly, Ferguson et al. (2000) showed that scores on personal statements were not predictive of subsequent success at University overall.

4.35   White et al. (2011; 2012) suggest that University applicants attempt to present themselves in personal statements and autobiographical submissions in ways that they perceived to be most desirable and that were not necessarily accurate. As such, the information captured by autobiographical submissions is likely to be both partial and subjective in nature.

4.36   Parry et al. (2006) reported inconsistency between UK medical schools in terms of how data from autobiographical submissions were used. Some medical schools formally used the information in making selection decisions, while others ignore this information due to concerns that it may unfairly bias selection decisions. White et al. (2102) also suggests that applicants are influenced by their understanding of what the admissions committees are looking for in candidates and they as applicants should behave in order to gain admission. Elam et al. (2002) reported that the contents of candidates' application forms are very unlikely to exert any significant influence on decisions made by admissions committees.

4.37 Despite the widespread use of personal statements and autobiographical submissions in selection, the research evidence fails to support the effectiveness and utility of such methods. There is a dearth of high-quality, generalisable evidence on the predictive validity of autobiographical submissions, and researchers have highlighted numerous limitations and sources of inconsistency with the use of these selection instruments.

Key message

Despite the widespread use of personal statements and autobiographical submissions in selection, the research evidence fails to support the validity, effectiveness and utility of such methods.

(4) Personality testing

4.38 The last thirty years have seen a substantial increase in the use of personality and related tests in personnel selection for a broad spectrum of jobs (Barrick et al., 2001; Barrick & Mount, 2012; Ones et al., 2007). Over several decades of research, personality researchers have agreed a general taxonomy of personality traits, the Big Five model, which is based on five factors or traits:

- Extraversion (i.e., outgoing, sociable, impulsive)
- Emotional Stability (i.e., calm, relaxed)
- Agreeableness (i.e. trusting, cooperative, helpful)
- Conscientiousness (i.e. hardworking, dutiful, organised) and
- Openness to Experience (i.e., artistic, cultured, creative).

4.39 Results from various meta-analytic studies (Salgado et al., 2003; Barrick & Mount, 2012) suggest that conscientiousness is a valid predictor of job performance across most jobs and organisational settings (with a criterion-related validity of approximately r=.23). Ones et al. (2007) for example, reports the conscientiousness facet 'achievement' as predictive of 'overall job performance and 'task performance' (r=.18 and r=.22, respectively).

4.40 In general, meta-analytic studies show that stronger relationships exist between personality and behaviour when performance is more discretionary and volitional (for example, citizenship behaviour, counterproductive behaviours and training success). Openness to experience in particular tends to be positively correlated with training performance across many job roles. Similarly, emotional stability has been shown to be positively associated with job performance across many organisational settings. Extraversion is generally found to correlate positively with performance in some jobs such as sales, where performance is judged within an interpersonal environment.

Key message

In relation to personality, conscientiousness, the disposition to be hardworking, thorough, and persistent, is the trait that most strongly predicts job performance in general.

4.41 Barrick & Mount (2012) highlight that validity of personality variables varies according to the outcome variable and job type, for example, conscientiousness correlates most highly with overall job performance, compared to all other personality dimensions. However, conscientiousness is also negatively correlated with creativity. In this way, depending on the target role, the facets of conscientiousness are differentially important. Hough & Furnham (2003) report that dependability (a facet of conscientiousness) correlates r=.18 with overall job performance in sales jobs and r=.03 with overall job performance in managerial jobs. In addition, when comparing the validity for facets of conscientiousness for managerial jobs in particular, results show that achievement correlates r=.17 with overall job performance whereas dependability correlated r=.03 with overall job performance. Overall, the research suggests that personality variables can be useful in predicting job performance and the predictive accuracy is increased when the predictor and criterion variables are closely matched in terms of their complexity, or job type for example.

Key message

Personality variables can be useful in predicting job performance and the predictive accuracy is increased when the predictor and criterion variables are closely matched in terms of their complexity, or job type.

4.42 Conscientiousness (the disposition to be hardworking, thorough, and persistent) is the trait that most strongly predicts intellectual academic performance (Noftle & Robins, 2007; Poropat, 2009). Lievens, Ones & Dilchert (2009) found that conscientiousness was the only trait that consistently predicted intellectual performance throughout medical school. In medicine, conscientiousness has been shown to be a positive predictor of pre-clinical medical school exam results (Ferguson et al., 2003; Ferguson et al., 2000; Lievens et al., 2002) and shows incremental validity over knowledge-based assessments (Ferguson et al., 2000; Ferguson et al., 2003). However, while positively associated with pre-clinical knowledge, conscientiousness has also been found to be a significant negative predictor of clinical skill (Ferguson et al., 2000). Therefore, the relationship between personality traits and performance in clinical education and training may be complex and possibly nonlinear.

4.43 In other healthcare settings, concerns over the strong reliance on academic predictors have led to the search for alternative selection methods. Specifically, some researchers have explored the role of personality tests in selection at undergraduate level. The Personal Quality Assessment (PQA) has been piloted for use in medical school selection in Australia and Scotland (Nicholson, 2005), and has been shown to be effective at differentiating candidates (Lumsden et al., 2005; Powis, 2009; Powis et al., 2005). However there is currently limited evidence of the predictive and construct validity of the PQA available (Prideaux et al., 2011).

4.44 In overview, the use of personality tests to assess characteristics of job applicants remains controversial. Critics argue that the predictive validity of personality traits for job performance is often low (Tett et al., 1999). Further, personality tests used by organisations are often poorly chosen (Murphy & Dzieweczynski, 2005), and "faking" can compromise the validity of personality tests (Birkeland et al., 2006; Ross et al., 1998).

4.45 A common concern for practitioners involved in recruitment is whether job applicants could fake (or intentionally distort) their responses on a personality measure and present themselves in a socially desirable manner. Morgeson et al. (2007) questioned whether the potential for applicant faking in the completion of personality measures, detracts from their effective usage for selection purposes. Many authors have debated the issue of faking and response distortion. Hough and Furnham (2003) suggest that when instructed to do so (in a laboratory setting, for example), people can distort their responses to self-report personality measures in either a positive or negative direction, depending on the instruction. However, in real-world settings, the majority of the evidence suggests that intentional distortion does exist but it may not have a substantial influence on the criterion-related validity of personality measures. There is also evidence to suggest that faking, or responding in a socially desirable way, does not compromise the predictive validity of personality tests (Hough et al., 1990). Research suggests that in real-life settings distortion is reduced when warnings about the detection of faking and the potential consequences are included in the administration instructions to applicants. In practice, most personality measures used in selection include a scale to assess social desirability and intentional distortion. Tett & Christiansen (2007) argue that of those predictive validity studies conducted in real-life settings, results demonstrate that faking, even if it does have an effect on validity, does not remove validity entirely.

Key message

Practically, in operational high stakes selection, there are concerns regarding faking with personality assessments.

Where there is a high risk of susceptibility to coaching, it is best practice to use personality assessment to drive more focused questioning at interviews (rather than a stand-alone instrument without verification).

---

(5) Emotional intelligence testing

4.46 Emotional intelligence (EI) has been defined as "the ability to perceive accurately, appraise, and express emotion; the ability to access and/or generate feelings when they facilitate thought; the ability to understand emotion and emotional knowledge; and the ability to regulate emotions to promote emotional and intellectual growth" (Sternberg & Kaufman, 1998). This ability is considered as an important for clinicians and healthcare workers in relation to clinical interactions with patients, their families, and other healthcare providers (Edwards et al., 2001). Thus, the use of EI tests has attracted recent attention within the research literature as a tool for student selection for those entering a healthcare professions in both nursing (Rankin, 2013) and in medicine (Carr, 2009; Carrothers, 2000; Leddy et al., 2011; Libbrecht et al., 2013). Each study suggests that EI relates to desirable personal and interpersonal skills for the healthcare professions.

4.47 Rankin (2013) examined the predictive relationship between emotional intelligence and the following programme outcomes for student nurses: clinical practice performance; academic performance and retention. Student nurse applicants (n=307) were asked to complete self-report scales to establish a total score and four sub-scores for emotional intelligence and

these scores were matched to individual student's performance on the undergraduate programme. A significant predictive relationship was found between emotional intelligence and programme outcomes after controlling for prior academic achievement, age and gender. Rankin (2013) argues that selection procedures could consider EI as a legitimate additional entrance criterion for student nurses. However, further research is required to examine the potential relationship between EI and compassionate care as this relationship is as yet unclear. The outcomes measured were end of first year nurse education studies so additional longitudinal studies are required. Furthermore, as with personality testing, there are significant limitations with regard to self-report measures and 'faking good' and susceptibility to coaching in selection.

4.48   In this study, participants were told completing the tool would have no bearing on the selection decision so were more likely to complete it honestly. This may well change if the tool was used operationally in future. Rankin (2013) also states, "Rejecting a candidate on the basis of a self-report measure would be potentially unfair for the individual who answered truthfully and would jeopardize the validity of self-reported emotional intelligence scores" (p.2723). However, this finding could be important for the institution and may help potential students to make an informed choice about enrolling in nurse education. It was concluded that an EI tool should not be used as a selection method in itself but including such a measure "may help potential students to make an informed choice about enrolling in nurse education….as a suitable self-selection tool". (p.2723)

4.49   In summary, there is literature emerging linking EI to values-based recruitment. For example, in a recent study Libbrecht et al. (2013) demonstrated predictive validity of EI using a situational judgement test methodology for medical students (see section on SJTs below). However, the majority of research identified on the use of EI testing in selection is sparse and at a very early stage of development. With a small number of notable exceptions, the majority research are typically pilot studies or opinion pieces citing evidence as to why EI may represent a valuable tool in future selection processes.

Key message

There is emerging literature on the use of emotional intelligence tools in selection for healthcare which may be helpful in self-selection in helping student make an informed choice about career choices.

6) Situational Judgement Tests (SJTs)

4.50 Selection practices within many healthcare professions have tended to focus on assessing academic ability alone, and yet research clearly shows that a range of (such as integrity, empathy and resilience, which can be mapped to the values of the NHS) are important for effective performance as a clinician (Lumsden et al., 2005). Historically, it has been difficult to measure such attributes on the scale required to assess the large numbers of applicants to Universities and in employment (Cleland et al., 2012). A key challenge for recruiters is how best to assess the desired values reliably, since for example, large scale interviewing can be costly and the use of personality tests is problematic since there is limited evidence to support their predictive validity for selection purposes in high stakes settings.

4.51 In an international review of selection practices for the healthcare professions, Prideaux et al. (2011) ask whether situational judgement tests (SJTs) may be a valid and reliable method for assessing a broad range of these important non-academic attributes for high volume selection. To what extent can a SJT measure values that are important in clinical practice, such as integrity, caring, empathy and compassion? Patterson et al. (2012) undertook a recent systematic literature review on the use of SJTs for selection into the healthcare professions (see also Table 4 for example SJT items and response formats). A summary of the evidence is presented in this section in relation to VBR. This is provided in some detail to provide the reader unfamiliar with SJTs with a clearer understanding of what they are, how they differ to other methods and the substantial evidence base that exists in healthcare settings.

4.52 Although the SJT methodology has been in existence for several decades, in the past 10 years their use has become increasingly popular in large-scale selection across all occupational groups internationally. SJTs are a measurement methodology rather than a single style of test presentation, (i.e. there is no single type of SJT, they can be constructed differently using different formats depending on the requirement of the role). SJTs are designed to assess an applicant's judgement regarding situations encountered in the work place (or in education settings), targeting professional attributes rather than knowledge.

Table 4. Example Situational Judgement Test items and response formats

SJT item using a multiple response format	SJT item using a ranking response format	Three SJT questions nested in one scenario using a rating response format
You review a patient on the surgical ward who has had an appendicectomy done earlier on in the day. You write a prescription for strong painkillers. The staff nurse challenges your decision and refuses to give the medication to the patient. **Choose the THREE most appropriate actions to take in this situation.** A Instruct the nurse to give the medication to the patient B Discuss with the nurse why she disagrees with the prescription C Ask a senior colleague for advice D Complete a clinical incident form E Cancel the prescription on the nurse's advice F Arrange to speak to the nurse later to discuss your working relationship G Write in the medical notes that the nurse has declined to give the medication H Review the case again	*You are reviewing a routine drug chart for a patient with rheumatoid arthritis during an overnight shift. You notice that your consultant has inappropriately prescribed methotrexate 7.5mg daily instead of weekly.* **Rank in order the following actions in response to this situation** (1= Most appropriate; 5= Least appropriate). A Ask the nurses if the consultant has made any other drug errors recently B Correct the prescription to 7.5mg weekly C Leave the prescription unchanged until the consultant ward round the following morning D Phone the consultant at home to ask about changing the prescription E Inform the patient of the error	*A consultation is taking place between a senior doctor and a patient; a medical student is observing. The senior doctor tells the patient that he requires some blood tests to rule out a terminal disease. The senior doctor is called away urgently, leaving the medical student alone with the patient. The patient tells the student that he is worried he is going to die and asks the student what the blood tests will show.* **How appropriate are each of the following responses by the medical student in this situation?** Q1 Explain to the patient that he is unable to comment on what the tests will show as he is a medical student Q2 Acknowledge the patient's concerns and ask whether he would like them to be raised with the senior doctor Q3 Tell the patient that he should not worry and that it is unlikely that he will die

4.53 In an SJT, applicants are presented with written or video-based depictions of **hypothetical** scenarios and asked to identify an appropriate response from a list of alternatives. A variety of response formats can be used and these are typically classified into one of two formats: knowledge based (i.e. what is the best option) or behavioural tendency (i.e. what would you most likely do). Response alternatives can be presented in either a written (low fidelity) or a video-based (medium fidelity) format. SJTs are typically scored by comparing applicants' responses to pre-determined scoring key agreed by subject matter experts. As a selection tool, SJTs are growing in popularity because they have useful levels of face, content and predictive validity and scenarios used in SJTs are typically derived from job analysis studies (Lievens et al., 2008; Christian, 2010; Patterson et al., 2009; Motowidlo et al., 1990; Whetzel et al., 2008).

4.54 SJTs are an established method of selection for use in high volume selection for many occupational groups and they can be used to reliably select for a range of professional attributes. Although there are other selection methods which can be used to select for these attributes, such as panel interviews or personality questionnaires, SJTs offer significant advantages over these methods. For example, panel interviews are often criticised for their potential to be biased or lack standardisation (McDaniel et al., 1994) whilst personality tests offer lower face validity and are less acceptable to candidates as a selection tool (Steiner & Gililand, 1996). By contrast, SJTs offer a standardised method of objectively assessing a broad range of attributes for large numbers of applicants, whilst being face valid to candidates since scenarios used in SJTs are based on job-relevant situations. SJTs can also be used in setting where applicants have no prior job specific experience (for example, at entry to University; Motowidlo & Beier, 2010).

4.55 There is good evidence to show that SJTs are a useful methodology to evaluate a range of professional attributes for selection into medicine and dentistry, for both admissions (for example, the UK Clinical Aptitude Test (UKCAT) SJT for entry onto medical and dental degrees) and for entry into postgraduate training (for example, the Foundation Year 1 SJT for Foundation Year doctors, and Dental foundation SJT). Long-term follow studies have shown an SJT measuring empathy, integrity and resilience (used to select candidates applying for training in UK general practice) to be the best single predictor of subsequent job performance and licensing outcomes compared to other selection methods (Lievens & Patterson, 2011; Patterson et al., 2013). An SJT has been used successfully to measure applicants' interpersonal awareness in medical and dental school admissions in Belgium (Lievens, 2013). Not only might SJTs offer an objective way of reliably assessing these attributes, but they are less susceptible to group differences than other selection tools (Clevenger et al., 2001).

Key message

Situational Judgement Tests are a machine-markable selection method that can be delivered on-line. Once developed, situational judgement tests offer significant cost savings for delivery over other methods that are hand-scored by trained assessors, such as in selection centre exercises, or in personal statements.

4.56   Using an SJT methodology, Libbrecht et al. (2013) examined whether EI would predict the performance of 367 medical students in courses on communication and interpersonal sensitivity. One of the dimensions of EI, the ability to regulate emotions, predicted performance in courses on communication and interpersonal sensitivity over the next three years of medical school, over and above measures of cognitive ability and conscientiousness. The authors measured EI using a Situational Test of Emotional Understanding (STEU) and the Situational Test of Emotion Management (STEM) (MacCann & Roberts, 2008). Students were presented with realistic, job-related situations and asked to indicate what should be done to handle each situation effectively (McDaniel, Hartman, Whetzel & Grubb, 2007; Weekley, Ployhart & Holtz, 2006). Scenarios were developed with basic emotional situations and asking respondents to indicate the most appropriate response in these situations, to target two main dimensions of EI including (i) emotional understanding (the ability to identify and understand the emotions that are most likely to be elicited by specific situations) and (ii) emotional management (the ability to effectively manage emotional situations, by choosing the most appropriate emotion regulation strategies.

4.57   The promising validity of SJTs over other measures including personality (McDaniel et al., 2007) and the possibility of varying the content of SJTs to capture a variety of performance domains (such as integrity and empathy) have made them a popular selection method (Christian et al., 2010; Weekley & Ployhart, 2006). By comparison, the reliability of interviews in assessing professional attributes such as integrity is variable (Albanese et al., 2003; Patterson et al., 2012; Christian et al., 2010; Lievens & Sackett, 2007).

Key message

Situational Judgement Tests show improved validity over other selection measures including cognitive ability and personality tests. Situational Judgement Tests can capture a variety of performance domains (such as empathy and integrity) which can be mapped to organisational values.

Situational Judgement Tests and values

4.58   The theory underlying SJTs suggests they measure implicit trait policies (ITPs) and general experience (and, depending on job level, specific job knowledge). ITPs are beliefs about the costs and benefits of expressing certain traits (which guides behaviour). Thus, implicit trait policies are an individual's judgement about the relative cost and benefits of expressing certain traits in certain situations (and so are related to trait expression rather than traits, per se). For example, making a judgement that generally being agreeable in a situation (perhaps towards a patient, a colleague or a supervisor) might be a more successful strategy in dealing with the situation than being disagreeable. In this way, SJTs measure the procedural awareness about what is effective behaviour in a given situation and this is likely to be linked to an individual's values.

4.59   Like values, it is thought that ITPs are shaped by experiences in fundamental socialisation processes, such as parental modelling during childhood. This may teach the utility of, for example; agreeable expressions, that is, helping others in need, turning the other cheek or looking after ones neighbours; or disagreeable expressions, that is, showing selfish

preoccupation with one's own interests, holding a grudge or getting even or advancing one's own interests at another person's expense. For example, having entered University in early adulthood, the challenge of educational supervisors is to teach students the utility of effective behaviour in the role of nurse or doctor, for example. This socialisation is tutored during supervised clinical practice at University, regarding effective behaviour by a clinician in any given situation.

4.60   Whilst personality generally represents the behaviours that come most naturally, values reflect effort (motivational goals) and a choice to behave a certain way. We do not generally think about or choose to be extraverted or introverted. However, there is an element of personal choice involved when we behave consistently with our values (Parks & Guay, 2009; 2012). Here, we argue that an individual also has procedural awareness about the costs/benefits of taking certain courses of action (as measured in an SJT), which also guides behaviour. In a given situation we may choose to be, for example, more extraverted than usual in order to obtain an outcome consistent with our values. We may also choose to 'treat others with respect' for example. Therefore, it is likely that SJTs can be constructed to measure aspects of an individual's values since SJTs measure implicit trait policies. SJTs are not measures of ethical values per se, but more measures of an individual's awareness about what is effective behaviour in work relevant contexts, for important interpersonal domains.

4.61   Although the design is costly, the use of SJTs could be a viable and scalable methodology for addressing the VBR agenda for some roles. SJTs are already used successfully with medicine and dentistry for both University admissions as well as for employment. However, the design parameters for an SJT linked to values will vary depending on whether the purpose is for attraction versus assessment (in medicine and dentistry the purpose is primarily assessment) and so a single tool for recruitment to all roles is not appropriate however could be useful for attraction or marketing purposes.

Key message

Situational Judgement Tests (SJTs) can be designed to measure a variety of non-academic attributes beyond clinical knowledge, which are especially relevant to medical education, training and practice. Since SJTs are scored using a pre-determined key, the tests themselves are systematically scored and often computer-delivered. This can produce substantial cost savings in high volume selection as the tests can be machine-marked.

---

7)   Selection/ assessment centres

4.62   A selection centre (SC) is a selection method, not a place. Selection centres, often called assessment centres (AC) involve a combination of selection techniques such as written exercises, interviews, and work simulations to assess candidates across a number of key skills, attitudes and behaviours. SCs allow the candidate multiple situations (interview, work simulation or written exercise) to demonstrate a key skill, and to be observed by a number of trained assessors. Thus a fairer and more reliable assessment can be made (due to multiple observations of key behaviours by multiple observers). With careful design, the increased reliability should equate to greater validity and more positive candidate reactions.

4.63　SCs were first used during World War II to select military personnel. However, it was not until the 1950s that the idea developed as a selection method, when the American company AT&T applied SCs to identify people with managerial potential. Since then, SCs have been widely used in recruitment. SCs are especially popular for graduate recruitment, with IRS Employment Review estimating that over half of recruiters, and over 95% of large organisations employing more than 10,000 individuals, use selection centres for graduate recruitment (IRS Employment Review, 2004). SCs have been used for many years especially for management roles across the NHS.

4.64　In medicine, Patterson and colleagues designed a SC to select General Practitioners and the results have shown good predictive validity in this context (Patterson et al., 2005; 2013). This work has been extended to select doctors for postgraduate training in other specialties such as obstetrics and gynaecology, and paediatrics (Randall, 2006a; Randall et al., 2006b). Patterson and colleagues have also piloted the use of SCs in the UK for graduate entry to medical school (Kidd et al., 2006).

4.65　The appeal of SCs lies in their generally good levels of criterion-related validity and face validity (Hough & Oswald, 2000). Although there has been much debate regarding the predictive validity of SCs, meta-analytic studies show the average validity of SC studies to be good.

4.66　There are certain criteria that an assessment process must fulfil in order to be defined as a SC. These include: (1) explicit dimensions (also referred to as competencies) derived from a job analysis which define the key knowledge skills and abilities required by a candidate in order to perform the role they are being assessed for; (2) multiple techniques (methods) to provide information relevant to the dimensions to be assessed and the context in which those dimensions are to be demonstrated (for example, different aspects of the role); (3) multiple, trained assessors to observe and evaluate each candidate; and (4) a systematic procedure to record and rate specific behaviours as they occur. Independent assessor ratings and reports are then pooled to form an overall rating for each candidate at what is often referred to as the 'wash-up' or moderation session.

Key message

Selection centres assess an applicant on multiple competencies using multiple job-related exercises and multiple trained assessors.

4.67　The design of a SC reflects the need to assess the extent to which applicants can demonstrate these competencies. Consequently a series of exercises and assessment tools are developed that: (1) are able to elicit the required behaviours; (2) reflect the actual content of the role; (3) assess applicants' performance in a variety of job-related situations; and (4) allow for different assessors to assess these competencies over different exercises.

4.68　Typically competencies are assessed multiple times within each SC, using different exercises to assess the same competencies. Exercises assess multiple competencies each, for instance, a group exercise may be assessing leadership skills and communication skills. However, exercises should never attempt to assess all of the competencies, as this would

overload assessors with too many sets of behavioural criteria to assess (Sackett & Tuzinski, 2001). Certain exercises are best-suited to assessing particular competencies. For example, a communication skills competency can be assessed in a presentation exercise and a group exercise.

Key message

When designed appropriately, selection centres are valid predictors of job performance and could be used for VBR. Construct validity is enhanced by ensuring the content is directly relevant to the attributes of interest in the target job role. However, selection centres are relatively expensive to design and deliver.

4.69   For VBR it is possible to map SC dimensions to the values and to create specific rating scales that relate to the values. No specific research literature is available that directly addresses this issue, although theoretically, as SCs are concerned with behaviour, it is likely that (when designed appropriately) SCs can be relatively good indicators of the extent to which an individual's values are congruent with those of the NHS Constitution.

Validity of selection centres

4.70   SCs are generally assumed to have good predictive validity because assessment is based upon direct observation of job-relevant behaviours. This enables assessors to predict how candidates will behave in the job by observing them engaging in job-relevant behaviour. Hermelin, Lievens & Robertson (2007) conducted a meta-analysis of 26 studies into the predictive validity of SCs, assessed against performance ratings provided by participants' supervisors. Hermelin et al. (2007) found a correlation of r=.28 between 'overall assessment ratings' and ratings of performance by supervisors, providing recent evidence for the predictive validity of SCs to predict job-related performance. Another explanation for why SCs work is that incremental validity is achieved by using multiple methods that assess separate and distinct aspects of performance. Therefore an SC is a better predictor because each exercise adds something to the predictive power of the process.

4.71   Research has been conducted into the predictive validity of particular assessment methods used within SCs. Slivinski (2008) investigated the predictive validity of SC pencil-and-paper tests compared with the predictive validity of SC situational test measures, finding that pencil and paper tests and situational test measures were equally valid, and that used together, both types of measures added incremental validity. Meriac et al. (2008) conducted a meta-analysis of 38 studies to explore whether seven SC dimensions ('organizing and planning'; 'influencing others'; 'drive'; 'problem solving'; 'stress tolerance'; 'consideration/awareness of others' and 'communication') could add incremental validity over cognitive ability and personality. Results showed that although there was some overlap between the SC dimensions and cognitive ability and personality, the relationship was small and demonstrated that SC dimensions are in the main, distinct from cognitive ability and personality.

4.72   In terms of the cost-effectiveness of SCs, compared to other methods they are costly to develop and deliver. However, Crail (2007) reported on a survey carried out by 'Employment Review', which contacted individuals in personnel roles within 91 different private and public

sector organisations. Results showed that although SCs are known for being a very expensive method of recruitment, though 53% of those surveyed felt this level of cost was 'justified'. Furthermore, Crail's results showed that over nine out of ten employers felt that SCs are effective for use in recruitment and, 47% of employers surveyed felt that SCs are 'very effective' in recruiting new employees. However, it is unlikely that SCs can be scalable for all job roles in the NHS to address the VBR agenda.

Summary of the quality of different selection methods to achieve VBR

4.73 Selection systems design concerns an overall program of assessment comprising a combination of methods (each with their distinctive psychometric properties) to make decisions about candidate selection. The focus of selection system design is not on how much validity a single assessment method adds, rather the question is for what is a selection method most valid?

4.74 Key issues in selection systems design are scalability, utility and cost efficiency, especially in organisations with large numbers of applicants (as in the NHS), as this could seriously limit the opportunities to use certain selection methods (such as selection centres; Prideaux et al., 2011). The initial investment in the development of bespoke selection measures may be expensive at the outset, but in the medium- to long-term, this investment can translate into significant gains in utility. For example, switching from a hand-scored application form personal statement method to a machine-markable test developed in partnership with key stakeholders could significantly reduce costs in the long-term (Lievens & Patterson, 2011; Irish et al, 2011).

4.75 Stakeholder buy-in is also an important consideration (Patterson et al., 2012). At present, there are some selection practices that display little or no predictive validity (for example, referees reports). However, these practices are viewed as acceptable because various groups of important stakeholders consider them credible.

4.76 Designing an accurate selection system is a complex process, especially when addressing the question of values as it is an area that has been relatively untouched in the research literature, especially in a healthcare context. Table 5 summarises the research evidence on the utility of different selection tools against key evaluation criteria.

Table 5 Summary of the research evidence relating to the effectiveness of selection methods

Selection Methods used for VBR	Reliability	Validity	Candidate acceptability	Cost (to the organisation)	Promotes diversity/ widening access	Susceptibility to coaching
Traditional Interviews	Low	Low	High	Moderate to high	Low	High
Structured Interviews e.g. competency-based, situational, multi-mini interviews	Moderate to high	Moderate	High	Moderate to high	Moderate	Moderate
Group Interviews	Low	Low	Moderate	Moderate	Low	High
Personal statements	Low	Low	High	Low to moderate	Low	High
References	Low	Low	High	Low to moderate	Low	N/A
Situational judgement tests	High	High (only if based on a robust psychometric methodology)	Moderate to high	Low to moderate	High	Low to moderate
Personality testing	High	Moderate	Low to moderate	Low to moderate	Moderate	Moderate to high
Selection centres using work samples e.g. group exercise, written/in-tray task, presentations, interactive exercises	Moderate to high	High (only if exercises are used in combination based on a multi-trait, method approach)	High	High	Moderate	Moderate

5. Summary, Implications and Recommended Next Steps

5.1 One of the central aims of the literature review is to ensure that the VBR programme is supported by research evidence. In each section we review relevant literature and outline key issues for evaluating selection methods for VBR. Here, we summarise the key messages and provide recommended next steps.

Literature review of VBR: summary of key messages

5.2 The research evidence supports the promotion of VBR as only one part of embedding values in the NHS and emphasises the need for a multifaceted approach to organisational values beyond recruitment issues alone.

5.3 There is a complex relationship between values and other psychological attributes such as personality, ability and motivation. In summarising the relationship, values are motivational goals that influence behaviour. The complex relationship means assessing and measuring values for recruitment is challenging and less straightforward than assessing abilities and skills.

5.4 The evidence relating to VBR directly is limited. However, this review provides links to other more established concepts in the literature that inform our understanding of how to best assess values in recruitment. This review notes the existence of a diversity of measurement tools claiming to be of use for values-based recruitment. A single VBR tool used for attraction and marketing purposes may be appropriate in conveying the values and behaviours expected across the NHS. However, if there is a need to assess values in order to make selection decisions (either at pre-screen or interview stage) then a single tool is unlikely to be appropriate. The research evidence suggests that a tailored approach is more likely to accurately assess the diverse requirements of different job roles across the NHS.

5.5 Recruiters responsible for making decisions during VBR will have a significant impact on the outcomes. It is important to ensure that those responsible for recruitment (particularly when interviewing) represent the values the organisation is seeking to attract.

Key concepts and evaluation criteria for selection: summary of key messages

5.6 Key concepts in best practice selection relevant to VBR are reviewed ranging from job analysis through to longitudinal validation. Evaluation criteria with which to judge the effectiveness and efficiency of selection methods in general are well established in the research literature. Sixteen key evaluation criteria relevant to the VBR agenda are identified with a description of how each criterion could be evidenced.

5.7 The evaluation criteria include fairness, reliability, validity, scalability, efficiency, utility, practicality, validation plans, ease of interpretation, expertise required for analysis and interpretation, generality of use, susceptibility to coaching, legality, positive employee, trainee and student perceptions, generates appropriate feedback, and educational

impact. Each criterion is not mutually exclusive and it is possible for some criteria to be at odds with one another. Decisions need to be made by recruiters in how best to resolve these tensions.

5.8 The design requirements for any longitudinal validation and evaluation of the impact of the VBR programme was explored. Longitudinal tracking in selection is challenging, and requires substantial efforts to design and deliver. Identifying appropriate outcome variables is a key priority at the outset.

Selection methods for Values Based Recruitment: summary of key messages

5.9 The established evidence on the selection methods available for VBR was reviewed. Much of the evidence base on selection methods in healthcare comes from the medical and dental recruitment literature with some contributions from nursing. Evidence from the broader occupational research literature internationally is also reviewed.

5.10 The selection methods reviewed include interviews (situational, behavioural, MMIs, and group based); references; application forms, personal statements, and autobiographical submissions; personality and emotional intelligence testing; situational judgement tests; selection and assessment centres. A key message is that is it necessary to clearly articulate what constitutes each of these selection methods as there is a great deal of variability in practice.

5.11 The current evidence for the effectiveness of each method was reviewed and an analysis of the quality of different selection methods to achieve VBR. Key messages for selection methods include:

- Interviews can be effective methods for VBR when they are designed appropriately, structured, ask relevant and standardised questions based on thorough role analysis, utilise a panel of interviewers trained in best practice interview techniques and use validated scoring criteria. Group interviews, whilst appealing on cost and efficiency, lack reliability compared to other methods.
- The evidence is weak for the use of references, application forms, personal statements and autobiographical submissions for VBR.
- The use of personality tests to assess characteristics of job applicants remains controversial. Practically, in high stakes selection, best practice is to use personality assessment to drive more focused questioning at interviews.
- There is some literature emerging linking emotional intelligence to values based recruitment. However, the majority of research identified on the use of emotional intelligence testing in selection is sparse and at a very early stage of development. With a small number of notable exceptions, the majority of research is typically pilot studies or opinion pieces citing evidence as to why emotional intelligence may represent a valuable tool in future selection processes.
- Situational Judgement Tests are an established method of selection for use in high volume selection for many occupational groups and research shows improved predictive validity over personality measures. As Situational Judgement Tests can be machine marked and/ or delivered on-line, they may offer significant advantages over other methods.

- Research shows that carefully designed and run selection centres can be effective at predicting job performance across a wide range of occupations. Gains are made in reliability and validity because selection centres make use of a combination of different exercises (using a multi-trait, multi-method approach) and use standardised scoring systems to measure the selection criteria. Practically selection centres are relatively expensive to design and deliver.

5.12 Research shows that candidates prefer high fidelity selection methods, such as simulation exercises in selection centres. This can present a dilemma for employers who also need to balance the costs of implementing high fidelity selection methods with the costs of implementing lower fidelity methods with sufficient validity, such as machine marked tests.

Implications for VBR

5.13 In practice recruitment is often delivered across two discrete phases; pre-screen which includes the initial activity to narrow down a series of applications into a manageable number (i.e. selecting candidates out) and 'selection' which involves the identification of suitable candidates to offer a job role or course place to (i.e. selecting candidates in). In Tables 6a and 6b, we provide a summary of the implications of the research evidence for VBR relating to both the pre-screening and selection phases of recruitment.

Table 6a Implements of the research evidence for implementing pre-screening methods for VBR

Selection Methods	Implications of the research evidence for VBR
Personal statements	The evidence is weak for the use of personal statements and whilst personal statements have high candidate acceptability, the susceptibility to coaching is also high. The evidence suggests they are *not an effective method for VBR.*
References	The use of referees' reports remains widespread in selection although there is little research evidence to support the validity or reliability. The evidence suggests they are *not an effective method for VBR.*
Situational judgement tests	Situational judgement tests (SJTs) show improved validity over other selection measures including cognitive ability and personality tests and can be mapped to organisational values. While they can be relatively costly to design, since SJTs are scored using a pre-determined key, SJTs are machine-markable and can be delivered on-line. This can produce substantial cost savings in high volume selection as the tests can be machine-marked. Given the evidence, *they can be an effective method for VBR.*
Personality testing	Practically, in operational high stakes selection, there are concerns regarding faking with personality assessments. Where there is a high risk of susceptibility to coaching, it is best practice to *use personality assessment to drive more focused questioning at interviews* (rather than a stand-alone instrument without verification). *Personality measures may be more useful at the attraction phases of VBR* as part of self-assessment or selection for roles within the NHS.

Table 6b. Implications of the research evidence for implementing final stage selection methods for VBR

Selection Methods	Implications of the research evidence for VBR
Traditional Interviews	Across many of the key evaluation criteria, traditional (e.g. unstructured) interviews perform poorly. The evidence for traditional interviews is they lack reliability and validity and therefore are *not suitable for VBR*.
Structured Interviews e.g. competency-based, situational, multiple-mini interviews	Where interviews are based on a thorough role analysis, use structured and standardised questions with trained interviewers and appropriate scoring, these *can be effective methods for VBR*, although they are relatively resource intensive.
Group Interviews	While group interviews may appear more cost efficient in terms of assessor time, evidence for reliability and validity is lacking. Group interviews are *unlikely to be an effective method for VBR.*
Selection centres using work samples e.g. group exercise, written/in-tray task, presentations, interactive exercises	When designed appropriately, selection centres are valid predictors of job performance when the exercises are used in combination. Effective SC exercises require time to design and assessor and role actor input, and so they can be costly, but they *can be an effective method for VBR.*

5.14 Where the research evidence suggests a selection method can or is likely to be effective for VBR, in practice the method needs to be examined in relation to the evaluation criteria to judge the likely effectiveness, especially with regard to scalability. For example, calling an interview a 'structured interview' does not mean it will automatically be effective for VBR. Recruiters will need to demonstrate how an interview meets the criteria for effectiveness (for example, validity through being based on the role analysis, reliability through standardised and structured questions).

Recommended next steps for VBR evaluation

5.15 The information presented here provides an evidence base to guide next steps in VBR Projects 1 and 2. The following recommendations are presented against the three core stages of VBR; Attraction; Pre-screen; and Selection Processes/Methods.

Attraction

Recommendation 1 (Feb-May 2014): Identify VBR selection methods that are appropriate and effective for self-selection purposes. Conduct a detailed desk review of identified case examples from HEIs and NHS providers. This review will seek to differentiate between VBR tools which are primarily used for attraction and self-selection as opposed to those tools which are designed primarily for selection, either pre-screen (shortlisting) or final stage selection (interviews). An important consideration will be the extent to which any identified tools are valid and credible to key stakeholders across different professional groups.

Pre-Screen

Recommendation 2 (Feb-Aug 2014): Develop guidelines on how to effectively map NHS values to competencies and job roles for assessment purposes at pre-screen stage. Conduct a detailed desk review of identified case examples from HEIs and NHS providers. This review will identify examples where this has been conducted successfully in practice and will document guidance on how best to map values to competencies effectively.

Selection Processes/Methods

Recommendation 3 (Jan-Feb 2014): Design a data template for the baseline evaluation measure. Using the definitions and descriptions from the literature review, define what constitutes each selection method. Pilot the template and gather feedback from key stakeholders to ensure it generates sufficient granularity of data regarding the various selection methods used for VBR. Appropriately designed, the template will help HEIs and NHS providers to effectively self-assess and achieve a more accurate baseline measure for longitudinal evaluation (differentiating between methods for attraction, pre-screening, and assessment for selection). Data gathered will ultimately inform return on investment of the VBR programme in the medium to long-term.

Recommendation 4 (Feb-Aug 2014): Review case examples of VBR selection methods (from HEIs and NHS providers) using the identified evaluation criteria. Using the examples identified by Project 1 and 2, conduct a detailed desk review of the case material by applying the evaluation criteria for VBR. Where methods are new or recently introduced, guidance will be provided about the most appropriate approach to validation to measure accuracy and effectiveness in the future.

Recommendation 5 (by October 2014): Generate best practice examples of VBR selection methods for shared learning to be included in any future VBR framework. These may include examples of:
- How to design and deliver effective values based interviews (for example, identifying differences between types of interviews and strengths for VBR).
- Best practice selection centre exercises and methods (based on research evidence).

- Guidelines and examples for recruiters in making judgements about appropriate/high quality selection methods (for example, what to look for in evaluating appropriateness for VBR).

Recommendation 6 (May 2014-March 2015): Design an appropriate VBR tool(s) for roles where there is an identified need. Most of the evidence of effective VBR tools is focused on medical and dental professions. This review suggests that there would be some merit in designing tools to address the recruitment of other professions. For example, in the recruitment of pre-registration nurses and healthcare assistants (HCAs), which together account for 59% of the care workforce (Cavendish, 2013) but where there is less evidence of robust VBR tools being used in practice. This presents a significant opportunity for HEE to more effectively address VBR for these roles by developing valid pre-screening tools (other than personal statements and reports from referees), building of current pilots. A significant challenge for these roles is in managing large applicant numbers in a robust yet efficient manner for VBR. Pilots could be instigated with evidence of validity gathered by March 2015.

Recommendation 7 (by Aug 2014): Design a specification for the outcome variables to be used in the longitudinal validation and evaluation. A detailed review of the parameters and data required for longitudinal validation will be undertaken. This will include identifying the outcome variables (the criterion) upon which the success of VBR will be measured. Practical challenges to delivery are anticipated therefore sampling of identified case material is likely to be most effective. The implications regarding the practicality of data collection, analysis and consent issues will be reviewed.

Recommendation 8 (Feb-Aug 2014): Develop a strategy for dissemination of findings from the evaluation. Identify key stakeholders to target the dissemination of findings. Agree how best to disseminate the output from the literature review to a broader audience. This could include, for example, targeting a publication in a journal.

Recommendation 9 (by Aug 2014): Update and extend the Project 1 Options Paper using the evidence presented in the literature review. Working with Project 1 the previously developed paper detailing options for the national approach taken for VBR will be reviewed and extended. In particular an analysis of the cost and scalability of designing and implementing selection methods (new and existing) for VBR will be documented.

6. References and Bibliography

Albanese MA, Snow MH and Skochelak SE et al. (2003). Assessing Personal Qualities in medical school sdmissions. *Academic Medicine*, *78*(3), 313-321.

Amos EA and Weathington BL. (2008). An analysis of the relation between employee-organisation value congruence and employee attitudes. *Journal of Psychology*. 142 (6): 615-631.

Anastasia A and Urbina S. (1997). *Psychological Testing.* (7th ed). Upper Saddle River, N.J: Prentice Hall.

Anderson N, Born M and Cunningham-Snell N. (2001). Recruitment and selection: Applicant perspectives and outcomes. In N. Anderson (Ed.), *Handbook of Industrial Work and Organizational Psychology.* London: Sage, pp. 200-218.

Arnold J, Coombs C, Wilkinson A, Loan-Clarke J, Park J and Preston D. (2003). Corporate images of the UK NHS: Implications for the recruitment and retention of nursing and allied health profession staff. *Corporate Reputation Review*. 6 (3): 223-238.

Arnold J, Randall R, Patterson F, Silvester J, Robertson I, Cooper C and Den Hartog D. (2010). *Work Psychology: Understanding Human Behaviour in the Workplace.* Financial Times / Prentice Hall.

Arthur W, Day EA and Woehr DJ. (2008). Mend it, don't end it: An alternate view of assessment center construct related validity evidence. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 105–111.

Barrick, M.R., & Mount, M. K. (2012). Nature and use of personality in selection. In N. Schmitt (Ed.), *The Oxford Handbook of Personnel Assessment and Selection*. New York: Oxford University Press (New York), pp 225-251.

Barrick MR, Mount MK and Judge TA. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment,* 9, 9-30.

Berry CM, Sackett PR and Landers RN. (2007). Revisiting interview-cognitive ability relationships: Attending to specific range restriction mechanisms in meta-analysis. *Personnel Psychology*, 60, 837−874.

Bilsberry J. (2007). Attracting for values: an empirical study of ASA's attraction proposition. *Journal of Managerial Psychology*, 22 (2): 132-149.

Birkeland SA, Manson TM and Kisamore JL et al. (2006). A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment*, 14, 317-335.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117, 187-215.

Bobko P, Roth PL and Potosky D. (1999). Derivation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology*, 52(3), 561-589.

Bowler, M. C., & Woehr, D. J. (2006). A meta-analytic evaluation of the impact of dimension and exercise factors on assessment center ratings. *Journal of Applied Psychology, 91*, 1114–1124.

British Medical Association Board of Medical Education. (2007). *Selection for specialty training*.

Byrnes DA and Kiger G. (2003). Evaluating the use of group interviews to select students into teacher-education programs. *Journal of Teacher Education*, 54, 163-172

Cable DM and Parsons CK. (2001). Socialisation tactics and person-organisation fit. *Personnel Psychology*, 54; 1-23.

Campion MA, Palmer DK and Campion JE. (1997). A review of structure in the selection interview. *Personnel Psychology*, 50, 655-702.

Campion MA, Pursell ED and Brown BK. (1988). Structured interviewing: raising the psychometric properties of the employement interview. *Personnel Psychology*, 41(1), 25-42.

Carr SE. (2009). Emotional intelligence in medical students: does it correlate with selection measures? *Medical Education*.;43(11):1069-77.

Carrothers RM, Gregory SW Jr, Gallagher TJ. (2000). Measuring emotional intelligence of medical school applicants. *Academic medicine,*75(5):456-63.

Chao GT, O'Leary-Kelly AM, Wolf S, Klein HJ and Gardner PD. (1994). Organisational socialisation: Its context and consequences. *Journal of Applied Psychology,* 79: 730-743.

Christian M, Edwards B and Bradley J. (2010). Situational judgement tests: constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology,* 63, 83-117.

Cleland J, Dowell J, McLachlan J, Nicholson S and Patterson F. (2012). Identifying best practice in the selection of medical students. London: *General Medical Council*. Available from: gmc-uk.org/about/research/14400.asp

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

Cohen J. (1988). *Statistical power analysis for the social sciences*. 2nd ed. London: Earlbaum.

Conway JM, Jako RA and Goodman DF. (1995). A meta-analysis of interrater and internal consistency reliability of selection interviews. Journal of Applied Psychology, 80, 565−579.

Cortina JM, Goldsten NB and Payne SC et al. (2000). The incremental validaity of interview scores over and above cognitive ability and conscientiousness scores. *Personnel Psychology*, 53(2), 325-351.

Crail M. (2007) Employers rate assessment centres as worth the high cost. Personnel Today, in Assessment centres, HR strategy, Recruitment & retention, Yardstick Update. http://www.personneltoday.com/hr/employers-rate-assessment-centres-as-worth-the-high-cost/

De Cooman R, De Gieter S, Pepermans R, Hermans S, Du Bois C, Caers R and Jegers M. (2009). Person-organisation fit: Testing socialisation and attraction-selection-attrition hypotheses. *Journal of Vocational Behaviour.* 74: 102-107.

Deary IJ, Watson R and Hogston R. (2003). A longitudinal Cohort Study of Burnout and Attrition in Nursing Students, *Journal of Advanced Nursing*, 43, 71-81

Dore KL, Kreuger S and Ladhani M et al. (2010). The reliability and acceptability of the Multiple Mini-Interview as a selection instrument for postgraduate admissions. *Academic Medicine,* 85(10), 60-63.

Dowell J, Lynch B and Till H et al. (2012). The multiple mini-interview in the U.K. context: 3 years of experience at Dundee. *Medical Teacher*, 34(4), 297-304.

Edwards JC, Elam CL, Wagoner NE. (2001). An admission model for medical schools. *Academic medicine*, 76(12):1207-12.

Edwards JC, Johnson EK and Molidor JB. (1990). The interview in the admission process. *Academic Medicine*, 65(3), 167-177.

Elam CL, Stratton TD, Scott KL, Wilson JF, Lieber A. (2002) Review, deliberation, and voting: a study of selection decisions in a medical school admission committee. *Teaching and Learning in Medicine*, 14(2):98-103

Eva K and Reiter H. (2004). Where judgement fails: pitfalls in the selection process for medical personnel. *Advances in Health Sciences Education. 9*(2), 161–174.

Eva K, Reiter H and Rosenfeld J et al. (2004a). The relationship between interviewers' characteristics and ratings assigned during a multiple mini-interview. *Academic Medicine,* 79(6), 602-609.

Eva K, Rosenfeld J and Reiter H et al. (2004b). An admissions OSCE: the multiple mini-interview. *Medical Education*, 38(3), 314-326.

Eva K, Reiter H and Trinh K et al. (2009). Predictive validity of the multiple mini-interview for selecting medical trainees. *Medical Education*, 43(8), 767-775.

Eva K, Wood T and Riddle J et al. (2010). How clinical features are presented matters to weaker diagnosticians. *Medical Education*, 44(8), 775-785.

Eva KW and Reiter HI. (2005). Reflecting the relative values of community, faculty and students in the admissions tools of medical school. *Teaching and Learning in Medicine*, 17, 4-8.

Francis, R. (2013). *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry: Executive Summary* (Vol. 947). TSO Shop.

Ferguson E, James D and Madeley L. (2002). Factors associated with success in medical school: systematic review of the literature. *BMJ, 324*(7343), 952-957.

Ferguson E, James D and O'Hehir F et al. (2003). Pilot study of the roles of personality, references, and personal statements in relation to performance over the five years of a medical degree. *BMJ,* 326(7386), 429-432.

Ferguson E, Sanders A and O'Hehir F et al. (2000). Predictive validity of personal statements and the role of the five factor model of personality in relation to medical training. *Journal of Occupational and Organisational Psychology*, 73, 321-344.

Finegan JE. (2000). The impact of person and organisational values on organisational commitment. *Journal of Occupational and Organisational Psychology*. 73: 149-169.

Goho J and Blackman A. (2006). The effectiveness of academic admission interviews: an exploratory meta-analysis. *Medical Teacher*, 28(4), 335-340.

Goldberg, L. R. (1992). The development of markers for the Big-Five Factor structure. *Psychological Assessment*, 4, 26–42.

Greengross S. (1997). What patients want from their doctors. In I. Allen, P. Brown, and P. Hughes (Eds.), *Choosing Tomorrow's Doctors,* pp. 9-12.

Harris S and Owen C. (2007). Discerning quality: using the multiple mini-interview in student selection for the Australian National University Medical School. *Medical Education*, 41(3), 234-241.

Hausknecht JP, Day DV and Thomas SC. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology*, 57(3), 639-683.

Hermelin E, Lievens F and Robertson IT. (2007). The validity of assessment centres for the prediction of supervisory performance ratings: A meta-analysis. *International Journal of Selection and Assessment,* 15(4), 405-411.

Highhouse S. (2008). Stubborn reliance on intuition and subjectivity in employee selection. *Industrial and Organizational Psychology*, 1(3), 333-342.

Hoffman BJ and Woehr DJ. (2005). A quantitative review of the relationship between person-organisation fit and behavioural outcomes. *Journal of Vocational Behaviour.* 68: 389-399.

Hofmeister M, Lockyer J and Crutcher R. (2009). The multiple mini-interview for selection of international medical graduates into family medicine residency education. *Medical Education*, 43(6), 573-579.

Hollup O. (2012). Nurses in Mauritius motivated by extrinsic rewards: A qualitative study of factors determining recruitment and career choices. *International Journal of Nursing Studies.* 49: 1291-1298.

Hough L. M., & Furnham A. (2003). Importance and use of personality variables in work settings. In I. B. Weiner (Ed-in-Chief) & W. Borman, D. Ilgen, & R. Klimoski (Vol. Eds.), Comprehensive handbook of psychology: Industrial organizational psychology. (Vol. 12, pp. 131–169). New York:Wiley & Sons.

Hough LM, Eaton NK and Dunnette et al. (1990).Criterion related validities of personality constructs and effect of response distortion on those validities. *Journal of Applied Psychology*, 75, 581-595.

Hough L. M., & Oswald F. L. (2000). Personnel selection: Looking toward the future—Remembering the past. *Annual Review of Psychology*, 51, 631–664.

Huffcutt AI and Arthur W Jr. (1994). Hunter and Hunter (1984) revisited: Interview validity for

entry-level jobs. *Journal of Applied Psychology*, 79, 184−190.

Hughes P. (2002). Can we improve on how we select medical students? *Journal of the Royal Society of Medicine*, 95(1), 18-22.

Irish B, Carr AS and Sowden D et al. (2011). Recruitment into specialty training in the UK [Online] (Updated 12 Jan 2011). Available at: http://careers.bmj.com/careers/advice/view-article.html?id=20001789 [Accessed 28 June 2012].

IRS Employment Review. (2002). Of good character: supplying references and providing access, 754, 6-34.

IRS Employment Review. (2004). Graduate Recruitment 2004/05: upturn and optimism, 8-40.

Judge TA and Ferris GR. (1992). The elusive criterion of fit in human resource staffing decisions. *Human Resource Planning* 154: 47-67.

Kapes, J. T., & Strickler, R. E. (1975). A longitudinal study of change in work values between 9th and 12th grades. *Journal of Vocational Behavior*, 6, 81–93.

Kidd J, Fuller J and Patterson F et al. (2006). Selection Centres: Initial description of a collaborative pilot project. *Proceedings for the Association for Medical Education in Europe (AMEE)*.

Klehe UC. (2004). Choosing how to choose: Institutional pressures affecting the adoption of personnel selection procedures. *International Journal of Selection and Assessment*, 12(4), 327-342.

Kreiter C, Yin P and Solow C et al. (2004). Investigating the reliability of the medical school admissions interview. *Advances in Health Sciences Education*, 9(2), 147-159.

Kristof-Brown AL. (2000). Perceived applicant fit: distinguishing between recruiters perceptions of person-job and person-organisation fit. *Personnel Psychology*. 53 (3): 643-671.

Kristof-Brown AL, Jansen KJ and Colbert AE. (2002). A policy-capturing study of the simultaneous effects of fit with jobs, groups and organisations. *Journal of Applied Psychology*. 87 (5): 985-993.

Kumar K, Roberts C and Rothnie I et al. (2009). Experiences of the multiple mini-interview: a qualitative analysis. *Medical Education*, 43(4), 360-367.

Land LM. (1994). The student nurse selection experience: a qualitative study, *Journal of Advanced Nursing,* 20, 1030-7

Leddy JJ, Moineau G, Puddester D, Wood TJ and Humphrey-Murto S. (2011). Does an emotional intelligence test correlate with traditional measures used to determine medical school admission? Academic medicine, 86(10), 39-41.

Lemay JF, Lockyer J and Collin V et al. (2007). Assessment of non-cognitive traits through the admissions multiple mini-interview. *Medical Education*, 41(6), 573-579.

Libbrecht N, Lievens F, Carette B and Côté S. (2013). Emotional Intelligence Predicts

Success in Medical School. *Emotion.* DOI: 10.1037/a0034392

Lievens F. (2013). Adjusting medical school admission: assessing interpersonal skills using situational judgement tests. Medical Education,47(2), 182-9

Lievens F, Coetsier P and De Fruyt F et al. (2002). Medical students' personality characteristics and academic performance: A five-factor model perspective. *Medical Education*, 36, 1050-1056.

Lievens F, De Corte W and Brysse K. (2003). Applicant perceptions of selection procedures: The role of selection information, belief in tests, and comparative anxiety. International Journal of Selection and Assessment, 11, 67−77.

Lievens F, Highhouse S and De Corte W. (2005). The importance of traits and abilities in supervisors' hirability decisions as a function of method of assessment. Journal of Occupational and Organizational Psychology, 78, 453−470.

Lievens F, Ones D, and Dilchert S. (2009) Personality scale validities increase through medical school *Journal of Applied Psychology*, 94(6), 1514-1535

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.

Lievens F, Peeters H and Schollaert E. (2008). Situational judgment tests: a review of recent research. *Personnel Review*, 37(4), 426-441.

Lievens F and Sackett P. (2007). Situational judgment tests in high-stakes settings: issues and strategies with generating alternate forms. *Journal of Applied Psychology*, 92(4), 1043-1055.

Lievens F and Thornton GC. (2005). Assessment centres: recent developments in practice and research. In A. Evers, O. Smit-Voskuijl and N. Anderson (Eds.). *Handbook of Personnel Selection*. Oxford: Blackwell, pp. 243-264.

Locke, E. A., & Henne, D. (1986). Work motivation theories. In Robertson Cooper (Ed.), *International review of industrial and organizational psychology* (pp. 1–34). Chichester, NY: John Wiley & Sons Ltd.

Locke, E. A. (1997). The motivation to work: What we know. In M. L. Maehr & P. R. Pintrick (Eds.). *Advances in motivation and achievement* (Vol. 10, pp. 375–412). JAI Press Inc.

Lumsden M, Bore M and Millar K et al. (2005). Assessment of personal qualities in relation to admission to medical school. *Medical Education*, 39, 258-265.

Macan T (2009). The employment interview: A review of current studies and directions for future research *Human Resource Management* Review 19, 203–218

Macan TH, Avedon MJ and Paese M et al. (1994). The effects of applicants' reactions to cognitive ability tests and an assessment center. *Personnel Psychology*, 47, 715-738.

MacCann R. (2008). A modification to Angoff and bookmarking cut scores to account for the imperfect reliability of test scores. *Educational and Psychological Measurement*, 68 (2), 224-254.

Maierhofer NI, Griffin MA and Sheehan M. (2000). Linking manager values and behaviour with employee values and behaviour: A study of values and safety in the hairdressing industry. *Journal of Occupational Health Psychology*. 5 (4): 417-427.

Maio, G. R., & Olson, J. M. (1998). Values as truisms: Evidence and implications. *Journal of Personality and Social Psychology*, 74, 294–311.

McCarthy JM and Goffin RD. (2001). Improving the validity of letters of recommendation: An investigation of three standardized reference forms. *Military Psychology*, 13(4), 199-222.

McConnell SL, Kogan LR. (2001) Subjective criteria as the sole method of selecting veterinary candidates at a US veterinary medical school. *Journal of Veterinary Medical Education*, 28(3):131-5.

McCrae, R. R., & Costa, P. T. Jr., (1997). Personality trait structure as a human universal. *American Psychologist,* 52, 509–516.

McDaniel M, Morgeson FP, Finnegan EB, Campion MA and Braverman EP. (2001). Use of situational judgement tests to predict job performance: a clarification of the literature. *J Appl Psychol*;86 (4):730–40).

McDaniel M and Whetzel D. (2007). Situational judgement tests. In: Whetzel DL, Wheaton GR, eds. Applied Measurement: *Industrial Psychology in Human Resources Management*. Mahwah, NJ: Lawrence Erlbaum Associates; 235–57.

McDaniel MA, Whetzel DL and Schmidt FL et al. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79(4), 599-616.

McManus IC. (1997). From selection to qualification: how and why medical students change. In I. Allen, P. Brown, and P. Hughes (Eds.), *Choosing Tomorrow's Doctors*. London: Policy Studies Institute, pp. 60-79.

McManus IC. (2003). Commentary: How to derive causes from correlations in educational studies. *BMJ,* 326(7386), 429-432.

McManus IC, Powis D and Wakeford R et al. (2005). Intellectual aptitude tests and A-levels for selecting UK school leaver entrants for medical school. *BMJ,* 331(7516), 555-559.

McManus IC and Richards P. (1986). Prospective survey of performance of medical students during preclinical years. *BMJ*, 293(6539), 124-127.

McManus IC, Smithers E and Partridge P et al. (2003). A levels and intelligence as predictors of medical careers in UK doctors: 20 year prospective study. *BMJ,* 327(7407), 139-142.

Meglino BM and Ravlin EC. (1998). Individual Values in Organisations: Concepts, Controversies, and Research. *Journal of Management*. 24: 351-389.

Meredith KE, Dunlap MR and Baker HH. (1982). Subjective and objective admissions factors as predictors of clinical clerkship performance. *Medical Education*, 57(10, Pt 1), 743-751.

Mirabile RJ. (1990). The power of job analysis. *Training*, 27(4), 70-72,74.

Morgeson FP and Campion MA. (1997). Social and cognitive sources of potential inaccuracy in job analysis. *Journal of Applied Psychology*, 82(5), 627-655.

Motowidlo SJ, Beier ME. Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgment test. *Journal of Applied Psychology,* 95(2):321-333

Motowidlo SJ, Dunnette MD and Carter GW. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology.* 75(6):640-647.

Mount, M. K., & Barrick, M. R. (1995). The big five personality dimensions: Implications for research and practice in human resource management. *Research in Personnel and Human Resource Management*, 13, 153–200.

Mount, M. K., & Barrick, M. R. (2002). *The personal characteristics inventory manual.* Libertyville, IL: The Wonderlic Corporation.

Muchinsky PM. (1979). The Use of Reference Reports in Personnel Selection: A Review and Evaluation. *Journal of Occupational Psychology*, 52(4), 287-297.

Murphy KR. (2002). Can conflicting perspectives on the role of g in personnel selection be resolved? *Human Performance*, 15(1-2), 173-186.

Murphy KR and Dzieweczynski JL. (2005).Why don't measures of broad dimensions of personality perform better as predictors of job performance? *Humance Performance*, 18, 343-357.

Nelson G. (2004). Hiring the Complete Nurse: Selecting the Right Nurses Means Looking Beyond Clinical Skills, *Nurse Leader*, 2, 36-39

Nicholson S. (2005). The benefits of aptitude testing for selecting medical students. *BMJ*, 331(7516), 559-560.

Oishi, S., Schimmack, U., Diener, E., & Suh, E. (1998). The measurement of values and individualism–collectivism. *Personality and Social Psychology Bulletin*, 24, 1177–1189. Ones DS, Dilchert S and Viswesvaran C et al. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, 60(4), 995-1027.

Oosterveld P, ten Cate O. Generalizability of a study sample assessment procedure for entrance selection for medical school. *Medical Teacher.* 2004;26(7):635-9.

Ostroff C, Yuhyung S and Kinicki AJ. (2005) Multiple perspectives of congruence: relationships between value congruence and employee attitude. *Journal of Organisational Behaviour.* 26: 591-623.

Ostroff C and Zhan. (2012) Person–Environment Fit in the Selection Process. In Schmidt, N. (Eds), *The Oxford Handbook of Personnel Selection and Assessment* Oxford University Press.

Oswald FL. (2003). Job analysis: Methods research and applications for human resource management in the new millennium. *Personnel Psychology*, 56(3), 800-803.

Outtz J. (2002). The role of cognitive ability tests in employment selection. *Human Performance*, 15, 161-171.

Parks, L. (2007). Personality and values: A meta-analysis. Paper presented at the *annual conference for the society of industrial and organizational psychology*, New York, New York.

Parks, L., & Guay, R.P. (2012). Can Personal Values Predict Performance? Evidence in an Academic Setting. *Applied Psychology: An International Review 61* (1), 149–173

Parks, L., & Guay, R.P. (2009). Personality, values, and motivation. *Personality and Individual Differences*, *47*, 675–684.

Parry J, Mathers J and Stevens A et al. (2006). Admissions processes for five year medical courses at English schools: review. *BMJ*, 332(7548), 1005-1009.

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.
http://www.gmc-uk.org/Identifying_best_practice_in_the_selection_of_medical_students.pdf_51119804.pdf

Patterson F, Ashworth V, Mehra S & Falcon H. (2012a). Can situational judgement tests be used to select for dental foundation training? *British Dental Journal*. 213, 23-26.

Patterson F, Ashworth V, Zibarras L, Coan P, Kerrin M and O'Neill P. (2012b). Evaluations of situational judgement tests to assess non-academic attributes in selection. *Medical Education*, 46(9):850 – 868.

Patterson F, Baron H, Carr V, Lane P and Plint S. (2009) Evaluation of three short-listing methodologies for selection into postgraduate training: the case of General Practice in the UK. *Medical Education,* 43: 50-57.

Patterson F, Carr V and Zibarras L et al (2009b). New machine- marked tests for selection into core medical training: evidence from two validation studies. *Clinical Medicine*, 9, 417-420.

Patterson F, Denney ML and Wakeford R et al. (2011). Fair and equal assessment in postgraduate training? A future research agenda. *British Journal of General Practice,* 61(593), 712-713.

Patterson F and Ferguson E. (2012). Testing non-cognitive attributes in selection centres: how to avoid being reliably wrong. *Medical Education*, 46(3), 240-242.

Patterson F and Ferguson E. (2007). Selection into medical education and training. *ASME Monographs*, 1-56.

Patterson F, Ferguson E and Lane P et al (2001). A new competency based selection system for general practitioners*. BMJ,* 323, 2-3.

Patterson F, Ferguson E and Lane P et al. (2005). A new selection system to recruit general practice registrars: preliminary findings from a validation study. *BMJ*, 330(7493), 711-714.

Patterson F, Ferguson E and Thomas S. (2008). Using job analysis to identify core and specific competencies: implications for selection and recruitment. *Medical Education*, 42(12), 1195-1204.

Patterson F, Ferguson E, Lane P, Farrell K, Martlew J and Wells A. (2000). A competency model for general practice: implications for selection and development. *British Journal of*

General Practice 50, 188-93.

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

Patterson F, Lievens F, Kerrin M, Zibarras L and Carette B. (2012). Designing selection systems for medicine: The importance of balancing predictive and political validity in high-stakes selection contexts. *International Journal of Selection and Assessment*, 20(4), 486-496.

Patterson F, Tavabie A, Denney ML, Kerrin M, Ashworth V, Koczwara A and Macleod S. (2013). A new competency model for general practice: implications for selection, training, and careers. *British Journal of General Practice* 05; 63(610):331-8

Patterson F and Zibarras L. (2011a). Exploring the construct of perceived job discrimination and a model of applicant propensity for case initiation in selection. *International Journal of Selection Assessment*, 19(3): 251 - 257.

Patterson F, Zibarras L, Carr V, Irish B and Gregory S. (2011b). Evaluating candidate reactions to selection practices using organisational justice theory. *Medical Education*, 45: 289 - 297.

Patterson F, Denney M, Wakeford R and Good D. (2011) Making assessment fair and equal in postgraduate training? A future research agenda. *British Journal of General Practice.* 61(593), 712-3.

Pawson R, Greenhalgh T, Harvey G and Walshe K. (2005). Realist review--a new method of systematic review designed for complex policy interventions. *J Health Serv Res Policy.* 2005;10, 1:21-34.

Plint S and Patterson F. (2010). Identifying critical success factors for designing selection processes into postgraduate specialty training: the case of UK general practice. *Postgraduate Medical Journal*, 86(1016), 323-327.

Posthuma RA, Morgeson FP and Campion MA. (2002). Beyond employment interview validity: a comprehensive narrative review of recent research and trends over time. *Personnel Psychology*, *55*(1), 1-81.

Powis D. (2009). Personality testing in the context of selecting health professionals. *Medical Teacher*, 31(12), 1045-1046.

Powis D, Bore M and Munro D et al. (2005). Developmet of the personal qualities assessment as a tool for selecting medical students. *Journal of Adult and Continuing Education* , 11(1), 3-14.

Prideaux D, Roberts C and Eva K et al. (2011). Assessment for selection for the health care professions and specialty training: Consensus statement and recommendations from the Ottawa 2010 Conference. *Medical Teacher*, 33(3), 215-223.

Quintero A, Segal L and King T et al. (2009). The personal interview: assessing the potential for personality similarity to bias the selection of orthopaedic residents. *Academic Medicine*, 84(10), 1364-1372.

Randall R, Davies H and Patterson F et al. (2006). Selecting doctors for postgraduate

training in paediatrics using a competency based assessment centre. *Archives of Disease in Childhood*, 91(5), 444-448.

Randall R, Ferguson E and Patterson F. (2000). Self-assessment accuracy and assessment centre decisions. *Journal of Occupational and Organisational Psychology*, 73(4): 443 - 459.

Randall R, Stewart P and Farrell K et al. (2006). Using an assessment centre to select doctors for postgraduate training in obstetrics and gynaecology. *Obstetrician and Gynaecologist*, 8(4), 257-262.

Rankin B. (2013). Emotional intelligence: enhancing values-based practice and compassionate care in nursing. *Journal of Advanced Nursing*. 69(12), 2717-2725

Rapping JA. (2009). You can't build on shakey ground: laying the foundation for indigent defense reform through values-based recruitment, training and mentoring. *Harvard Law and Policy Review.*

Ravlin, E. C., & Meglino, B. M. (1987a). Issues in work values measurement. In L. Preston (Ed.). *Research in corporate social performance and policy* (Vol. 9, pp. 153–183). JAI Press Inc.

Razack S, Faremo S and Drolet F et al. (2009). Multiple mini-interviews versus traditional interviews: stakeholder acceptability comparison. *Medical Education*, 43(10), 993-1000.

Reede JY. (1999). Predictors of Success in Medicine. *Clinical Orthopaedics and Related Research*, 362.

Rhodes CA and Nyawata ID. (2011). Service user and carer involvement in student nurse selection: Key stakeholder perspectives, *Nurse Education Today*, 31, 439–443

Robertson IT and Smith M. (2001). Personnel selection. *Journal of Occupational and Organizational psychology*, 74(4), 441-472.

Roberts C, Walton M and Rothnie I et al.(2008). Factors affecting the utility of the multiple mini-interview in selecting candidates for graduate-entry medical school. *Medical Education*, 42, 396-404.

Roccas, S., Sagiv, L., Schwartz, S. H., & Knafo, A. (2002). The Big Five personality factors and personal values. *Personality and Social Psychology Bulletin*, 28, 789–801.

Rokeach, M. (1972). *Beliefs, attitudes, and values: A theory of organization and change.* San Francisco, CA: Jossey-Bass Inc.

Rokeach, M. (1973). *The nature of human values.* New York: Free Press.

Rosenfeld J, Reiter H and Trinh K et al. (2008). A cost efficiency comparison between the multiple mini-interview and traditional admissions interviews. *Advances in Health Sciences Education,* 13(1), 43-58.

Rosse JG, Stecher MD and Miller JL et al. (1998).The impact of response distortion on pre-employment personality testing and hiring decisions. *Journal of Applied Psychology*, 83, 634-644.

Roth PL, Van Iddekinge CH, Huffcutt AI, Eidson CE Jr and Schmit MJ. (2005). Personality saturation in structured interviews. International Journal of Selection and Assessment, 13,

261−273.

Ryan AM, McFarland L, Baron H and Page R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology*, 52(2), 359-392.

Ryan AM and Ployhart RE. (2013). A Century of Selection. *Annual Review of Psychology*. 2014. 65:20.1–20.25

Rynes SL and Connerley ML. (1993). Applicant reactions to alternative selection procedures. *Journal of Business and Psychology*, 7(3), 261-277.

Sackett P and Lievens F. (2008). Personnel selection. *Annual Review of Psychology*, 59: 419 - 445.

Salgado JF and Anderson N. (2002).Cognitive and GMA testing in the European Community: Issues and evidence. *Human Performance*, 15, 75-96.

Salgado JF and Moscoso S. (2002). Comprehensive meta-analysis of the construct validity of the employment interview. European Journal of Work and Organizational Psychology, 11, 299−324.

Salgado J, Anderson N and Moscoso S et al. (2003). A meta-analytic study of general mental ability validity for different occupations in the European community. *Journal of Applied Psychology*, 88(6), 1068-1081.

Salgado JF, Viswesvaran C and Ones D. (2001). Predictors used for personnel selection: an overview of constructs, methods, and techniques. In N. Anderson, D. S. Ones and H. K. Sinangil et al. (Eds.), *Handbook of Industrial, Work and Organizational Psychology*. Volume 1. London: Sage, pp. 99-165.

Schmidt FL and Hunter JE. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262-274.

Schmitt, N. (2012). *The Oxford Handbook of Personnel Assessment and Selection.* Oxford University Press: USA

Schmitt N, Gooding RZ and Noe RA et al. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.

Schneider B. (1987). The people make the place. *Personnel Psychology*. 40; 437-453.

Schneider B, Goldstein HW and Smith DB. (1995). The ASA Framework: An Update. *Personnel Psychology,* 48: 474-773.
Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–65). San Diego, CA: Academic Press.

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *Journal of Social Issues*, 50, 19–45.

Schwartz SH. (2012). An Overview of the Schwartz Theory of Basic Values. *Online Readings in Psychology and Culture*, 2(1).

Schwartz, S. H., & Bilsky, W. (1987). Toward a psychological structure of human values. *Journal of Personality and Social Psychology,* 53, 550–562.

Schwartz, S. H., & Bilsky, W. (1990). Theory of the universal content and structure of values: Extensions and cross-cultural replications. *Journal of Personality and Social Psychology*, 58, 878–891.

Siddique C. (2004). Job analysis: a strategic human resource management practice. *International Journal of Human Resource Management*, 15(1), 219-244.

Smith, M. & Robertson, I. (1993). *The Theory and Practice of Systematic Personnel Selection.* Palgrave Macmillan.

Stedman JM, Hatch JP, Schoenfeld LS. (2009) Letters of recommendation for the predoctoral internship in medical schools and other settings: do they enhance decision making in the selection process? *Journal of Clinical Psychology in Medical Settings*, 16(4):339-45.

Steiner DD and Gilliland SW.(1996). Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology.* 81(2):134-141.

Sternberg RJ and Kaufman JC. (1998). Human abilities. *Annual Review of Psychology.* 49:479-502

Terpstra DA, Mohamed AA and Kethley RB. (1999). An Analysis of Federal Court Cases Involving Nine Selection Devices. *International Journal of Selection and Assessment*, 7(1), 26-34.

Tett RP, Jackson DN and Rothstein M et al. (1999). Meta-analysis of bi-directional relations in personality-job performance research. *Human Performance*, 12, 1-29.

Tiffin PA, Dowell JS and McLachlan JC. (2012). Widening access to UK medical education for under-represented socioeconomic groups: modelling the impact of the UKCAT in the 2009 cohort. *BMJ*, 344, e1805.

Thomson W, Ferry P and King J et al. (2003). Increasing access to medical education for students from medically underserved communities: one program's success. *Academic Medicine,* 78(5), 454-459.

Thornton G and Rupp D. (2006). *Assessment centers in human resource management: Strategies for prediction, diagnosis, and development.* Mahwah, NJ: Erlbaum.

Tooke J. (2007). Aspiring to Excellence: Findings and recommendations of the independent inquiry into modernising medical careers. *MMC Inquiry* [Online] (Updated 28 Feb 2008)*.* Available at: http://www.mmcinquiry.org.uk/.

Topor DJ, Colarelli SM and Han K. (2007). Influences of traits and assessment methods on human resource practitioners' evaluations of job applicants. Journal of Business and Psychology, 21, 361−376.

Tran T and Blackman MC. (2006). The dynamics and validity of the group selection interview. *The Journal of Social Psychology*, 146, 183-201

Turnbull D, Buckley P and Robinson J et al. (2003). Increasing the evidence base for

selection for undergraduate medicine: four case studies investigating process and interim outcomes. *Medical Education*, *37*(12), 1115-1120.

Van der Vleuten CPM. (1996). The assessment of professional competence: developments, research and practical implications. *Advances in Health Sciences Education,* 1, 41-67.

Van der Vleuten C and Schuwirth L. (2005). Assessing professional competence: from methods to programmes. *Medical Education*, 39, 309-317.

Van Vianen AEM. (2000). Person-organisation fit: The match between newcomers' and recruiters preferences for organisational cultures. *Personnel Psychology.* 53:113-149.

Verquer ML, Beehr TA and Wagner SH. (2003). A meta-analysis of relations between Person-Organisation fit and work attitudes. *Journal of Vocational Behaviour.* 63: 473-489.

Weiss, H. M. (1978). Social learning of work values in organizations. *Journal of Applied Psychology,* 63, 711-718

Whetzel D, McDaniel M and Nguyen N. (2008). Subgroup differences in situational judgment test performance: A meta-analysis. *Human Performance.* 21:291-309.

White J, Brownell K, Lemay JF, Lockyer JM. (2012) "What do they want me to say?" The hidden curriculum at work in the medical school selection process: a qualitative study. *BMC medical education*, 12:17.

White J, Lemay JF, Brownell K, Lockyer J. (2011) "A chance to show yourself" - how do applicants approach medical school admission essays? *Medical Teacher,* 33(10):e541-8.

Wiesner WH and Cronshaw SF. (1988). A meta-analytic investigation of the impact of interview format and degree of structure on the validity of the employment interview. *Journal of Occupational Psychology*, 61(4), 275-290.

Wiley A and Koenig J. (1996). The validity of the Medical College Admission Test for predicting performance in the first two years of medical school. *Academic Medicine*, 71, 5-83.

Wilk SL and Cappelli P. (2003). Understanding the determinants of employer use of selection methods. Personnel Psychology, 56, 103−124.

Williamson LG, Campion JE and Malos SB et al. (1997). Employment interview on trial: Linking interview structure with litigation outcomes. *Journal of Applied Psychology*, 82(6), 900-912.

Wong G, Greenhalgh T and Westhorp G.(2012). Realist methods in medical education research: what are they and what can they contribute? *Medical Education.* 46(1):89–96.

Woodruffe C. (2000). *Development and Assessment Centres: Identifying and Developing Competence.* 3rd ed.. London: Institute of Personnel and Development.

Woolf K, Potts HW and McManus IC. (2011). Ethnicity and academic performance in UK trained doctors and medical students: Systematic review and meta-analysis. *BMJ*, 342, 901.

Wyatt MRR, Pathak SB and Zibarras L. (2010). Advancing selection in an SME: Is best practice methodology applicable? *International Small Business Journal.* 28(3):258-273.

7. Appendix 1

Search terms for the values-based recruitment literature review

Values based recruitment
Values based assessment
Selection methods*
Person-organisation fit
Professional attributes*
National Health Service*
Healthcare*
Principles*
Professional standards*
Morals*
Ethics*
Ideals*
Evaluation*
Doctor*
Nurse*
NHS employees*

Glossary of terms

Selection methods

1.1   Selection/Assessment centres are selection methods involving multiple exercises or techniques and multiple assessors to rate a candidate's performance on a series of job-relevant competencies.

1.2   Curriculum vitae (CV) can enable the systematic collection of candidates' biographical information. Information collected from CVs is typically objective and verifiable; such as educational background and previous work experience.

1.3   Multiple mini-interviews (MMI) is a type of interviewing technique. It comprises a series of short test stations within a circuit that the candidate will go through; each station employs a single standardised short interview scenario and a single assessor.

1.4   Personal statement is a self-report of the candidate's perception of his or her suitability for the position that they are competing in the selection process for. Personal statements may

include an elaboration of the biographical information that is presented in a CV and/or a self-assessment of the target competencies that are assessed in the recruitment process.

1.5   Personality measures are self-report questionnaires that assess a candidate's personality traits.

1.6   Situational Judgement Tests are designed to assess a candidate's judgement regarding work-relevant scenarios are presented to candidates, who then make

judgements about possible responses, and these responses are then assessed against a predetermined scoring key (defined by subject matter experts).

1.7 Structured interviewing is a selection method whereby the assessor asks the candidate a set of question that is specified in advance and mapped to the selection criteria. This approach improves standardisation across interviewers.

Psychological terms

1.8 Personality traits are underlying dimensions of psychological characteristics along which individuals differ from one another. Personality traits can influence the way an individual behaves. These traits are relatively stable especially during adulthood. There are five fundamental dimensions of personality: extraversion, emotionality, agreeableness, conscientious and openness to experience.

1.9 Attitudes relate to an individual's feelings, thoughts and predispositions to act towards some aspect of their environment. Attitudes are evaluative; that is, they reflect a person's tendency to feel, think or behave in a positive or negative manner towards the specific target or object.

1.10 Behaviours are external, observable displays of an individual's conduct.

1.11 Cognitive abilities may also be referred to as intelligence or general mental ability.

1.12 Construct validity measures the extent to which the selection methods measure the domains or qualities that they are intended to measure.

1.13 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 Matters Experts (SMEs).

1.14 Face validity is achieved when the selection tool content appears to be relevant towards the target role.

1.15 Internal reliability measures the extent to which different parts of the same measure (questions designed to measure the same thing within the assessment) produces results that are consistent with each other.

1.16 Motivation is factors that determine the effort, direction and persistence of a person's behaviour.

1.17 Parallel-form reliability measures the extent to which two equivalent test forms (similar content and equal level of difficulty) produces similar candidate scores on both test forms. It is intended to measure the external reliability between scales.

1.18 Predictive validity measures the extent to which outcomes of the selection process are accurate predictors of performance in role. For example, a high predictive validity may indicate that the candidate's score on the qualifying test predicted future work performance.

1.19 Reliability refers to the extent to which the selection tool is able to measure consistently across conditions. Internal reliability, parallel-form reliability and test-retest reliability are three common forms of reliability.

1.20 Test-retest reliability is form of external reliability that measures the extent the same candidate who completes the same assessment across a significant time lag of administration would produce the same assessment score. In assessing this form of reliability, it is very important that testing conditions are standardised and that the construct to be measured in relatively stable.

1.21 Validity of tools/processes: A selection method is valid to the extent that it is able to measure what it claims to measure. To minimise error, best practice advises the evaluation of five types of validity: Face validity, content validity, construct validity and predictive validity.

1.22 Values are a person's beliefs about what is good or desirable in life. They are long-term guides for a person's choices and experiences. Values affect one's perspective on how things should be and guides an individual's approach to life. Values are relatively stable with little change over long periods of time.