Assessing professionalism for the selection of mental health clinicians: the development and validation of a situational judgement test

Lauren Margaret Ellen Aylott

PhD in Medical Sciences

The University of Hull and The University of York

Hull York Medical School

August 2022

Abstract

Patients using mental health services are vulnerable to abuse and exploitation, yet there is no consensus on assessment for values-based recruitment (VBR) in this setting. VBR requires evaluating attitudes, knowledge and traits relevant to the delivery of professional and effective care. Structured interviews are one approach to VBR but are resource intensive to deploy. In contrast, situational judgement tests (SJT) are generally valid predictors of job performance and can be delivered more cost-effectively, and at scale. This project developed and validated a SJT that aimed to assess one's knowledge of professionalism in a mental health services context.

A mixed methods programme of work was conducted, incorporating a rapid systematic review to define professionalism in mental health services, a qualitative study exploring stakeholder perspectives, and a quantitative study that assessed the criterion-related validity of the resulting SJT. The literature review included 70 articles and resulted in two operational definitions of professionalism. The subsequent qualitative study incorporated interviews and focus groups with 56 patients, carers and staff members from a range of professions, who facilitated the development of the SJT. Finally, a pilot study, which involved 170 mental health professionals, was conducted to evaluate the validity of the SJT.

The SJT scores validly predicted workplace supervisor ratings of professionalism and effectiveness for nurses and allied health professionals. Its predictive ability was comparable to that previously reported for face-to-face interviews. In conclusion, the SJT can be reliably deployed in mental health services to assist values-based recruitment by identifying candidates who are unable to demonstrate reasonable levels of knowledge related to professional behaviours. This novel tool can support our efforts to reduce the risk of patient abuse and reported instances of malpractice in mental health services.

List of contents

List of tables

List of figures

Acknowledgements

I would not have been able to undertake this work had my former boss Dr Angela Kennedy not sent me the initial advertisement for the PhD; she was continually presenting me with good opportunities. Had Hull York Medical School and the University of York not provided me with the financial assistance they did, I would have struggled to make ends meet.

There are many people that have contributed to this project, and I could not have completed the research or the thesis without them. I am firstly indebted to my supervisors Professor Gabrielle Finn and Professor Paul Tiffin; they helped conceptualise the project's aims and objectives and guided and supported me to establish and conduct the methods that are reported on in this thesis. Without their constructive feedback and continual encouragement, I would never have completed these works. I am grateful for the opportunity they gave me to undertake the project, as well as the support they have provided throughout, in good times and in bad. They have inspired and encouraged me and kept me going in times of crisis. I must also acknowledge my thesis advisory panel: Professor Dean McMillan and Jan Boehnke. Dean and Jan have provided an opportunity for me to reflect critically on my work and focus on the necessary tasks to complete the thesis without getting distracted by other objectives.

Many individuals, including academics and NHS employees have helped with this work and there are far too many to mention: from administrators, to support staff, and individual mentors, to people who have facilitated recruitment, and to others sharing their expertise and/or belief in the importance of this project, to name but a few. Of course, it is also necessary to acknowledge the participation of patients, carers, and staff members whose insights and pearls of wisdom have proved invaluable – thank you.

Last, but not least, I must turn my attention to those particularly close to me. I am grateful to all my friends, and in particular my fiancée Rachel; she has continued to believe in me and has never doubted my capability to complete this doctorate. Rachel and I first met

when I started the PhD and she has been there to see it through with me, despite all the challenges posed. The opportunity to commence a PhD would never have arisen without the love and lifelong support of my parents and my siblings (Ben and Megan). Mum and Dad's teachings have led to my desire to undertake this work and the values I have learned from them underly my yearning to generate an assessment that would enhance recruitment into mental health services. It is to them that I dedicate this thesis. Finally, I must acknowledge my cats – Toby and Tinker; Toby has kept my knees warm whilst I have been working from home post-Covid-19; meanwhile, Tinker has kindly woken me up on a morning ready for the writing ahead.

Statement of contribution

This thesis reports on the findings of a systematic review, a qualitative study, and a quantitative study. The author is grateful, and credit is due to the following individuals given their level of contribution to these works.

The overarching thesis

I cannot thank Professor Gabrielle Finn and Professor Paul Tiffin enough for their part in this doctoral work - they have played a key role throughout. Gabrielle and Paul have provided significant input into all aspects of the project. Alongside supervision and helping conceptualise the project, Gabrielle and Paul have provided specific input into the literature review, and each study, providing support with obtaining a favourable ethical opinion, co-facilitating focus groups, helping analyse the data and supporting the final write up.

Chapter 4: A rapid systematic review

It is good practice to have a minimum of two researchers involved in a literature review to minimise the risk of error and bias (Centre for Reviews and Dissemination, 2009). In this instance, the author was supported with the review by both supervisors as well as further individuals from Hull York Medical School (Dr Mona Saad), the University of York (Alexis Llewellyn, and Adrian Clark) and Tees, Esk and Wear Valleys (TEWV) NHS Foundation Trust (Naomi Hay-Gibson). Mona Saad (FY2 doctor) supported with various elements of the review process including the development of the review protocol and search strategy, the screening of articles, the quality checking of a sample of data extracted from articles by the lead reviewer, and facilitating the data analysis and write-up. The author had direct support from the Centre for Reviews and Dissemination, and Alexis Llewellyn helped to devise the review protocol and provided additional support with issues that presented throughout the review period. Both Naomi Hay Gibson (healthcare librarian) and Adrian Clark (university librarian) provided advice on the search strategy, and Naomi provided mentoring support to the author during the initial phase of the review, which included commenting on the

protocol. Frederic Hafferty and Madeline Carter kindly commented on earlier drafts of the manuscript prior to its publication.

Chapter 5: Understanding professionalism (a qualitative study)

Alongside both supervisors, Dr Sally Brown at Edinburgh Napier University provided support with the qualitative study. Specialising in qualitative research, Sally provided supervisory input to the lead researcher (LA) throughout the study process. Sally supported with the analysis of data also and the subsequent write up of the study for publication. The study could not have been completed without the support of Tees, Esk and Wear Valley's NHS Foundation Trust as well as the kind involvement of staff, patients and carers that participated in the interviews and focus groups.

Chapter 6: Piloting the SJT (a quantitative study)

It is important to acknowledge the staff at TEWV for their assistance in bringing the pilot study to fruition. In particular, Mr David Levy was instrumental in enabling this study to occur and supporting the development of the idea. TEWVs Research and Development department, and professional leads working in the organisation helped facilitate recruitment. Dr Sarah Dexter-Smith also drove the study forward and provided additional support to implement the SJT within the personnel selection process at Tees, Esk and Wear Valleys NHS Foundation Trust. This process is ongoing at the time of submission of this thesis. Thanks are also due to PCMIS for their part in developing a final electronic version of the SJT.

List of scientific contributions resulting from this thesis

Publications

Aylott, L., Tiffin, P. A., Brown, S. & Finn, G. M. 2022. Great expectations: Views and perceptions of professionalism amongst mental health services staff, patients and carers. *Journal of Mental Health*, 31, 139-146, DOI: 10.1080/09638237.2020.1818195 (see Appendix A).

Aylott, L., Tiffin, P. A., Saad, M., Llewellyn, A. R. & Finn, G. M. 2019. Defining professionalism for mental health services: a rapid systematic review. *Journal of Mental Health*, 28, 546-565, DOI: 10.1080/09638237.2018.1521933.

Conference presentations

Aylott, L., Finn, G. M, Tiffin, P. A. & Brown, S. "Defining Professionalism for Mental Health Services: A Qualitative Study." ASME (Association for the study of Medical Education) Annual Scientific Meeting: Sustainability, Transformation and Innovation in Medical Education 2019, 3-5 July, Glasgow.

Aylott, L., Tiffin, P. A. & Finn, G. M. "The Predictive Validity of a Situational Judgement Test for Assessing Professional Judgement in Mental Health Services." HYMS (Hull York Medical School) Postgraduate Research Conference 2021, 8-9 July, Virtual Conference.

The lead author was awarded the best Poster Presentation Prize for their contribution at this conference.

Aylott, L., Tiffin, P. A., Saad, M., Llewellyn, A. R. & Finn, G. M. "Professionalism of Mental Healthcare." ASME (Association for the study of Medical Education) Annual Scientific Meeting 2018, 11-13 July, Gateshead.

Aylott, L., Tiffin, P. A., Saad, M., Llewellyn, A. R. & Finn, G. M. "The Professionalism of Mental Healthcare: a Rapid Review." HYMS (Hull York Medical School) Postgraduate Research Conference 2018, 3 July, Hull.

Author's declaration

I confirm that this work is original and that if any passage(s) or diagram(s) have been copied from academic papers, books, the internet, or any other sources these are clearly identified by the use of quotation marks and the references(s) is fully cited. I certify that, other than where indicated, this is my own work and does not breach the regulations of HYMS, the University of Hull, or the University of York regarding plagiarism or academic conduct in examinations. I have read the HYMS Code of Practice on Academic Misconduct, and state that this piece of work is my own and does not contain any unacknowledged work from any other sources. I confirm that any patient information obtained to produce this piece of work has been appropriately anonymised.

Chapter 1: Introduction

This thesis reports on the background, rationale, and methods used to develop and validate a Situational Judgement Test (SJT) that assesses individuals' knowledge of professionalism for mental health services. This initial chapter discusses the context for this work and provides a brief overview of the background underlying the doctoral project. The initial aims and objectives of the project are highlighted, and the structure of the thesis is presented. Whilst this thesis reports on work completed within the NHS, healthcare services are delivered internationally, and the findings of this doctoral work are hopefully relevant, to some extent, to all services and staff that deliver mental healthcare.

1.1   The NHS

The National Health Service (NHS) was founded, in England, in 1948 to deliver healthcare free at the point of delivery for all United Kingdom (UK) residents (Triggle, 2018). The principles and values of the NHS are established within the NHS Constitution, which sets out the rights and responsibilities expected of the public, patients and staff members with regard to healthcare delivery under the NHS (Department of Health & Social Care, 2021); the Constitution is renewed every 10 years in collaboration with the public, patients and staff members and comprises of seven key principles that guide the NHS:

1. The NHS provides a comprehensive service, available to all
2. Access to NHS services is based on clinical need, not an individual's ability to pay
3. The NHS aspires to the highest standards of excellence and professionalism
4. The patient will be at the heart of everything the NHS does
5. The NHS works across organisational boundaries
6. The NHS is committed to providing best value for taxpayers' money
7. The NHS is accountable to the public, communities and patients that it serves

The above principles are underpinned by six NHS core values, including 'Working together for patients', 'Respect and dignity', 'Commitment to quality of care', 'Compassion', 'Improving lives'; and 'Everyone counts' (see Department of Health & Social Care, 2021). Expanding on the third principle (i.e., that 'the NHS aspires to the highest standards of

excellence and professionalism), the constitution states that the NHS '*provides high quality care that is safe, effective and focused on patient experience*'. Recognising the need for suitable staff and leadership, the Constitution acknowledges that high quality care will be achieved and delivered by staff members receiving the right support, education, training and development and among other factors, through research to improve the current and future health and care of the population (Department of Health & Social Care, 2021).

In keeping with the principles and values of the NHS Constitution, staff members wishing to work as healthcare professionals in the UK are required to undertake training and obtain the appropriate qualifications to register with a chosen health regulator (e.g., the Nursing and Midwifery Council; NMC). A list of health regulators relevant to mental health services is provided in Table 1.1.

Table 1.1 Health regulators applicable to mental health services

Health regulator	Professional group
Health & Care Professions Council (HCPC)	Health, psychological and other allied health professionals
Nursing and Midwifery Council (NMC)	Nurses and midwives in the UK, and nursing associates in England
The General Medical Council (GMC)	Doctors
Social Work England	Social Workers

Each of the health regulators have their own standards and expectations that professionals working within that discipline are expected to adhere to; for example, the Nursing and Midwifery Council have 'The Code', which outlines the professional standards expected of all nurses, midwives and nursing associates according to four themes: prioritise people, practise effectively, preserve safety and promote professionalism and trust (Nursing and Midwifery Council, 2021). Similarly, the General Medical Council has guidance that reports on what it means to be a good doctor; the guidance document Good Medical Practice outlines the values and behaviours expected of doctors, and thus psychiatrists, across four

sections: knowledge, skills and performance; safety and quality; communication, partnership and teamwork; and maintaining trust (General Medical Council, 2021).

Despite the expectations placed on healthcare professionals, there remains instances when healthcare goes wrong and the services, and staff, do not live up to these standards; well cited examples include the scandal at Mid Staffordshire NHS Foundation Trust, and more pertinent to mental health and learning disability services, the abuses uncovered at Winterbourne View care home (e.g. Cafe, 2012).

1.1.1 Mid-Staffordshire NHS Foundation Trust Scandal

Between 2005 and 2008, Mid-Staffordshire NHS Foundation Trust (NHSFT) failed to protect patients who were faced with "*unacceptable risks of harm and from unacceptable, and in some cases inhumane, treatment*" (Francis, 2013, p.8). This example demonstrates an unprecedented lack of professionalism in the culture of an acute care organisation. The failings at Mid Staffordshire NHSFT were not picked up by external regulatory bodies, such as the Care Quality Commission (CQC), as they should have been. Instead they were uncovered when attention was paid to the true implications of the Trust's mortality rates and persistent complaints were made by a number of patients and relatives (Francis, 2013).

Following a public inquiry into the failings at Mid-Staffordshire NHSFT, the Francis report made 290 recommendations, which in turn led to a values-based recruitment (VBR) approach. For example, the Francis report recommended that the NMC should consider introducing an aptitude test that would assess aspirant nurses attitude toward caring, compassion and other essential values prior to them entering the profession (Francis, 2013). As with the situation at Mid Staffordshire NHSFT, a significant lack of professionalism was evident in the case of Winterbourne View.

1.1.2 Winterbourne View scandal

Winterbourne View was a private Treatment and Assessment service that was established to support adults with learning disabilities and autism; the care home is now closed. In

2011, a televised programme highlighted the abuse that occurred at Winterbourne View; staff members continually abused and mistreated patients and management failed to deal with these unprofessional practices (Flynn and Citarella, 2012). Despite an individual whistle blower, the abuses were not picked up by external health or regulatory bodies and instead continued (Department of Health, 2012). The abuses were eventually aired by an undercover journalist, as seen in the serious case review (Flynn and Citarella, 2012). In the case of Winterbourne View, patients were vulnerable and unable to protect themselves against this unprofessional and abusive behaviour.

Despite the case of Winterbourne View receiving significant attention in the media, there have been many further instances where staff have failed to protect patients and have instead mistreated them. Again, an undercover journalist in a recent Panorama episode demonstrated a *"toxic culture of humiliation, verbal abuse and bullying at the Edenfield Centre in Prestwich near Manchester"* (BBC, 2022). Less than two weeks later, a dispatches episode broadcast scenes at Essex Partnership University NHS Foundation Trust, where patients were dragged along the floor by staff, staff used unnecessary force against patients, and mocked patients whilst they were experiencing distress (see Channel Four Television Corporation, 2022); an independent inquiry is currently underway at Essex Partnership University NHS Foundation Trust, some of the findings of which are touched upon in chapters 6 and 7 of this thesis (see sections 6.5.5 and 7.2.3). Given these circumstances, it is even more concerning that complaints raised by psychiatric patients may fall on deaf ears, as was the case when doctors William Kerr and Michael Haslam sexually abused patients (Mannion et al., 2019). VBR is, therefore, especially important in mental health services, where patients are vulnerable and may be less able to safeguard themselves against abuse and exploitation.

1.1.3 The additional impact of the Covid-19 pandemic

On January 30, 2020, the World Health Organization declared the outbreak of a Public Health Emergency of International Concern (Gourret Baumgart et al., 2021). On March 11,

2020, a pandemic was declared – that being the Covid-19 coronavirus disease (Fond-Harmant, 2021). When thinking about the impact of Covid-19, aside from having to self-isolate, many people may recall the slides being aired on the daily news that reported the national mortality rates; there was a significant focus at the time on the nation's physical health. This physical health concern led to individuals being unable to see their friends or family, for fear of people contracting the virus. Not only did people choose not to see their friends and family, but the government ordered them not to do so; people would receive a fixed penalty notice and be fined for breaking these rules. Social distancing measures may have had an increased impact on mental health services in comparison to other settings due to the interpersonal nature of mental healthcare (see Kane et al., 2022). During these times, patients being cared for in mental health services were also unable to have friends or family visit, and caring often went on behind closed doors, thus increasing the potential for abuse and misconduct.

Alongside NHS colleagues, the Centre for Mental Health forecast an increase of almost 20% of the population in England needing new or additional mental health care as a result of the Covid-19 pandemic (O'Shea, 2020; 2021), thus increasing the workload of professionals that provided care in this setting. Healthcare professionals, including mental health workers are at increased risk of burnout, depression, or post-traumatic stress disorder as a result of the pandemic (see Gourret Baumgart et al., 2021). Thus, it is possible that staff were less likely to meet patients' mental health needs at this time (or thereafter). Whilst some patients may have difficulty protecting themselves against malpractice, carers can mitigate against this by speaking up and ensuring a patient's needs are acknowledged. Yet, as highlighted above, during the pandemic, patients could not receive visitors due to the overarching concern for peoples' physical health, thus resulting in an even greater need to recruit staff with the right knowledge and behaviours to deliver safe, effective and person-centred care.

1.2   Nuances in mental health services

Mental health services are in increasingly high demand; yet, they are in receipt of inadequate funding, an understaffed workforce, and insufficient training (British Medical Association, 2017). As the complexity of mental health care increases, professionalism becomes increasingly challenged. For instance, when patients are deprived of their liberty, the principle of patient autonomy, one of the three fundamental principles of medical professionalism becomes problematic (Project of the ABIM Foundation, 2002). There are common values across all healthcare specialties, for example, all must adhere to the NHS constitution. Nevertheless, professionalism is, to some extent, context dependent (Brendel et al., 2007; Brody and Doukas, 2014; Harris and Kurpius, 2014; Ikkos and Mace, 2009; Malhi, 2008; Rees and Knight, 2007; van de Camp et al., 2004; Wear and Kuczewski, 2004; Wise, 2008), and attention must therefore be paid to the specific elements of professionalism within each of the healthcare disciplines (van de Camp et al., 2004). In addition, there are patient and workforce issues prevalent in mental health services that increase the need to select suitable personnel to work in this setting.

1.2.1   Patient issues

Patients using mental health and learning disability services are especially vulnerable to becoming victims of unprofessional practice and misconduct. Patients may lack the mental capacity to make decisions in their own best interests and thus rely more on practitioners, and others, such as relatives, to make decisions for them. Indeed, the Independent Mental Capacity Advocate role was introduced as part of the Mental Capacity Act (2005) to ensure that the views of patients that lack capacity to make potentially life-changing decisions can be made by substitute decision makers (Redley et al., 2010). This dependence on clinicians leaves patients vulnerable to abuse and exploitation. Patients may be detained using the Deprivation of Liberty Safeguards procedure (Mental Capacity Act, 2005), or under part II of the Mental Health Act (1983). Such involuntary detention may be due to the severity of their mental health problem and/or learning disability.

Whilst, legally, patients can challenge their detention by taking this to a mental health review tribunal, they may choose not to do so. A study that sought to explore why most patients did not exercise their right to appeal their detention under Section 2 of the Mental Health Act 1983 observed that patients with a previous admission or educated to an A-level standard were more likely to appeal their detention; however, those with a diagnosis of depression or dementia were less likely to appeal (Bradley, 1995). Those that did not appeal also demonstrated less understanding of a booklet that described their rights. Patients may find it more difficult to represent their case given the mental health difficulties they are experiencing. Patients may be subject to coercive practices, such as physical restraint and seclusion, they may suffer from psychoses, or other issues that impact on their ability to have insight into their problems. This may make it difficult for them to communicate their difficulties to a mental health practitioner. Not only do patients experience difficulties because of their mental health problem and/or learning disability, but practitioners too may face their own challenges due to the nature of the patients' difficulties. For example, patients may be 'non-compliant' with treatment or may disengage with services making the delivery of care more challenging for professionals. Patients may also display 'behaviours that challenge, which can include verbal or even physical aggression. Patients may also complain, sometimes unfairly about their care.

1.2.2 Workforce issues

High levels of stress in patients has been found to partly account for high levels of psychological stress among healthcare professionals (Heponiemi et al., 2014). Nurses working in mental health departments have been observed to have higher levels of perceived stress than oncology nurses, nurses working in intensive care units, and nurses working in emergency rooms, and medical surgical wards (Masa'deh et al., 2017). Higher rates of depression, stress and burnout have also been found in psychiatrists compared to doctors from other healthcare specialties (Bressi et al., 2009; Deary et al., 1996; Firth-Cozens, 2007). Furthermore, high levels of burnout and stress have been observed among

occupational therapists (Brown and Pranger, 1992; Gupta et al., 2012), and social workers (Lloyd et al., 2002) compared to other occupational groups.

Historical evidence has suggested that medical students with high levels of psychological distress are more likely to pursue a career in psychiatry (Eron, 1955); psychiatrists also have been found to have higher levels of self-criticism as students (Firth-Cozens et al., 1999). A study with clinical psychology trainees found that 73% of individuals also reported clinically significant levels of psychological distress as reported on the General Health Questionnaire – 28 (Stafford-Brown and Pakenham, 2012). Women working in the mental health field have reported higher rates of physical abuse, sexual molestation, parental alcoholism, hospitalisation of a parent for mental illness, and death of a parent or sibling compared to other professionals, such as accountants (see Elliott and Guy, 1993).

Psychiatrists are more likely to receive disciplinary action than other healthcare specialties (Reich and Schatzberg, 2014), which is often due to having sexual relationships with patients (Morrison and Morrison, 2001). Psychiatrists often keep their problems to themselves (Rossler, 2012) and have been found to use more benzodiazepines than other healthcare specialties (Braquehais et al., 2014; Domenighetti et al. 1991). Prescription-type drug use has also been observed to be more prevalent among nurses that work in oncology, rehabilitation, and psychiatry, in comparison to other nursing specialties (Trinkoff and Storr, 1998). Trinkoff and Storr's study also observed mental health nurses, following oncology nurses, to have the highest past-year prevalence of substance misuse compared to nurses working in other healthcare specialties. Similarly, psychiatry residents also have been found to have the highest lifetime use of cocaine, LSD, and marijuana compared to other specialties (Firth-Cozens, 2007).

Psychiatrists have significantly high rates of suicide compared to doctors in general health medicine (Hawton et al., 2001). Nurses also have been reported to be at higher risk of suicide than the general population, with depression being a contributory factor (see a

review by Alderson et al., 2015). Of critical importance, however, is that the clinical symptoms of depression in doctors have been found to have negative implications for patient care (Firth-Cozens, 2001). Patient safety and wellbeing is paramount, yet in mental health services vulnerable patients are cared for by staff that are also vulnerable. Such factors have been worsened by the Covid-19 pandemic (see section 1.1.3). Fitness to practice cases often result in increased stress faced by professionals, however, the effects of high stress and poor wellbeing on professionals can also lead to fitness to practice situations (Coffey, 2020). Furthermore, increased vacancies, alongside increased demand for mental health services (British Medical Association, 2022), is likely to result in more pressure on current employees and, subsequently, higher levels of stress in staff. It is therefore important that we can assess the professionalism of staff applying for mental health staff roles, in order to enhance the health-related outcomes of patients using these services.

1.3 Mandating of VBR

Following the failings of Mid-Staffordshire NHSFT, the Francis report highlighted that staff being recruited to the profession should evidence the appropriate values, attitudes and behaviours (Francis, 2013). Similar recommendations have also been made in the Cavendish Review (Department of Health, 2013a). Francis noted that there are commonalities across healthcare professions and that all healthcare staff should apply the Nolan principles (Francis, 2013). In 2012, a new strategy 'Compassion in Practice' set out to promote high quality, compassionate care among nurses, midwives, and care staff. This strategy built on the values set out in the NHS Constitution and incorporated 6Cs, including; care, compassion, courage, communication, commitment and competence (Cummings and Bennett, 2012). In 2014, plans were made to roll out the 6Cs to all healthcare professions, from Hospital porters to trust chief executives (Stephenson, 2014).

To develop an effective workforce, the Health Education England (HEE) Mandate (April 2014 to March 2015), superseding the Health Education England (HEE) Mandate (April 2013 to March 2015; Department of Health, 2013b), identified the need to develop evidence-based

approaches to recruitment and selection for training programmes based not just on technical and academic skills, but also on values and behaviours (Department of Health., 2014). It is vital that individuals with the right values are selected, not just to any educational course or training, but also to healthcare roles in general.

1.4   What is VBR?

Processes were in place to assess desirable non-academic (personal) attributes in healthcare applicants before VBR was mandated (Groothuizen et al., 2018). However, following the publication of the Francis Report, VBR was implemented in the NHS to help recruit and select staff whose personal values and behaviours align with the values outlined in the NHS Constitution (Francis, 2013). It is argued that staff need to commit to the constitution's core values and principles, because unless there is a large enough workforce with the right skills, values and behaviours to deliver new care models, these standards would simply not become a reality (Health Education England, 2014).

Having critically evaluated documents relating to VBR, specifically in nurse education, it was concluded that one may question whether the values listed in the NHS constitution are the appropriate values to underpin VBR (Groothuizen et al., 2018). The authors also highlighted that as organisations develop their values in line with the NHS constitution, the specified values may '*get lost in translation*', resulting in the recruitment of individuals based on unrelated criteria (Groothuizen et al., 2018, p.1072). Despite the importance of values for healthcare delivery, there is extensive variation in the approaches and processes used for assessing values (Spilsbury et al., 2022), and there is currently no singular assessment tool used for the recruitment and selection of staff in mental health services.

1.5   Implications of VBR

There is mixed evidence regarding the impact of VBR on patient care and service delivery. Consilium Research and Consultancy Ltd. reported on a study that was conducted with social care organisations, which assessed the outcomes of a values-based approach to

recruitment and retention. A values-based approach to recruitment and retention incorporates not just recruiting staff with the right values and attitudes, but also supporting staff to develop their skills and knowledge whilst working in a caring role (Consilium Research and Consultancy Ltd, 2016). Using an online survey and qualitative interviews to assess the impact of this values-based approach, the study found that 72% of employers reported that staff employed and supported using the 'values-based' approach perform better than those recruited through 'traditional' methods. In addition, more than 70% of employers found that staff recruited using this approach possess stronger care values (e.g. respect) in comparison to those selected using 'traditional' methods (Consilium Research and Consultancy Ltd, 2016). The values-based recruitment and retention approach not only impacted the social care staff being recruited, but it also had an impact on costs. A financial analysis found an estimated return on investment of 22.8% for implementing this approach (Consilium Research and Consultancy Ltd, 2016); also, where additional resources were used, employers reported that the benefits outweighed the costs.

In contrast to the study undertaken by Consilium Research and Consultancy Ltd., a report detailing the findings of an evaluation of VBR, which was commissioned and funded by the Policy Research Programme, documented that the effect of VBR is not clear (Spilsbury et al., 2022). The authors concluded that their findings were not sufficient to support the assumption that VBR helps select individuals with values that are more aligned to those of the NHS, and instead the culture of the organisation and workplace practices were perceived to be more influential. Nevertheless, that does not mean that VBR is an unnecessary feat. As highlighted by Spilsbury et al. (2022), VBR processes may signal the organisation's values to potential employees, and therefore influence their decision as to whether to apply for a vacancy or not.

1.6   How can VBR be done effectively?

Whilst values influence behaviour, people do not always behave in ways that are consistent with their values; other factors, including knowledge, skills, experience and personality also

play a role (Health Education England, 2014b). The culture of an organisation and workplace practices are also highly influential (Spilsbury, 2022). Parks and Guay (2009) differentiated between values as preferences and values as principles; values as preferences are primarily attitudes, whereas values as principles regard how one ought to behave and are believed to directly impact on an individual's motivation.

There are numerous selection methods currently used for VBR and the feasibility and validity of these approaches has been reported previously (see Health Education England, 2014b; Patterson et al., 2016a; 2016b). Given the relevance of the different selection procedures to the choices underlying the current research project (i.e., to develop an SJT), several alternative selection methods will be summarised here.

1.7 VBR measures

A literature review conducted by Patterson and colleagues split VBR methods into two separate categories; these being shortlisting and final stage selection methods (Patterson et al., 2016a; 2016b). Shortlisting methods include personal statements, references, SJTs and personality tests and are typically used to screen out applicants. Final stage selection methods include the use of structured interviews, group interviews, and multiple-mini-interviews (MMIs), among other approaches. The author will address some of these methods before providing a rationale for the current doctoral project.

1.7.1 Personal statements

Individuals are frequently asked to provide a personal statement when applying for a training course or vacancy; however, evidence regarding the effectiveness of personal statements is mixed (Patterson et al., 2016a). The amount of content in a personal statement has been found to predict better clinical performance (Ferguson et al., 2003), yet personal statements have been shown to have low reliability (Patterson et al., 2016b) and the information provided is subjective in nature (Patterson et al., 2016a). Whilst personal statements are typically accepted by candidates as a means of selection (Patterson et al.,

2016b), individuals may provide a false account of themselves and respond in a way that they believe is expected of them (White et al., 2012). Further concern is raised that personal statements may cloud the judgement of those making decisions about the recruitment of staff (Patterson et al., 2016a). Individuals can be taught how to write better personal statements or may have someone else write a personal statement for them (Patterson et al., 2016a), resulting in an unfair selection process.

1.7.2 References

It has been claimed that references have no predictive value (Poole et al., 2009). Despite this, references, like personal statements, are used widely in selection and recruitment. Recruiters search references for positive comments or red flags (Wagoner, 2006), yet evidence suggests that they are biased and fail to differentiate between applicants (Stedman et al., 2009). There is wide agreement that references are an unreliable method of selection (Patterson et al., 2016a).

1.7.3 Personality tests

It is not disputed that personality influences one's behaviours in the workplace. Conscientiousness, in comparison to the other Big Five personality traits, has been frequently found to be most closely related to individuals' success in both educational and work settings (see Lievens et al., 2009); conscientious individuals have been found to be more likely to set goals, commit to goals, and perform better. Conscientiousness is also directly related to supervisor ratings of performance (Barrick et al., 1993).

Nevertheless, the use of personality assessments for recruitment has been widely debated (see Patterson et al., 2016b). Personality tests are cheap and efficient to use, score and interpret (Finn et al., 2018); however, when assessed for educational performance, these self-report measures have been found to have very little predictive validity (Finn et al., 2018). Personality tests have been shown to have incremental validity over cognitive methods when used for medical school selection (Patterson et al., 2016a); similar to personal statements however, there are concerns that individuals fake their scores on

personality measures to portray themselves more positively (Birkeland et al., 2006); this social desirability response is not always conscious faking.

1.7.4 Structured interviews

Interviews are commonly used and accepted as part of the selection process (see Health Education England, 2014b). Structured interviews predict job performance and have increased validity in comparison to unstructured interviews (McDaniel et al., 1994). Structured interviews have also been found to have incremental validity over and above cognitive ability and conscientiousness (Cortina et al., 2000); yet, whilst they have high validity, they are resource intensive and can be costly to construct and use (Health Education England, 2014b; Schmidt and Hunter, 1998).

Research that explored the incremental validity of face to face selection centres over paper and pencil tests for GP selection, which include an SJT and clinical problem solving test, found that, when age and sex were not controlled for, the paper and pencil tests accounted for 55.3% and 36.7% of the variance in Applied Knowledge Test (AKT) and Clinical Skills Assessment (CSA) scores, respectively (Patterson et al., 2015). Selection centres had statistically significant incremental value over the paper and pencil tests, but only by 0.4% for the AKT and 4.3% for the CSA. A further evaluation of GP specialty selection obtained similar results (Davison et al., 2016). Noting the limited amount of additional prediction provided, the authors commented "*whether the additional predictive benefit provided by [a selection centre] is cost-effective is open to dispute*" (Davison et al., 2016, p.158). Research regarding the selection of medical students has also found interviews to be susceptible to interviewer bias in that interviewers ranked candidates more favourably if they shared certain personality preferences (Quintero et al., 2009).

1.7.4.1 *Multiple mini-interviews (MMI)*

As the term indicates, a multiple mini-interview includes a series of short interviews or interpersonally orientated tasks, which are performed at different test stations consecutively. Research in medical education has found MMIs to be a reliable selection

procedure that is accepted by both candidates and assessors (Dore et al., 2010). MMIs have also been found to predict performance (Eva et al., 2009). However, whilst MMIs show improved reliability and validity in comparison to traditional interview approaches (Patterson et al., 2016a), there is limited research exploring what construct or constructs they actually measure (Health Education England, 2014b). Additionally, MMIs are costly to construct and use, as the stations require a thorough job analysis (Health Education England, 2014b) and the recruitment of multiple assessors, which are usually academic staff.

1.7.5 Situational judgement tests (SJTs)

SJTs can be conceptualised as 'low fidelity simulations' (Motowidlo et al., 1990), whereby test-takers are provided verbal or written descriptions of hypothetical work-related scenarios and are subsequently asked to provide a response regarding the most appropriate course of action; this could be in a forced choice or Likert scale format. Respondents can also be asked to provide an open-ended response where they comment on certain aspects of the scenario, for example, why the scenario provided is ethically dubious. An example SJT item is provided in chapter 2. SJTs may be able to evaluate a broader range of constructs, in comparison to MMIs, due to the amount of items that can be used in an SJT assessment (Patterson et al., 2018). However, SJTs and MMIs are often considered to be complementary and are often used at different stages of the selection process (Patterson et al., 2018).

SJTs can be delivered at scale as they can be computerised or implemented in online format and machine marked (Patterson et al., 2018); SJTs are therefore less costly than 'high fidelity simulations', involving live actors or patients, and are relatively easy to administer. This has led to the increasing popularity in their use within personnel selection generally (Weekley and Ployhart, 2006). Most commonly, within the recruitment process SJT format tests are used as 'screen-out' selection procedures, whereby lower scoring candidates do not progress to more resource intensive selection assessments such as multiple mini-interviews (MMIs; Patterson et al., 2018).

Whilst challenges have been noted regarding the development of SJT items, it is highlighted that some of these are due to a lack of resources that organisations have available to invest in the development, refinement, piloting and evaluation of such selection assessments (Health Education England, 2014a). It can be both time consuming and resource intensive to develop a SJT; nevertheless, it is widely accepted that SJTs, when properly constructed are a valid and cost-effective selection method (Patterson et al., 2016a). SJT scores can predict job performance (McDaniel et al., 2001), which is usually rated by respondents' supervisors in SJT validation studies.

Another strength of SJT scores, is that they have, at times, been shown to hold incremental validity over cognitive ability, and the Big 5, as well as a composite of the two (Mcdaniel et al., 2007). An SJT format assessment could therefore feasibly be an effective approach to selecting a more compassionate and caring workforce, with the correct interpersonal skills and values to deliver mental healthcare.

1.8   Focus of the doctoral project

This introductory chapter highlights the importance of having VBR in mental health services. As evidenced in a report commissioned by Health Education England (2014a), organisations in the UK currently use many alternative approaches to VBR and there is no one agreed form of assessment. MMIs are shown to have increased validity in comparison to other selection measures, yet they require multiple assessors, are resource intensive and can be costly to develop and use. The reliability of MMIs increases with the number of stations (Dore et al., 2010); more raters are thus required to generate robust reliability values. Scores resulting from SJTs also demonstrate good predictive validity, as well as test-retest reliability, yet SJTs can be delivered online and machine marked meaning that they are much less expensive to deliver.

1.8.1 The value of professionalism

This thesis places an emphasis on the need for VBR in mental health services; yet, there are concerns regarding the validity and effectiveness of this approach (Groothuizen et al., 2018). There is limited understanding, for instance, as to why the values listed in the NHS constitution were those chosen to underpin VBR, values also are subjective (Rankin, 2013), and it is therefore necessary to determine what it is we aim to assess for (see Groothuizen et al., 2018). Whilst one would hope that mental health workers possess suitable values to deliver safe and effective mental health care, values are not always enacted. Values may also conflict, meaning that one value will be prioritised over the other and prior research has demonstrated values attrition over time (Groothuizen, 2020).

Job analysis studies evidence the importance of non-academic (personal) attributes for successful performance in multiple healthcare roles (Patterson et al., 2016d). Unlike cognitive attributes, non-academic attributes include personal qualities such as empathy or emotional intelligence, however, it is acknowledged that even non-academic attributes will have cognitive elements to them (e.g. 'situational cognition'; Finn et al., 2018). An SJT could be developed to assess desirable non-academic attributes, which would, in turn, facilitate the selection of suitable staff for mental health services. Interpersonal skills are particularly important for staff working in mental health services and the patient-practitioner relationship is viewed as fundamental to professionalism (Groves and Kerson, 2011; Schreiber et al., 2016). Given that staff are judged on their 'fitness to practice' and unprofessional practice, the opportunity to objectively assess professionalism is deemed of great importance. This doctoral project thus sought to assess one's professionalism, or knowledge of professionalism within this context, as opposed to assessing one's values. This was deemed a more suitable approach, because values are not easily operationalised, are unlikely to be unidimensional in most cases, and are not easily amenable to quantification.

SJTs are relatively easy tests and can differentiate better between low scoring applicants; they are therefore likely to identify applicants that are at risk of poor practise and more

likely to face disciplinary action during their career. Whilst there is currently no single 'gold standard' assessment used for VBR, SJTs are the most practical way of evaluating one's procedural knowledge of professionalism, within a mental health services context, given a trade-off between cost and validity. Thus, the primary aim of this doctoral project is to develop and validate an SJT that will assess staff members' knowledge of professionalism for a mental health services context. Used alongside other selection procedures, such as MMIs, an SJT would facilitate the VBR process, thus promoting the selection of a more compassionate and competent workforce that is skilled to deliver mental healthcare.

1.9   Aims and objectives

The overall aim of the research project was to 'develop and validate an SJT that would assess staff members' knowledge of professionalism for a mental health services context. Objectives of this project are as follows.

1.9.1   Objective 1

To define 'professionalism' for a mental health services context.

1.9.2   Objective 2

To develop a pool of SJT items that could validly evaluate 'knowledge of professionalism' in a mental health setting.

1.9.3   Objective 3

To develop and validate an SJT for personnel selection in mental health services.

1.9.4   Objective 4

To develop a bespoke workplace behaviours rating tool.

Please note that there were some departures from the original research protocol, which are addressed in chapter 3.

1.10 Thesis structure

Having provided an overview of the literature resulting in the current research project, this thesis continues with a topic-based structure. Over the course of the doctoral project, two articles were submitted to peer-reviewed journals for publication. Chapter outlines, and subsequent publications are documented here.

Chapter 2 supports objectives 2 and 3. Chapter 2 discusses the history and current use of SJTs for personnel selection. The strengths and limitations of this selection method are also addressed.

Chapter 3 supports objectives 1, 2 and 3. Chapter 3 highlights the author's position regarding the research project; the methodology underlying this work is discussed.

Chapter 4 supports objective 1. Chapter 4 reports on a systematic literature review that was conducted to derive an operational definition of professionalism for a mental health services context. The work performed for this aspect of the project has been compiled in the following publication:

> Aylott, L., Tiffin, PA, Saad, M., Llewellyn, AR. & Finn, GM. 2019. Defining professionalism for mental health services: a rapid systematic review. *Journal of Mental Health*, 28, 546-565.

Chapter 5 supports objectives 1 and 2. Chapter 5 reports on a qualitative study that sought to 1) establish the domains of professionalism according to key stakeholders and 2) develop a pool of SJT items that covered the concept of professionalism. Findings from the qualitative study have been reported in the following publication:

> Aylott, L., Tiffin, PA., Brown, S. & Finn, GM. 2022. Great expectations: Views and perceptions of professionalism amongst mental health services staff, patients and carers. *Journal of Mental Health*, 31, 139-146.

Chapter 6 supports objective 3. Chapter 6 reports on the development, pilot, and validation of the SJT.

Chapter 7 provides an overall discussion regarding the content of the thesis. Recommendations for policy and practice are proposed.

Chapter 2: SJT Literature Review

Chapter 1 sets the context for this doctoral project, and thus highlights the importance of Values-Based Recruitment in mental health services. The reader was introduced to various selection procedures that are currently being used; the advantages and disadvantages of each approach were highlighted prior to the author providing a rationale for choosing to develop an SJT for selection of the mental healthcare workforce. This chapter provides a historical perspective regarding the use SJTs for personnel selection, noting when they were first developed and how they are being used in more recent times. The theory underlying SJTs, as well as evidence demonstrating their reliability and validity is explored and described.

2.1 The use of SJTs for personnel selection

SJTs have been used for personnel selection for many years, the earliest example of which depends on how the SJT is defined (Weekley and Ployhart, 2006). Similarities can be observed between SJTs and other selection procedures, such as the situational interview (Latham et al., 1980), work samples (Asher and Sciarrino, 1974), and other situational-based assessment centre exercises (see Weekley and Ployhart, 2006). The author will hereby provide an initial description of an SJT prior to discussing the history of their use in personnel selection.

2.1.1 An introduction to SJTs

SJTs are designed to assess peoples' judgement in work settings by presenting test-takers with hypothetical job-related situations (item stems), alongside various responses (item responses; McDaniel and Nguyen, 2001). Responses are typically provided in a multiple choice format and candidates are asked to choose the best or worst option, rank options, or rate the appropriateness or effectiveness of each option (Weekley and Ployhart, 2006). An example SJT item with a rating format is provided in Table 2.1.

Table 2.1 An example SJT item with a rating format (replicated from UCAT Situational Judgement Test Question Bank, UKCAT Consortium, 2019)

Merryn, a medical student, is walking through a hospital corridor when she sees a small pool of liquid in the middle of the floor. She is already five minutes late for the start of her shift, however does not want to leave without cleaning it up. Merryn must decide how to deal with the situation.

How **important** to take into account are the following considerations for **Merryn** when deciding how to respond to the situation?

The potential risk of infection that the liquid presents to patients in the hospital

- Very important
- Important
- Of minor importance
- Not important at all

2.1.2   A brief history on the usage of SJTs

Having conducted a review on the history of SJTs, McDaniel et al. (2001), highlight that one of the first widely used tests to assess judgement was the George Washington Social Intelligence Test in the 1920s. The George Washington Social Intelligence Test contained various parts and included a section where test-takers were provided 30 scenarios that incorporated difficult social problems; test-takers were subsequently asked to exercise their judgement by determining the best response for each situation (Moss, 1926). SJTs have been used to assess the judgement of soldiers throughout World War II (McDaniel et al., 2001; Tiffin et al., 2020), and were used to assess supervisory and managerial potential between the 1940s and 1960s (Weekley and Ployhart, 2006). It was in the 1990s that SJTs experienced a sudden surge in popularity for their use in personnel selection (Corstjens et al., 2017). Their increased usage was likely stimulated by Sternberg and colleagues describing the importance of 'tacit knowledge' to job performance (Weekley and Ployhart, 2006). Moreover, Motowidlo's reconceptualisation of SJTs as 'low fidelity simulations' no doubt also provided impetus to their adoption for staff selection purposes (Motowidlo et al., 1990). As noted by Motowidlo et al. (1990, p.640), *fidelity decreases as stimulus*

materials and responses become less and less exact approximations of actual job stimuli and responses, hence the term low fidelity simulation. SJTs are also referred to as low fidelity simulations because they typically ask what someone would do, but do not assess the actual behaviour observed (Lievens, 2017b). The faithfulness of the tests to reality can vary. For example, video and multimedia SJTs increase their fidelity through the rich portrayal of the situation in question (Lievens et al., 2008); high fidelity simulations are a very realistic representation of a task that provides applicants an opportunity to respond almost exactly as they would in an actual job situation (Motowidlo et al., 1990).

SJTs have been purported to assess various constructs, including emotional intelligence (Sharma et al., 2013), integrity (de Meijer et al., 2010) and personal initiative (Bledow and Frese, 2009) among other traits. Some would suggest that these are 'construct-driven SJTs', however, which differ from traditional SJTs that are often used for personnel selection. Traditional SJTs are more atheoretical and are typically derived through blueprinting, generating critical incidents, and creating a scoring key with subject matter experts (SMEs). Whilst SJT domains have labels, the actual construct being assessed is often not made explicit (Christian et al., 2010). SJTs have been used for selection across occupations (e.g. the recruitment of police officers; de Meijer et al., 2010, Kanning et al., 2006), and for the past 10 years have been used for selection into medicine. SJT scores have predicted performance across a range of medical professions from the early years to general practice (Patterson et al., 2016d). For example, an SJT is currently being used as one part of the selection to the Foundation Programme in the United Kingdom (UK) to target the key professional attributes required of clinicians; a study evaluating this SJT observed a relationship between SJT scores and the odds of successful completion of the Foundation Programme (Smith and Tiffin, 2018).

2.2   The theoretical underpinnings of SJTs for personnel selection

SJTs are purported to assess a variety of non-academic (personal) attributes, as well as procedural knowledge; the latter is peoples' knowledge about what is effective and

ineffective behaviour at work (Motowidlo et al., 2006b). It is worth noting that procedural knowledge and declarative knowledge differ; whereas declarative knowledge relates to things that individuals are aware they know and can describe to others (e.g. 3 + 4 = 7), procedural knowledge is knowledge that individuals display in their behaviour, but that they are not conscious of (Anderson and Lebiere, 2014). Procedural knowledge is believed to be a combination of specific job knowledge and implicit trait policies (ITPs; Motowidlo and Beier, 2010; see below). Such procedural knowledge is postulated as having a causal effect on job performance, because it is assumed that people who know what actions are effective are more likely to perform effectively in their role (Motowidlo et al., 2006b). There are two main underlying theories that justify the use of SJTs for personnel selection (Patterson et al., 2016d), these being behavioural consistency theory and theory relating to ITPs; the latter requires more discussion due to its complexity.

2.2.1   Behavioural consistency

The concept of behavioural consistency theory is relatively simple; according to behavioural consistency theory, an individual's current behaviour can predict their future behaviour (Wernimont and Campbell, 1968). Based on this tenet, Motowidlo et al. (1990) noted that simulations, which are designed to closely resemble actual work conditions, should be good indicators of future job performance. Therefore, one would expect higher fidelity tests to have increased predictive validity, due to their point-to-point correspondence (Asher and Sciarrino, 1974). As expected, research findings demonstrate that individual's are more consistent across situations that are similar in comparison to situations that are dissimilar (Funder, 2006). A recent meta-analysis that explored the criterion-related validity of personnel selection measures found that higher levels of consistency between the predictor and criterion resulted in higher predictive validity (Arnold, 2020). Whilst SJTs may assess individuals' intentions or knowledge regarding certain behaviours in hypothetical situations, it is worthy to note that individuals' intentions and behaviours may differ, potentially due to an individuals motivation and/or skills (Arnold, 2020).

2.2.2   Implicit trait policies

Motowidlo, Hooper and Jackson (2006b) developed a theory to explain how SJTs assess procedural knowledge. SJTs are believed to assess implicit trait policies (ITPs), which are conceptualised as the beliefs people hold with regard to how effective various levels of trait expression are (Motowidlo et al., 2006a; 2006b). Motowidlo et al. (2016, p.332) later developed an SJT to assess prosocial implicit trait policy; essentially, one's knowledge *'about the utility of prosocial expressions in social interactions'*. The term prosocial has been defined as '*denoting or exhibiting behavior that benefits one or more other people'* (American Psychological Association, 2020); thus, 'prosocial ITPs' related to healthcare would include beliefs regarding the effectiveness of expressing traits, such as compassion, respect, and caring for patients (Patterson et al., 2016b).

ITP theory partly derives from the 'dispositional fit' argument, which postulates that peoples' beliefs regarding the effectiveness of different behaviours develop in line with their own basic traits (Motowidlo et al., 2006a). For example, individuals that are highly agreeable are more likely than individuals that are disagreeable to believe that agreeable actions are highly effective. Therefore, if a work situation 'requires agreeable action,' agreeable individuals will have more accurate knowledge about how to behave effectively in that situation. ITPs are therefore believed to mediate the relationship between one's personality traits and their procedural knowledge (Motowidlo et al., 2006b). Whilst personality traits are believed to have causal effects on an individual's ITPs, and in turn affect their procedural knowledge, experience also influences ITPs over time (Motowidlo et al., 2006b). Through exposure to various work situations, an agreeable individual can learn, for example, that in certain circumstances disagreeable actions may be more effective regardless of their own standing on that trait (Motowidlo et al., 2006a; 2006b). Thus, ITPs are not entirely dependent on personality traits, and instead, can also be influenced by experience on the job. Henceforth, if agreeable behaviour is required in a given situation, disagreeable individuals are still likely to judge that agreeable behaviour would be more effective than

disagreeable behaviour; yet, in contrast to agreeable individuals, disagreeable individuals may judge agreeable behaviour to be only *slightly* more effective than disagreeable behaviour (Motowidlo et al., 2006a).

2.2.3   What do SJTs assess in this context?

SJTs are popular, because they provide contextualised information and enable the assessment of interpersonally-oriented skills among a large population early in the selection process (Lievens et al., 2008). It is proposed that SJTs are best considered a measurement method or methodology, as opposed to a single type of assessment, as they differ in their development and purport to assess a variety of constructs (Christian et al., 2010; Patterson and Driver, 2018). However, SJTs have sometimes been criticised as they lack a measurement model and it is not always clear what is actually being assessed (e.g. Tiffin et al., 2020). Results from a meta-analysis highlighted that a third of SJTs reported on in the literature either did not document the constructs they assessed, or they provided a composite score for an SJT that assessed multiple constructs (Christian et al., 2010). Where the constructs were documented, the meta-analysis found that SJT scenarios are most often related to the construct domain of applied social skills. Authors of the meta-analysis urged researchers to '*maintain the distinction between methods (e.g., SJTs) and constructs (e.g., leadership skills)*' (Christian et al., 2010, p.107) by clearly reporting on the specific constructs that SJTs assess. Work-related behaviours require the use of multiple skills and knowledge however, and it is therefore common for SJT scenarios to assess various constructs simultaneously (Christian et al., 2010).

Many SJTs assess heterogeneous composites, however some researchers set out to develop SJTs that assess specific constructs, which, as noted above, have been referred to as 'construct-driven SJTs' (Guenole et al., 2017; Tiffin et al., 2020). Motowidlo (2006a), for example, maintained that SJTs can be built to assess implicit trait policies for targeted traits, such as agreeableness. In contrast to traditional SJTs, which assess multiple constructs simultaneously, construct-driven SJTs present a test-taker with a trait activating situation,

and a set of response options that lie on a continuum to depict varying trait levels (Lievens, 2017a). A detailed overview of SJTs as assessments of context-dependent knowledge versus context-independent knowledge is provided in a paper by Krumm et al. (2015); Krumm et al., view the two theories as sitting on either end of a continuum with most SJT items sitting somewhere in between. Based on the tenet of behavioural consistency theory, the current programme of work sought to develop a traditional SJT, which would present a test-taker with a set of situations that they would encounter on the job. It was hoped that this would assess the test-taker's procedural knowledge and would, in turn, predict their future job performance.

2.3 The development process

There are key differences in the development and design of traditional SJTs, that are typically used for personnel selection, as opposed to construct-driven SJTs (Lievens, 2017a; Tiffin et al., 2020). For example, whilst subject matter experts are typically involved in identifying critical incidents for SJT item content in traditional SJTs, psychologists are typically in charge of developing item stems for construct-driven SJTs so that the item stem is appropriate for eliciting the trait of interest. A breakdown of the differences in development between traditional SJTs and construct-driven SJTs can be viewed in Figure 2.1. Not only do development options differ between traditional SJTs and construct-driven SJTs, but there are differences among traditional SJTs themselves, and likewise, between construct-driven SJTs. Furthermore, the construct and criterion-related validity of SJT scores are dependent on such characteristics, including; the mode of presentation, the level of fidelity, the response instructions, the scoring method, the scenario content, and the cognitive complexity of items (Arthur Jr et al., 2014, McDaniel et al., 2013).

2.3.1 Format

Various aspects of SJT design can vary, including the scenario content, response instructions, response formats and approaches to scoring (Patterson et al., 2016d). In fact, with regards to SJT development, Campion et al. highlight 12 attributes that differ, which include;

situation and response development, key development, scoring methods, scenario presentation, stimulus medium, response medium, response format, instruction format, context, constructs assessed, research design, purpose of study, number of items, sample size, and number of dimensions (Campion et al., 2014).

Traditional approach

- Define test specification by identifying the personal qualities deemed essential or important to effective workplace performance. This usually involves interviews or focus groups with subject matter experts (SMEs)

- Elicit "critical incidents" that can serve as potential scenarios in the test, usually via focus groups with SMEs and other stakeholders

- Decide on a response format and create items relating to the scenarios

- Involve SMEs in designing a provisional scoring key, based on expert consensus

- Trial pilot items in a sample of test-takers, similar to those for whom the test is intended as a selection measure. Ideally, some data relating to validity should also be obtained at this stage

- Review item construction, inclusion and scoring in light of psychometric analysis of pilot response data

- Pilot final version in the selection process before being implemented as part of the actual decision-making process

- Conduct regular, general, psychometric review and refresh test content, building up an 'item bank' of questions with known properties

Construct-driven approach

- Decide which construct(s) are to be captured by the test on the basis of theory and prior research evidence

- Scenarios, created primarily by psychologists, are developed that are considered likely to activate the relevant trait or construct. Stem construction may also be informed by "taxonomies of situations", which categorise the scenarios

- Psychologists create response options aiming to tap into a single trait. Format usually in the form of a rating scale to capture differing degrees of trait

- Provisional scoring system created. This can initially be based on theory, with increasing scores indicating higher levels of relevant trait. However, consensus ("wisdom of the crowd") and expert scoring may be alternatives

- Trial pilot items in a sample of test-takers, similar to those for whom test intended as a selection measure. Ideally some data relating to validity should also be obtained at this stage

- Review item and scale construction, inclusion and scoring in light of psychometric analysis of pilot response data. Emphasis is on achieving high levels of reliability and/or desirable test information properties

- Pilot final version before being used as part of selection or for research purposes

Figure 2.1 The traditional and construct-driven approaches to developing situational judgement tests for personnel selection (replicated from Tiffin et al., 2020a)

2.3.1.1 *Mode of presentation and level of fidelity*

Prior to the recent resurgence of SJTs, Motowidlo and colleagues referred to the 'low fidelity' simulation (Motowidlo et al., 1990); a paper and pencil test whereby test-takers are

provided a hypothetical scenario and several response options to choose from. However, SJT scenarios can be presented in a variety of formats, which range from text to avatar-based (Tiffin et al., 2020). The format of the scenarios and response options influence the fidelity of the SJT and, in turn, its resulting predictive validity. Theoretically, one would expect higher fidelity tests to have increased predictive validity, due to their point-to-point correspondence (Asher and Sciarrino, 1974). This theory is supported by a study performed by Christian et al. (2010), who found that scores on video-based SJTs were more strongly associated with performance than paper-and-pencil SJTs across varying construct domains. Nevertheless, despite some alternative formats having increased predictive validity, Motowidlo et al. (1990) found that in a sample of approximately 120 management incumbents, scores on a low-fidelity SJT-like paper and pencil test correlated from .28 (p < .01) to .37 (p < .01) with supervisory ratings of performance. Patterson et al. (2017), argue that when assessing for non-academic outcomes in medical and dental students, a text-based SJT may be equally as effective as a video-based SJT. Whilst video-format SJTs have demonstrated increased predictive validity compared to lower fidelity simulations, a paper and pencil format assessment can be a valid predictor of job performance that is developed at much less expense than an animated or video format.

Despite the increased use of SJTs more recently, there is little consensus in the literature as to how SJTs should best be developed, scaled, or scored (Hashmi, 2018). Nonetheless, all SJTs possess common features, such as having an item stem, response options, and response instructions.

2.3.1.2 *Item stems*

Every SJT item poses the test taker with a scenario or job-related dilemma, which is termed the item stem. Weekley et al. (2006) note that item stems differ in four key ways; the source used to develop the item stems can differ, item stems can change in both complexity, and fidelity, and the content of items stems also may vary. SJT item stems must provide enough information about a scenario for a test-taker to be able to decide how it is best to act,

however, too much information can result in test-taker fatigue (Brooks and Highhouse, 2006). A test developer must therefore consider the level and complexity of information contained within each item stem.

The most common approach to developing SJT scenarios is by utilising the critical incident technique (Flanagan, 1954), thus utilising an inductive approach. This incorporates SMEs being asked to recall incidents that they have encountered in the workplace. Another inductive approach for developing item stems would be to obtain critical incidents using archival records (Hashmi, 2018). To develop item stems deductively, a theoretical model can be utilised instead (Hashmi, 2018). For example, a literature review could be conducted to identify the key attributes required of a role. Finally, item stems can be developed using a joint inductive and deductive approach, thus benefitting from the strengths of each approach (De Leng, 2019).

2.3.1.3 *Response options*

Alongside the provision of a job-related dilemma (item stem), test-takers are typically shown various options as to how one may respond to the situation in question; these can be referred to as response options. Response options are typically presented in a written format, even when the item stem is displayed in a video or alternative format (Hashmi, 2018). Response options can be developed at the same time as item stems, or they can be developed after, with a different group of SMEs. When developing construct-driven SJTs, that are intended to assess specific personality traits, response options are typically written by psychologists in order to ensure that the various options represent different levels of the targeted traits (Lievens, 2017b). Like item stems, response options can also be developed deductively, using a theoretical model.

2.3.1.4 *Response instructions*

There are typically two forms of response instructions used in an SJT; these include, knowledge-based instructions that ask what one 'should' do and behavioural-tendency instructions that ask what one 'would' do (Mcdaniel et al., 2007). A sample of 372 medical

school applicants perceived behavioural-tendency instructions as easier to cheat than knowledge-based instructions (De Leng et al., 2018). Whilst SJTs are believed to assess procedural knowledge, this assessment might be diluted by response instructions that ask what one would do (Motowidlo et al., 2006b). This is because, despite the instruction, some respondents, particularly in high stakes selection, might respond with what they think they ought to do as opposed to what they think they would do in that situation, thus demonstrating social desirability bias. Behavioural-tendency instructions are more likely to assess one's typical performance, whereas knowledge-based instructions are more likely to assess one's maximal performance. As one may expect, responses to knowledge-based SJTs correlate more highly with cognitive ability and scores on knowledge tests, whereas responses to behavioural-tendency instructions are more highly associated with personality traits (Mcdaniel et al., 2007).

Given the potential for faking and social desirability bias to influence a test-takers SJT scores (Peeters and Lievens, 2005), knowledge-based instructions may be preferred to a behavioural-tendency format, especially when used for personnel selection. After all, knowing what to do is also a pre-requisite to behaving appropriately on the job. It is worthy to note also that ethnic minority medical school applicants rate should-do instructions more favourably that would-do instructions, thus demonstrating a preference towards a knowledge-based format SJT (De Leng et al., 2018). Response instructions can be provided in a written or video based format, which can also influence candidate reactions (Patterson et al., 2016d).

Job applicants may be provided multiple SJT responses at a time, or they can be provided one response only (a single-response SJT format). Other response instruction formats can include a free-text response format (Webster et al., 2020), or test-takers being asked to rank options in order, or alternatively to select the best and worst options. Although these response instructions can result in valid assessment tools, responses are harder to model, and item and reliability analysis of the instrument is difficult because the responses are

partially ipsative (Hicks, 1970; McDaniel and Nguyen, 2001). A rating scale format reduces the problem of partial ipsativity and makes item analysis easier. If a rating format is utilised however, test-developers must consider extreme response bias. That is, a preference for test-takers to choose categories at extreme ends of the rating scale (e.g., 'very appropriate' as opposed to 'appropriate'). Extreme response tendency influences a test-takers SJT scores in that their resulting score is the product of both their extreme response tendency and the level they possess of the construct being assessed. For this reason, one may prefer to use fewer points on a Likert scale.

2.4   Scoring SJTs

The next stage of SJT development is to develop a scoring key. Research findings have previously demonstrated that the scoring method used for an SJT can reduce the impact of extreme response tendency on SJT scores, as well as the presence of faking effects (De Leng et al., 2019). In addition, by controlling for systematic error, an SJT developer can increase the internal consistency reliability of an SJT and lower its adverse impact (De Leng et al., 2017). Unlike cognitive ability tests, SJTs do not always have objectively correct answers and it is often a case of which answer is 'best', rather than which answer is 'right' (Bergman et al., 2006). There are various scoring approaches documented in the SJT literature, which include empirical scoring, theoretical scoring, expert-based scoring, and hybrid keying (Bergman et al., 2006). With empirical scoring, SJT items are scored according to their association with a criterion measure. With theoretical scoring, SJT items are scored according to a theory; for example, in the case of construct-driven SJTs, test-takers would receive higher scores for endorsing response options that expressed high levels of the targeted trait and lower scores for endorsing options with low levels of the targeted trait (Lievens, 2017b). With expert-based scoring, a scoring key is typically determined by an SME panel who have specific expertise on the competencies required for a role. In the latter instance, SMEs are asked to judge how effective the various response options are; this method of scoring items can also be referred to as rational scoring. With hybridised scoring,

different scoring keys that have been generated using different scoring approaches are combined Thus, one scoring key can be used as a primary key, yet a secondary key can influence some of the item scores, such as replacing scores of zero that are generated by the primary key (Bergman et al., 2006). Bergman et al., note that using an empirical and theoretical key together can help overcome some of the concerns raised by each approach, because it recognises theory and relies less on solely empiricism.

Various methods are used to score SJTs; for example, using the Chan and Schmitt method, participants are awarded points for each item dependent on what percentage of SMEs endorse the same point on a rating scale; therefore, higher scores represent better situational judgement (Chan and Schmitt, 1997; 2002). Similar to the Chan and Schmitt method, MacCann and colleagues (2004) refer to the proportion scoring method, noting that this allocates a test-taker a score for each item based on how many other individuals endorsed that response. Some other consensus scoring methods noted in the literature include mode scoring, lenient mode scoring, dichotomous scoring, distance scoring, and adjusted distance scoring (MacCann et al., 2004; Weng et al., 2018). A description of each can be viewed in Table 2.2. An alternative scoring approach that has received increasing attention is pairwise comparison scoring (Gold and Holodynski, 2015; Reiser et al., 2022; Rosman et al., 2016). In sum, SJTs can differ in their format as well as their scoring approach, all of which can influence the internal consistency reliability and construct and criterion-related validity of an SJT (De Leng et al., 2017; 2019).

Table 2.2 Five consensus methods (adapted from MacCann et al., 2004; Weng et al., 2017)

Scoring method	Description
Proportion scoring	A test-taker is allocated a score for each SJT item dependent on the proportion of individuals that endorse that response
Mode scoring	A test-taker is allocated a positive score for each rating or category that is chosen as most effective by the largest proportion of individuals
Lenient mode scoring	A test-taker is allocated marks as per mode scoring, but is also allocated a positive score for items chosen on either side of the mode rating
Dichotomous scoring	A test-taker is allocated a score, of one, if their response falls within the range of response options that is considered correct; otherwise, they receive a score of zero.
Distance	A test-taker is allocated a score for each SJT item dependent on how far away their rating/response is from the mean rating provided by all individuals for that item
Adjusted distance	An individual's distance scores are converted into z scores

2.5   Reliability (and its problems)

Assessing the reliability of SJTs is problematic (Whetzel and McDaniel, 2009). With traditional SJTs, item stems (i.e. scenarios) and their response options are construct heterogeneous (McDaniel and Whetzel, 2005) and assess multiple constructs simultaneously. As a result, the use of traditional metrics of reliability, such as Cronbach's alpha, can be inappropriate. The reliance of such approaches to estimating test reliability can lead to an underestimation of the internal consistency of items (i.e., the extent to which item scores intercorrelate; Whetzel and McDaniel, 2009, Catano et al., 2012). This can be due to failing to accommodate for various method effects, such as the dependency of responses to items related to the same stem. A meta-analysis has previously reported internal consistency coefficients ranging from 0.43 to 0.94 across 39 SJTs (McDaniel et al., 2001). A separate study reported internal consistency coefficients ranging from 0.32 to 0.69, within subjects, which appeared dependent on the scoring method and response instruction used (Ployhart and Ehrhart, 2003). In the latter study, rating formats had the highest internal consistency reliabilities; and forms where test-takers were asked to select one

response option only had slightly smaller internal consistency reliabilities in comparison to forms with two choices (e.g., what would you most and least likely do).

Findings from factor analytic SJT research also demonstrate that the responses to SJTs are 'multidimensional', which, in turn, provides a rationale for their low internal consistencies (Lievens et al., 2008). However, in this sense, the response patterns may not be well described by a multidimensional factor model. Rather, many SJTs used in personnel selection have what might be referred to as 'an essential unidimensional' structure. That is, one, main factor may explain a substantial portion of the variance in responses, but there may be several smaller factors with items loading on them indistinctly (Nandakumar, 1991). Instead of assessing the internal consistency of SJT scores, it is recommended that reliability is assessed using test-retest reliability, or parallel form reliability (Whetzel and McDaniel, 2009). Findings of a longitudinal study demonstrated that SJT test-retest reliability was higher than internal consistency reliability across two populations, and for two different types of SJT response instructions (Catano et al., 2012). Despite this, many studies assessing SJTs continue to report on internal consistency estimates. Chan and Schmitt (2002) used two parallel SJT forms to assess overall job performance as well as task performance, motivational contextual performance and interpersonal contextual performance. The authors reported the parallel form reliability as 0.76.

To summarise, caution must be given when interpreting the results of any meta-analysis regarding the reliability of SJTs. Noting that an assessment can be 'reliably wrong', it has previously been suggested that *'greater emphasis should be placed on establishing the predictive and concurrent validity of selection tools instead'* (Patterson, 2018, p.4).

2.6   Validity

Three types of validity are typically evaluated to determine whether an assessment tool measures what it purports to measure; these include evaluating a tool's construct validity, criterion-related validity, and incremental validity.

2.6.1 Construct validity

The construct validity of an assessment provides an indication of how much an assessment measures what it intends to measure (e.g. integrity or empathy; Patterson, 2018). There are many ways of evaluating an assessments construct validity and two means of doing so include exploring an assessment's convergent or criterion-related validity.

2.6.1.1 *Convergent and discriminant validity*

To establish an assessment's convergent validity, one can make comparisons between scores on the assessment and scores on other measures of similar constructs (Patterson et al., 2016d). For example, one could evaluate whether scores on an assessment evaluating ones' knowledge of professional behaviours are positively correlated with their level of agreeableness. To demonstrate convergent validity, a prior study reported that scores on an SJT that was developed to assess integrity for medical school selection significantly correlated with all other integrity related-measures used in the study (De Leng et al., 2018). In addition, the same authors evaluated whether SJT scores were correlated with scores on the self-efficacy subscale of the Motivated Strategies for Learning Questionnaire (Pintrich et al., 1991). De Leng et al. (2018), observed that none of the SJT scores were significantly correlated with these subscale scores, thus demonstrating discriminant validity.

One meta-analytic study reported that scores from SJTs had a mean correlation of .46 with general cognitive ability tests (McDaniel et al., 2001). However, there was great variability around this mean with some SJT scores being more highly correlated with cognitive ability than others. McDaniel et al. (2001), suggest that moderators of this relationship include whether the SJTs were based on a job analysis and the amount of detail in SJT questions. As reported by the McDaniel et al. (2001), SJTs that generate scores which are more highly related to general cognitive ability include those that are based on a job analysis, and those with more detailed questions. Tests with knowledge-based instructions have also been observed to have higher correlations with cognitive ability than tests with behavioural-tendency instructions, which instead show higher correlations with personality constructs

(Mcdaniel et al., 2007). Chan and Schmitt found that performance on an SJT, which was developed to predict overall job performance, was uncorrelated with cognitive ability (r = -.02, 2002), thus demonstrating that, where desired, an SJT can be designed to tap less into general cognitive ability as opposed to other constructs.

2.6.1.2 *Criterion-related validity*

The criterion-related validity of an assessment indicates the extent to which the SJT scores can predict subsequent job performance or behaviour (Patterson et al., 2016d), therefore, also providing evidence for an instrument's construct validity. A meta-analysis that obtained 102 correlation coefficients between SJT scores and job performance reported an estimated population validity of 0.34 across a range of measures and samples (McDaniel et al., 2001); as McDaniel et al., highlight, this is a similar level to that reported in a meta-analysis of the validity of assessment centres (Gaugler et al., 1987). A more recent study by Lievens and Patterson (2011) found also that the criterion-related validity of scores on an SJT and an assessment centre were similar with regard to supervisor-rated job performance for advanced-level high-stakes testing. The scores resulting from SJTs have been found to demonstrate varying levels of criterion-related validity however, due to the SJTs differing characteristics and the construct domains they are built to assess. Another meta-analysis reported estimated population validities of 0.26 for both knowledge-based and behavioural-tendency based instructions; however, when the SJT content was kept constant, knowledge-based instructions had substantially higher criterion-related validity (Mcdaniel et al., 2007). Whilst some validities reported in the literature may be smaller, criterion-related validities are increased when the content domains of predictors and criteria are appropriately matched (Christian et al., 2010). It is therefore recommended that appropriate outcome markers are chosen when determining how effective a selection assessment is (Patterson, 2018).

2.6.2 Incremental validity

In numerous, though not all instances, SJT scores have been found to have incremental validity above and beyond other predictors of performance, including cognitive ability and personality, as well as job experience (Chan and Schmitt, 2002; Clevenger et al., 2001; Mcdaniel et al., 2007). These findings indicate that the SJT is assessing something distinct and unique, that is not observed using traditional selection procedures (Weekley and Ployhart, 2006). Whilst most studies report modest values of incremental validity, often in the range of 5 to 10% variance, it is proposed that incremental validity values will be larger when adjusting for the effects of constructs that are different from those assessed by SJTs (Webster et al., 2020). Ratings derived from assessment centre exercises have demonstrated incremental validity over and above SJT scores in predicting supervisor-rated job performance; nevertheless, an SJT provides a promising cost-effective method that has demonstrated similar criterion related validity to an assessment centre (Lievens and Patterson, 2011).

2.7   Issues (challenges to development)

Whilst there is a lot of evidence that supports the use of SJTs as a valid predictor of performance, there are also challenges to their development that can influence their effectiveness in practice. SJTs can be prone to faking, practice, and coaching effects (Lievens et al., 2008), however these issues can be minimised through the careful design of an SJT.

2.7.1 Faking

In high stakes selection, applicants want to provide the best answers to an assessment to obtain higher test scores. Thus, when provided a behavioural-tendency instruction on an SJT (e.g., 'what would you do'), it is unlikely that a candidate would want to provide an undesirable response, and they may instead, select what they think they ought to do in that situation. A study that explored the fakability of an SJT on college students' performance found that students' whom were asked to fake their answers on an SJT in order to provide the best impression, scored significantly higher than those that were asked to answer the

SJT honestly (d=.89; Peeters and Lievens, 2005). This finding is concerning for high-stakes selection, however, evidence highlights that knowledge-based instructions are less prone to faking than behavioural-tendency instructions (McDaniel and Nguyen, 2001; Nguyen et al., 2005). Therefore, it may be more appropriate to use knowledge-based questions, as opposed to behavioural-tendency instructions, when developing an SJT for selection purposes as you cannot fake what you do not know.

2.7.2 Coaching

Whilst organisationally-endorsed coaching can improve a person's responses on an SJT, it is argued that this does not affect the criterion-related validity of the SJT scores; also, coaching effects may be reduced if all applicants are given the opportunity to practice similar SJT items (Stemig et al., 2015). Knowledge-based instructions are also less susceptible to coaching (Patterson et al., 2016a), providing a further argument for utilising knowledge-based instructions in high stakes selection settings.

2.8   Subgroup differences

Prior research has demonstrated that, on average, white test takers perform better on SJTs than test takers from a Black, Asian and Minority Ethnic (BAME) background; also, female test takers typically perform slightly better than male test takers (Whetzel et al., 2008). Whilst subgroup differences have been observed however, SJTs are high-validity assessments that result in less adverse impact against minority subgroups than those observed for traditional selection procedures, such as cognitive ability tests (Weekley and Ployhart, 2006). Students from more privileged backgrounds typically have stronger academic records than those from less privileged backgrounds (Patterson, 2018), which can result in an unfair selection procedure. This concern can be partially addressed using selection methods that assess for non-academic attributes, such as an SJT. SJTs may therefore be more favourable to organisations that embrace diversity in the workforce. The importance of a diverse healthcare workforce has previously been highlighted; we must

have well-trained staff that understand the culture of the communities they serve (Patterson et al., 2018).

SJTs can be designed to minimise subgroup differences; for example, if the cognitive loading of an SJT is low, they have less adverse impact towards minority groups (Lievens et al., 2008). Black test takers tend to provide extreme ratings on a Likert-scale SJT more frequently than White test takers (McDaniel et al., 2011); therefore, McDaniel et al. (2011) recommend controlling for these response tendencies when scoring SJTs. Also, an SJT developer could include less points on a rating scale (e.g., having four Likert points instead of eight) to minimise the impact of extreme response style.

2.9   Applicant reactions

If the selection process is viewed unfair by candidates, there is an increased risk of litigation (Patterson, 2018). It is therefore important that applicants judge assessments as fair. Patterson (2018) suggests that to attract potential employees to a role, it is important that assessments are perceived relevant to the position, thus possessing face validity. Studies have previously demonstrated that SJTs are rated favourably by candidates (see Lievens et al., 2008; Patterson et al., 2016b).

2.10   Conclusion

There is currently no widely used, single, structured assessment tool used for the selection of staff into mental health service roles. However, SJTs have generally been found to be a valid and relatively cost-effective method for staff selection, especially where interpersonal skills are part of the job role (Lievens, 2013; Lievens and Sackett, 2012). Whilst challenges have been noted regarding the development of SJT items, it is highlighted that some of these challenges are due to a lack of resources that organisations have available to invest in the development, refinement, piloting and evaluation of this process (Health Education England, 2014a). It can be both time consuming and resource intensive to develop an SJT, however, it is widely accepted that SJTs, when properly constructed, are a valid and cost-

effective selection method (Patterson et al., 2016a), that are good predictors of job performance (McDaniel et al., 2001; Motowidlo et al., 1990). Furthermore, compared to cognitive ability and personality tests, scores resulting from SJTs show increased criterion-related validity and incremental validity (Lievens et al., 2008). Prior studies demonstrate that SJTs can assess non-academic attributes for medical education and training (Patterson et al., 2012), including an individual's values (Groothuizen, 2020). The acceptability of SJTs to applicants is positive. SJTs can also be delivered to large populations at the same time via the internet (Lievens et al., 2008), and can be marked mechanically against standardised scoring criteria (Weekley and Ployhart, 2006). For this reason, SJTs may be easier to implement in large-scale testing programs (Weekley and Ployhart, 2006).

Motowidlo and Beier (2010) propose that where ITPs have some level of stability across a domain of jobs and organisational settings, it could be beneficial to develop an SJT that would assess those ITPs, even if the specific knowledge required for the role did not generalise across jobs. Whilst a thorough job analysis is recommended to assess the key attributes and competencies required of a role when developing an SJT (Patterson and Driver, 2018), there are specific attributes expected of all healthcare professionals, which include compassion, respect, and caring for patients (Patterson et al., 2016b). It is conceivable that a single SJT could assess the interpersonal attributes desired of all staff working in mental health services and that such assessment would, in turn, facilitate the selection of staff with the appropriate values and qualities to work in mental health services. An SJT therefore has the potential to help select a more compassionate and caring workforce, well suited to the nature of the work.

Chapter 3: Philosophical Foundations

Chapter 2 of this thesis provided a detailed description of Situational Judgement Tests (SJTs), while presenting the evidence base supporting their use for personnel selection. Having chosen to develop an SJT to assess professional judgement in mental health services, this chapter outlines the philosophical foundations underlying the methods and methodology used for this mixed-methods programme of work.

3.1   Reflexivity: my starting point

Researcher bias can be present at any stage in the research process, from developing a research question to interpreting the research findings. Reflexivity however increases the validity of qualitative research (Pillow, 2003), as researchers consider how their background, characteristics, and engagement with the research process may influence the data generated and their subsequent interpretation of findings (Berger, 2015).

Before detailing the methodology underlying this project, it is felt appropriate to first discuss the author's motivations for embarking on the project itself. To fulfil this objective, the author wishes to speak in the first person, recognizing the influence their experience and beliefs have had on the research process. It is not uncommon for a researcher's personal experience to influence their choice of research approach (Creswell, 2014). I applied to undertake this PhD in 2016; at that time, I had received care from mental health services and had previously been detained under the Mental Health Act (1983; that is, being detained, in hospital, involuntarily). During that period of hospital treatment, I was physically restrained and given medication against my will; all whilst experiencing psychosis. As one may expect, I was left with some misgivings about the mental health system, as well as some staff members working within that system.

Not only had I received care from mental health services, but I had also worked within the organisation that provided this care; job roles I had held within the organisation included working as a care assistant for adults with learning disabilities, training and completing a

diploma in Health and Social Care, working as an assistant psychologist and a research assistant, and working in a clinical audit and compliance capacity. Whilst the former roles were more clinical in nature, the latter two roles were in corporate positions and involved reviewing the Trust's compliance with health care regulation, guidance, and standards.

I was familiar with many staff members and patients in the organisation where the research was performed; some knew me as a patient, supporting me in either inpatient or outpatient services, others knew me as a colleague. Referring to focus groups, some argue that it can be unhelpful to choose a moderator who is an insider (Barbour, 2007, p. 50), however, Berger (2015) notes that her 'insider' perspective facilitated the conduct of her research as it helped her gain entry to the field she was researching and increased her ability to recruit participants and process the data generated. The personal knowledge I held regarding the organisation certainly helped facilitate the conduct of this doctoral work, which included the recruitment of staff.

As highlighted in the acknowledgements section of this thesis, I received an advertisement for this doctoral project from my former boss Dr Angela Kennedy while I worked as a research assistant. The thought of undertaking this project appealed to me greatly. I have always wanted to play a part in improving mental health services, not only due to my own experience as a patient, and someone that lives with mental health difficulties, but also having been a child of parents that had experienced mental health difficulties too. Recognising my own psychological struggles of working directly with individuals that have mental health difficulties, I believed that research and service development was my forte. The opportunity to develop an assessment that would help select suitable staff into mental health services appealed to me greatly. I was forever grateful for this opportunity and, additionally, the stipend that Hull York Medical School provided to undertake this work.

A key component of this project was to 'develop ways of constructing and then measuring peoples' knowledge of professionalism in mental health services'; with this framework in

mind, it was intended that a pool of SJT items would be developed and validated. The project was funded by an Athena SWAN award, presented to Professor Finn, that was granted to promote gender equality and advance the careers of women researchers. It is worthy to note that this doctoral work also sat within a wider project led by Professor Paul Tiffin, namely 'Enhancing selection of the healthcare workforce: a programme of psychometric epidemiology.' The project led by Professor Paul Tiffin was part of a Career Development Fellowship funded by the National Institute for Health Research.

Since the commencement of this doctoral project, I have been discharged from secondary (specialist) mental health services. I have had the fortune of meeting staff from many disciplines and am pleased to have developed a more positive image towards professionals working in the field. I am delighted that I have had the opportunity to develop an SJT that will help select more staff with the appropriate values, knowledge, and behaviours to deliver effective mental health care.

3.2   Paradigms – 'researchers' beliefs about their efforts to create knowledge'

Noting that paradigms underpin and inform the design of quantitative and qualitative research, Brown and Dueñas (2020) recommend that researchers select their paradigm before they commence their research. Thomas Kuhn is well known for his use of the term 'paradigm', and this likely led to the popularity of its use for summarising researchers' beliefs about their efforts to create knowledge (Morgan, 2007). Whilst the term paradigms has proved helpful for some, it is not uncommon for social scientists to use the term and mean completely different things (Morgan, 2007).

Morgan (2007) categorised uses of the term 'paradigms' into four basic versions; Paradigms as Worldviews, Paradigms as Epistemological Stances, Paradigms as Shared Beliefs Among Members of a Specialty Area, and Paradigms as Model Examples of Research. The author will briefly summarise the first two of these, Paradigms as Worldviews and Paradigms as

Epistemological Stances. According to the use of the term 'paradigms as worldviews', paradigms can be described as the way that people experience and think about the world in its entirety (Morgan, 2007). Morgan notes, however, that it is not very helpful to think of worldviews as encompassing 'everything' that someone thinks and believes, and that researchers should instead focus on their thoughts about the nature of research. In discussing 'Paradigms as epistemological stances', Morgan notes that this version of the term takes a narrower approach than using the term 'paradigms as worldviews' as it concentrates on one's beliefs regarding issues within the philosophy of knowledge, including the nature of knowledge and knowing (Morgan, 2007). Morgan's use of the term 'epistemological stances' aligns with Creswell's use of the term worldview (see below; Creswell, 2014). It is worth noting at this point that it is not unusual for a lack of clarity regarding what a paradigm consists of (Brown and Dueñas, 2020); as mentioned earlier, social scientists often use the term 'paradigm' and mean completely different things (Morgan, 2007).

It has been suggested that paradigms are constructed of several 'building blocks', which include axiology *(a theory of value)*, ontology *(the nature of reality)*, epistemology *(the nature of knowledge)*, methodology *(the nature of research)*, methods *(research procedures),* and sources *(available research data;* Biedenbach and Jacobsson, 2016; Brown and Dueñas, 2020; Bunniss and Kelly, 2010); social scientists may have differing beliefs regarding these 'building blocks' and thus align to different theoretical orientations. Figure 3.1 displays these building blocks as part of an iceberg metaphor, recognising that all these elements are intertwined and influence one another. The blocks beneath the surface influence our methods, yet it is only the methods we use that are clearly visible to those around us.
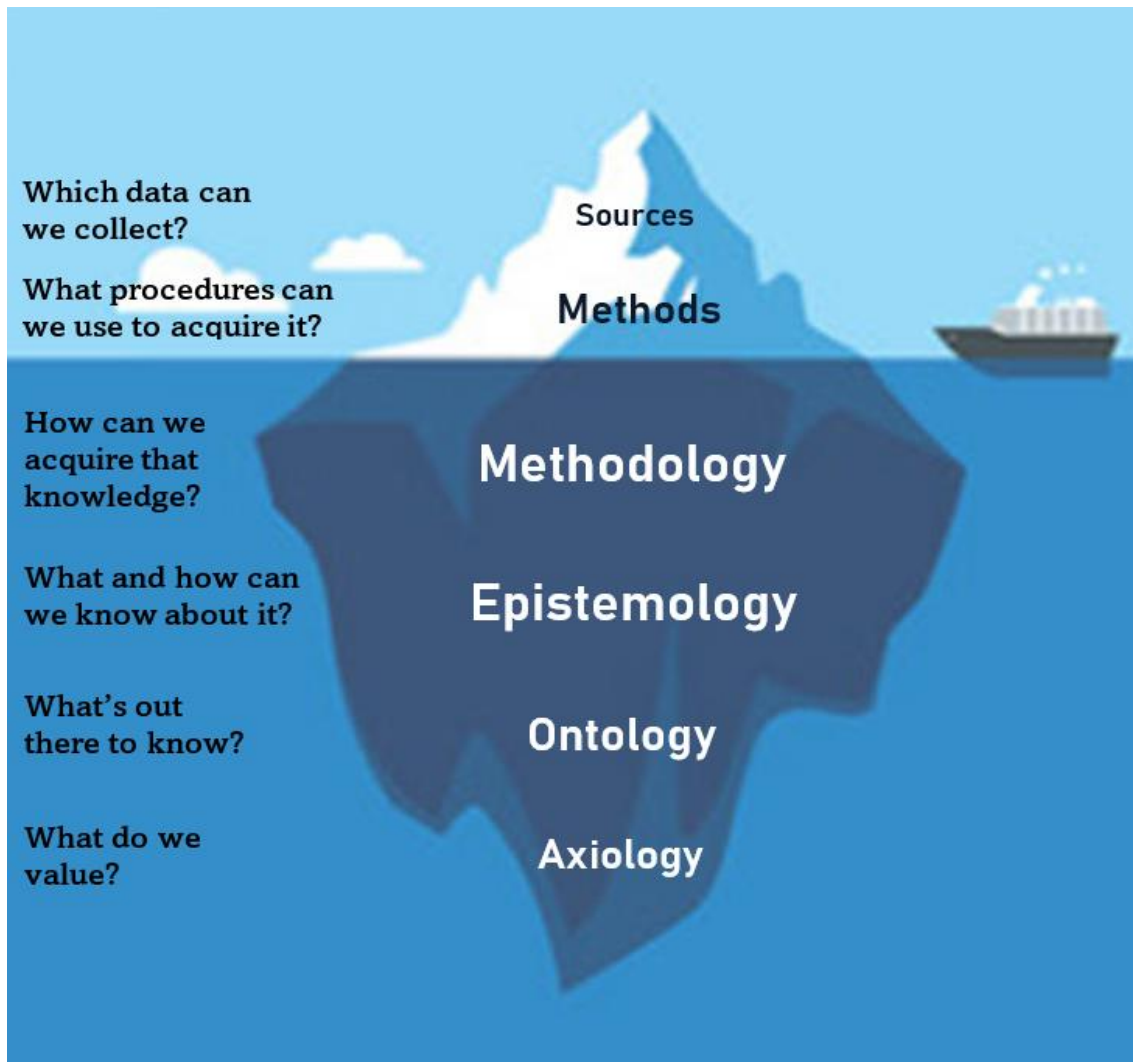
Figure 3.1 The building blocks that form a research paradigm (adapted from Brown and Dueñas, 2020; the iceberg image was obtained via an Adobe Stock license)

3.3 Theoretical orientations

In a book on research design, Creswell (2014) details four key beliefs held by researchers that are widely cited in the literature, including; the Postpositivist worldview, the Constructivist worldview, the Transformative worldview and the Pragmatic worldview. A summary of these theoretical orientations is provided below. Prior to detailing these, however, the author wishes to first discuss positivism.

3.3.1 Positivist research

The positivism doctrine was created by Auguste Comte in the 19th century and was based on the conviction that human behaviour could be examined using the same laws of the physical

and natural sciences (Miller and Brewer, 2003; Salkind, 2010). In Comte's system, the researcher could apply methods used in the positive sciences to the humanities in order to obtain knowledge (Salkind, 2010). It has been proposed that positivism is a misunderstood term (O'Reilly, 2009) that has different meanings for many current researchers and academics (Salkind, 2010). Nevertheless, the paradigm has several central tenets (see Lewis-Beck, 2004). A few of the beliefs held by positivists include: that all our knowledge of the world is received through our senses as we experience it; that we can only know what we observe; that we can test anything we claim to be true through the use of observation and experimentation; that recurring patterns of experience can be stated as scientific laws; that knowing these laws should enable us to predict future occurrences of that phenomena; and that scientific objectivity rests on testable factual statements, as opposed to subjective value judgements (O'Reilly, 2009). Essentially, there is a 'real world' out there that is independent of people's perceptions of it; this reality is fixed and unchanging and research can accurately access this (Miller and Brewer, 2003).

Lincoln and Guba have summarised several assumptions that sit within the positivism paradigm, including: 1) that there is a single, tangible reality; 2) that the knower is independent of the known; 3) that knowledge is generalisable; 4) that there are no effects without causes and vice versa; and 5) that the use of appropriate methods mean that the study is independent and not influenced by the researcher's values (Lincoln and Guba, 1985). It is worth noting that in the social sciences, positivism is commonly rejected by the other paradigms, because of its belief that science can discover definitive objective facts (Corry et al., 2019).

3.3.2 Postpositivist research

Like positivism, the term 'postpositivism' is riven with ambiguity among researchers and there is no singular agreed definition (Corry et al., 2019). It has been suggested that

"Postpositivism was a reaction to the widely discredited axioms of positivism, and many of its tenets were in direct opposition to those of its predecessor" (Clark and Creswell, 2008,

p.12). Creswell and Poth (2007, p.23) note that postpositivists do not believe in strict cause and effect, but instead believe that '*all cause and effect is a probability that may or may not occur*'. Furthermore, the authors suggest that postpositivists believe in multiple perspectives from participants as opposed to a single reality. In contrast however, an article by Bunniss and Kelly (2010) asserts that postpositivists, like positivists, believe that reality is fixed and static and there is an overarching objective truth. Postpositivist tenets include the ideas that: the values held by investigators influence research (value-ladenness of inquiry); the theory, hypotheses, or framework that an investigator uses also influence research (theory-ladenness of facts); and investigators' understanding of reality is constructed (nature of reality; Clark and Creswell, 2008).

3.3.3 Constructivist research

The discrediting of positivism also resulted in more radical paradigms than postpositivism, which included constructivism (Clark and Creswell, 2008). Referring to constructivism, Guba and Lincoln utilised the term 'naturalist' (Clark and Creswell, 2008), whilst subsequently listing several clear distinctions between naturalistic inquiry and positivism (see below; Guba and Lincoln, 1982):

1) whereas positivists believe in a single, tangible reality, naturalists believe in multiple, intangible realities.

2) naturalists believe that the knower and the known are interrelated and not independent of one another, as believed by positivists.

3) whilst positivists believe that generalisations are possible, regardless of context, naturalists believe that this is not possible.

4) in comparison to positivists' belief that cause-effect relationships can be established, unequivocally, for every action, naturalists believe that inquirers can, at best, establish plausible inferences regarding such patterns.

5) Naturalists believe that inquiry is value-bound, as opposed to value-free.

This list has since been added to by Clark and Creswell (2008), who note that in constructivism *'there is an emphasis on arguing from the particular to the general'*.

3.3.4 Transformative research

The transformative approach arose during the 1980s and 1990s from individuals that believed the postpositivist assumptions imposed laws and theories that did not fit marginalised individuals in society; transformative researchers also believed that the constructivist paradigm did not support strongly enough an action agenda to help marginalised individuals (Creswell, 2014). A fundamental belief behind this approach is that knowledge reflects the power and social relationships in society, and is thus not neutral (Creswell and Poth, 2007). Transformative researchers include critical theorists and participatory action researchers. Furthermore, transformative research typically contains an action agenda for reform and would incorporate studies that may change the lives of participants or institutions in which they work or reside (Creswell, 2014).

The transformative approach holds that ethical research needs to promote social justice and further human rights (Mertens, 2017). In addition, the framework assumes that there are multiple versions of reality influenced by peoples' different societal positionalities, such as their gender, more or less privilege, ethnicity, and disability status, among other factors (Mertens, 2017). According to Merton (2017), transformative researchers need to be aware of their own power and cultural standpoint when undertaking research, and be mindful of how these factors influence their relationships with the research participants. Transformative researchers' methods are not dictated by methodological assumptions (Mertens, 2017) and researchers that adopt this paradigm may therefore utilise mixed methods.

3.3.5 Pragmatist research

Pragmatism has previously been presented as a third paradigm, sitting on a continuum between the qualitative and quantitative research, whilst utilising a mixed methods approach (Johnson and Onwuegbuzie, 2004). A pragmatic researcher will use multiple

methods of data collection and may use multiple sources of data to best address the research problem. The primary focus of the pragmatist is on the practical implications of the research, rather than antecedent conditions (Creswell, 2014; Creswell and Poth, 2007). Whilst other theoretical orientations are particularly concerned with truth, pragmatists are only concerned with what works at the time; therefore, it does not matter whether reality is independent of the mind or not (Creswell, 2014). Instead of being concerned about the nature of reality and the possibility of truth, pragmatists are instead concerned with 'what difference' the research will make (Morgan, 2007).

3.4   Paradigm wars

The debate between the alternative paradigms, particularly the qualitative and quantitative debate (i.e., the positivist paradigm versus the constructivist paradigm) has been referred to as the 'paradigm wars' (Clark and Creswell, 2008). The 'paradigm wars' date back to the late nineteenth century (Smith and Heshusius, 1986) and applies at various levels, including data, design and analysis, and the interpretation of results (Howe, 1988). Advocates of the 'incompatibility thesis' believe that quantitative and qualitative methods are incompatible and mutually exclusive, due to their differing beliefs regarding reality, truth, and the relationship between the investigator and the object of investigation (Howe, 1988). According to these advocates, research that tries to combine the two approaches are doomed to failure due to differences in the philosophies underlying them (Clark and Creswell, 2008).

In comparison to advocates of the 'incompatibility thesis', pacifists believe that qualitative and quantitative approaches are compatible (Clark and Creswell, 2008; Howe, 1988), and that qualitative and quantitative methods are merely different approaches to research that should be used pragmatically, dependent on the research question (Bryman, 2012, as cited in Holloway and Galvin, 2017). Paley and Lilford (2011) believe that it is unnecessary for qualitative research to align itself to a particular philosophy and instead suggest that qualitative and quantitative methods are 'alternative tools' that can be used for different

tasks in research. Indeed, the core assumption of using both qualitative and quantitative methods is that the combination of these approaches provides a better understanding of a research problem than what would be obtained using either method alone (Creswell, 2014).

3.4.1 The author's research position

The focus of this doctoral work was to develop an SJT that would assess practitioners' knowledge of professionalism for mental health services. The author started these doctoral works as a pragmatist and was not overly concerned regarding the 'nature of reality' or the 'possibility of truth'; instead, their primary intention was to develop an instrument that would effectively select professionals with the appropriate knowledge and behaviours to work in mental health services. Nevertheless, for transparency, the author believes it is important to share some of their views here.

3.4.1.1 *Axiology – a theory of value*

The author makes no secret of the fact that they have previously been hospitalised in mental health services. It is predominantly due to this experience that the author wanted to undertake this project and, in turn, felt it valuable to develop an assessment that would help select appropriate staff for this setting.

3.4.1.2 *Ontology – the nature of reality*

Psychological attributes differ from physical objects (Guyon et al., 2018). The author agrees with Guyon, that mental attributes can be viewed as characteristics that can be inferred from the behaviour of a person; they are a state of '*equilibrium'* in an individual and can therefore be considered a reality (Guyon et al., 2018, p.154). Like Guyon, the author believes that these attributes need to be analysed in relation to their dynamic social environment.

3.4.1.3 *Epistemology – the nature of knowledge*

Whilst the author believes that mental attributes can be considered a reality to some degree, the author also believes that knowledge is subjective and that there are multiple

versions of reality. The author believes that this knowledge is co-constructed among groups and individuals and it is mediated by power relations, thus leaning towards the theoretical orientations of interpretivism and critical theory (Bunniss and Kelly, 2010). The author is aware of various power differentials that exist in mental health services, ranging from the care worker to the patient, and across and within professional groups. An example of which is the opposing views regarding the 'social model' and 'medical model' of disability (Hogan, 2019), noting that the bio-medical model has historically dominated the debate; that is, a narrow reductionist view of medicine that predominantly focuses on the biological factors without consideration of the psychosocial factors and influences. Henceforth, to promote inclusivity, the author was keen to involve patients and carers early on in the research process and continually throughout the project.

3.4.1.4 *Methodology – the nature of research*

A mixed methods approach was utilised for the current project, which aligns well to the author's original paradigm, pragmatism. Mixed methods research is an emerging field and incorporates the collection, analysis and mixing of qualitative and quantitative data in order to understand a problem more fully (Ivankova and Creswell, 2009). The adoption of mixed methods can help minimise the weaknesses and utilise the strengths of both quantitative and qualitative approaches in a single study, providing an innovative approach to address issues in health services (Fetters et al., 2013; Johnson and Onwuegbuzie, 2004). In line with pragmatism, the author agrees that qualitative and quantitative approaches to research should not be viewed as polar opposites, but instead lie on different ends of a continuum (Creswell, 2014). Furthermore, the author agrees with Paley and Lilford (2011), that qualitative and quantitative methods can be used for different tasks and do not necessarily need to align themselves with a particular philosophy at all. A mixed-methods process is common in the development of traditional SJTs (Tiffin et al., 2020), which leads the author to recollect Morgan's use of the term 'Paradigms as Model Examples of Research' (Morgan, 2007). As Morgan highlights, 'paradigmatic examples' highlight to newcomers how the field

address their central issues. Despite use of the term 'Paradigms as Model Examples of Research' receiving little attention, the current author recognises how model examples in the field have guided their work (e.g. Tiffin et al., 2020; see Figure 3.2).

The four mixed methods designs most frequently used by researchers include Explanatory Design, Exploratory Design, Triangulation Design, and Embedded Design (Ivankova and Creswell, 2009); a sequential exploratory mixed methods design with instrument development was adopted for this project (Clark and Creswell, 2008). The latter approach involves the researcher first conducting qualitative research to explore the views of participants, before analysing and using the data in a quantitative study (Creswell, 2014). In psychometrics, one hopes to evaluate psychological phenomena (e.g., intelligence, personality or attitudes); it is necessary to operationalise these phenomena in order that one can attempt to assess quantitative manifestations of these, via test scores. As with a study conducted by Milton et al., the mixed-methods approach used for this project commenced with a qualitative phase, progressed to instrument development, and was followed by a quantitative study (Clark and Creswell, 2008; Milton et al., 2003). Some complementary quantitative and qualitative data was also collected during the qualitative and quantitative phases, respectively, which enhanced the development and subsequent validation of the SJT (the development and validation process is summarised in Figure 3.2). It is acknowledged that there is a lot of debate regarding the circumstances that psychological assessments represent fundamental measurement. This topic is therefore expanded upon in the discussion chapter of this thesis.

```
┌─────────────────────────────┐
│ Define test specification by│        • Rapid, systematic literature review
│ identifying the personal    │        • Qualitative study (data collection and analysis)
│ qualities deemed essential  │
│ or important to effective   │
│ workplace performance. This │
│ usually involves interviews │
│ or focus groups with subject│
│ matter experts (SMEs) study │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Elicit ''critical incidents''│       • Qualitative study (data collection)
│ that can serve as potential │
│ scenarios in the test,      │
│ usually via focus groups    │
│ with SMEs and other         │
│ stakeholders                │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Decide on a response format │        • SJT literature review
│ and create items relating to│        • Inductive and deductive item development
│ the scenarios               │          and review
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Involve SMEs in designing a │        • Follow up from qualitative study (quantitative
│ provisional scoring key,    │          data collection and analysis)
│ based on expert consensus   │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Trial pilot items in a      │        • Quantitative study (data collection, with
│ sample of test-takers,      │          supplementary qualitative data collected)
│ similar to those for whom   │
│ the test is intended as a   │
│ selection measure. Ideally, │
│ some data relating to       │
│ validity should also be     │
│ obtained at this stage      │
└─────────────────────────────┘
              ↓
┌─────────────────────────────┐
│ Review item construction,   │        • Quantitative study (data analysis, with review
│ inclusion and scoring in    │          of supplementary qualitative data)
│ light of psychometric       │
│ analysis of pilot response  │
│ data                        │
└─────────────────────────────┘
```

Figure 3.2 The 'traditional' approach to SJT development (adapted from Tiffin et al., 2020)

3.5   Methods

A systematic review and qualitative study, which incorporated both interviews and focus groups, were undertaken to develop an understanding of the concept of professionalism, as well as identify the professional attributes desired of professionals working in mental health services. Thematic analysis and framework analysis were utilised to analyse the data generated during the review and qualitative study, respectively. In comparison to a quantitative study, a qualitative approach was deemed most appropriate at this stage of the project as qualitative research is recommended when exploring the meaning individuals ascribe to a particular social problem (Creswell, 2014). During the qualitative phase, the critical incident technique was utilised to generate examples of professional scenarios and identify the challenges that staff members face (Flanagan, 1954). Noting that SJT scenarios should be based on a thorough job analysis to identify the key attributes associated with effective performance in the role (Patterson et al., 2016d), this was deemed to be a suitable approach.

Noting that cognition cannot be understood as a purely internal cognitive process (Merriam, 2010), but is instead intrinsically linked to the social and cultural contexts in which it occurs (Cobb, 2001), SJT items were drafted to assess the 'situated cognition' of professionals working in mental health services. It was hoped that the SJT scores would predict which professionals would be more effective in practice and thus have the potential to facilitate the selection of staff into mental health services. Items were appraised by an SME panel, who also rated the SJT item responses. SME consensus was evaluated to determine the best response to each item. During development of the initial SJT instrument, a 'blue printing' process was undertaken, which involved the author assigning pilot SJT items to one of six professional attribute domains - domain labels were determined through analysis of the interview and focus group data. This blueprinting process facilitated the development of two SJT instruments that incorporated a similar range of items.

Whilst the qualitative study served its purpose, it was yet to be determined whether the SJT would work as intended (i.e., the capability of the SJT to assess ones' workplace effectiveness was yet to be established). A subsequent quantitative study enabled the evaluation of the utility of the SJT. The collection of hard outcomes related to patient benefit would have been most desirable (for example, a patient's duration in hospital); however, pragmatically this data is very hard to obtain. Therefore, a decision was made to collect supervisor ratings of job performance instead; the collection of supervisor and manager ratings of effectiveness is common in SJT validation studies (e.g. Bledow and Frese, 2009; Cousans et al., 2017; Crook et al., 2011; Motowidlo et al., 2009). A bespoke supervisor rating tool was developed, and a pilot study was undertaken. Quantitative research is a suitable approach for testing objective theories by examining relationships between variables (Creswell, 2014); a mixed methods approach was thus utilised for this project.

3.5.1 Aim

The overall aim of the research project was to 'develop and validate an SJT that would assess staff members' knowledge of professionalism for a mental health services context'.

3.5.2 Original objectives

1. To define 'professionalism' for a mental health services context.

2. To develop a pool of SJT items that could validly evaluate 'knowledge of professionalism' in a mental health setting.

3. To develop and validate an SJT for personnel selection in mental health services.

4. To develop a bespoke workplace behaviours rating tool.

3.5.3 Departures from the original research protocol

Initially, the researcher hoped to develop a conscientiousness index for mental health services, thus building on the work of McLachlan et al (2009). Recognising that many facets incorporated within the conscientiousness index would be difficult to obtain in the NHS

(e.g., scoring staff dependent on whether they were on time for shifts), a bespoke supervisor rating tool was developed instead; it was believed that colleagues would be best placed to evaluate how effective staff were, overall, on the job.

To help validate the SJT, there were initial plans also to collect data through the established Independent Assessment of Clinical Skills Programme (IACS) in the local NHS site; the IACS program has several stations, like Observed Structured Clinical Examinations, where core psychiatric trainees participate in roleplay whilst their clinical skills are assessed by qualified psychiatrists. These plans were put on hold however, as 1) supervisor ratings alone provided support for the validity of the scores resulting from the SJT, and 2) the Covid-19 pandemic prevented the IACS from going ahead on the dates originally scheduled. Other changes to the study protocol, as well as enhanced details regarding the methods used for each study are discussed in their associated parts of this thesis.

3.6   Conclusion

Having laid out the aims and objectives of this doctoral work, the remainder of this thesis reports on the programme of work undertaken to meet these objectives. Chapter 4 reports on the findings of a rapid, systematic review, which met objective (1) of the research project. The review findings were built upon by a qualitative study, which is reported on in chapter 5 of this thesis. The qualitative study also helped generate item content for the SJT, thus fulfilling objective (2) of the research project. Chapter 6 reports on the development and validation of the SJT, which met objective (3).

Chapter 4: A Rapid Systematic Review: Defining professionalism for Mental Health Services

All professions emphasise the importance of *professionalism* for working in mental health services. Yet, whilst the term is used widely, it is not always clear what this means. Having discussed the methodology and methods used for this doctoral work in the prior chapter (chapter 3), this chapter reports on the method and findings of a rapid systematic review, which sought to derive an operational definition of professionalism for a mental health services context. A definition of professionalism was required to inform the development of the situational judgement test (SJT).

4.1　Background

Professionalism has received increasing attention over the last 30 years, in both clinical medicine and medical education (Hodges et al., 2011). Nevertheless, little attention has been paid to the concept in psychiatric services. Patients using mental health services differ to that of other healthcare specialties. For example, in mental health services, patients may be deprived of their liberty and may also be the recipient of coercive practices (Barbui et al., 2021; Mental Capacity Act, 2005). Patients may suffer from various illnesses, including psychoses, which makes them vulnerable and less able to fend for themselves against professional malpractice and misconduct. Patients therefore rely on practitioners to

safeguard them against abuse. Patients also often have needs that rely on their interpersonal relationships with professionals, as well as others.

Given that staff are judged on their 'fitness to practice' and 'unprofessional practice' (see General Medical Council, 2022; Health & Care Professions Council, 2022; Nursing & Midwifery Council, 2022; Social Work England, 2022), practitioners must be aware of their own professionalism; a clear operational definition of 'professionalism' is thus needed. Such a definition would also provide a framework for the development of an SJT, thus enabling the assessment of practitioners' knowledge of professionalism for mental health services.

The health sciences' and medical professions' literature typically define professionalism as upholding professional values, exhibiting professional behaviours or demonstrating professional attitudes (Aguilar et al., 2011). There is an argument that professionalism varies, dependent on context (Brody and Doukas, 2014; Rees and Knight, 2007; van de Camp et al., 2004; Wear and Kuczewski, 2004) and attention must therefore be paid to the concept across healthcare specialties (van de Camp et al., 2004). Literature reviews have sought to define professionalism in medicine (van de Camp et al., 2004), medical education (Birden et al., 2014), surgical education (Deptula and Chun, 2013), dental education (Zijlstra-Shaw et al., 2012), nursing (Ghadirian et al., 2014), occupational therapy (Hordichuk et al., 2015), and across healthcare disciplines (Aguilar et al., 2011). Each professional group working in mental health services has their own code of ethics and professional regulations that they must adhere to; nevertheless, all professions working in this setting have the common goal of providing the best possible care to patients in this multi-disciplinary service.

4.2   Aim

This review aimed to define professionalism for a mental health services context. This was achieved by answering two research questions.

1) How is professionalism conceptualised within mental health services?
2) Does the definition vary across the mental health professions?

4.3   Theoretical framework and methodology

The review sought to derive an operational definition of professionalism for mental health services in order to inform the development of an SJT that would assess practitioners' knowledge of professionalism for this context. To allow ample time to develop the SJT, the review was conducted in a rapid, systematic manner, noting that systematic reviews help identify, evaluate and summarise the available evidence, making this more accessible to decision makers (Centre for Reviews and Dissemination, 2009). Whilst the review was conducted in a rapid manner, rapid reviews have been found to produce similar results to more thorough systematic reviews (Watt et al., 2008) and can be useful for informing health care delivery and policy (Khangura et al., 2012). A rapid approach was thus deemed to be a suitable and pragmatic choice for this review.

An integrative review was performed, as these reviews help generate new frameworks and perspectives on a topic (Torraco, 2005), and also facilitate the development of theory (Whittemore and Knafl, 2005). An integrative approach was deemed appropriate for the review question as it allows the use of all literature to be included in the review. Thematic analysis (Braun and Clarke, 2006) was performed and a narrative synthesis (Popay et al., 2006) employed to report on the findings.

4.4   Criteria

Whilst the review questions were more qualitative in nature, which some may argue is more suited to a SPICE (Setting, Perspective, Intervention, Comparison, Evaluation; Booth, 2006) or SPIDER (Sample, Phenomenon of Interest, Design, Evaluation and Research type; Cooke et al., 2012) framework, criteria were developed using the PICOS (Population, Intervention, Comparison, Outcome, Study Design) framework instead (Centre for Reviews and Dissemination, 2009), to facilitate early registration of the review. The PICOS framework demonstrates equal or higher sensitivity than SPIDER searches, which minimises the risk of relevant papers being omitted during the search (Methley et al., 2014). Early registration is recommended by the Centre for Reviews and Dissemination who note that specifying the

methods in advance minimises the risk of introducing bias during the review process (Centre for Reviews and Dissemination, 2009). The review was registered on the International prospective register of systematic reviews (PROSPERO) on 10 February 2017 (registration number: CRD42017056594). The protocol was subsequently amended on 12 June 2017; amendments were discussed with the review team as part of an iterative review process.

4.4.1 Population

Articles were considered if they referred to either 1) a mental health profession or 2) a mental health setting, whether that be generally or with a focus on a particular specialty (e.g., mental health services for older people). Learning disability services were considered as part of the wider mental health services, as learning disability services are delivered alongside mental health services in the United Kingdom (UK). Perceptions of all stakeholders were considered, including the opinion of authors, practitioners, patients, carers, students, or teachers, among other individuals. Exclusion criteria were as follows:

4.4.1.1 *Exclusion criteria*

- Articles that did not make reference to either the mental health setting or a registered mental health profession (as identified via the NHS jobs website; NHS Jobs, 2018)
- Articles referring to the practice of non-registered practitioners (e.g., peer support workers)
- Articles referring to substance misuse services or primary care settings
- Articles focusing on a subgroup of people based on their diagnosis or symptomatology
- Articles relating to low- or middle-income countries

Literature was excluded if it referred to substance misuse or primary care services as the therapeutic relationship and provision of care in these services differ to more specialist mental health services; it was hypothesised that these findings may skew any subsequent

definition derived during the review. Whilst the literature was initially included if it related to any mental health patients, a revision was made to the protocol in June 2017, to also exclude all articles that focused on a specific subset of patients; this amendment was made as it was believed that focusing on such articles would look at interventions and practices for these particular groups, and a subsequent definition may not be generalisable to the wider mental health patient population. Articles were also excluded if they focused on low- or middle-income countries; this decision was made in order to source data from a population where the provision of care is similar to that of the United Kingdom. Low-, middle- and high-income countries were identified via the World Bank Group (The World Bank Group, 2018).

4.4.2 Intervention

All articles that provided a definition or description of professionalism within a mental health services context and/or a definition or description of unprofessionalism within a mental health services context were considered. Certain exclusion criteria applied, as listed below:

4.4.2.1 *Exclusion criteria*

- Articles that focused on certain characteristics, but did not directly attribute these to professionalism (e.g., articles that discuss ethics, but make no reference to professionalism)

4.4.3 Comparison

The review did not set out to compare interventions or procedure and thus had no comparison group.

4.4.4 Outcome

The review outcomes were adapted from an earlier review (Birden et al., 2014) and included:

1) A comprehensive, universally accepted definition of professionalism pertaining to a mental health services context.

2) Definitions of professionalism pertaining to a mental health services context, according to profession.

An operational definition of professionalism was generated during the review, however, and was not directly obtained from the literature.

4.4.5 Study design

As quantitative research alone would not sufficiently answer the review question, all literature, except books, was included if it met the inclusion criteria. This is in keeping with integrative reviews, which allow the inclusion of theoretical literature as well as empirical research (Whittemore and Knafl, 2005). Certain exclusion criteria applied, as listed below:

4.4.5.1 *Exclusion criteria*

- Articles not written in English

- Articles that were not published between 2006 and 02nd March 2017 (date of final search)

- Literature reviews (revised and later included)

The research team was UK based and thus all articles were excluded if not written in English. Whilst this exclusion criteria would typically warrant concern in meta-analyses (Begg and Mazumdar, 1994; Duval and Tweedie, 2000; Easterbrook et al., 1991), this was a practical decision and was not a concern for this review. Articles were restricted by date of publication as a current definition of professionalism for mental health services was desired; this restriction was in place due to awareness that mental health practices have changed over the past few decades and the desire to keep the research findings contemporary.

4.5   Search strategy

The initial search string was modified from an earlier review (Birden et al., 2014), by three reviewers (LA, GF and PT), with support of an academic librarian who specialised in healthcare and an additional NHS librarian. Whilst the PICOS framework demonstrates equal or higher sensitivity than SPIDER searches, the SPIDER search demonstrates equal of higher specificity than PICOS searches (Methley et al., 2014); the SPIDER search is therefore likely to identify only those articles that are very closely linked to the review question. The review sought to identify a broad range of articles on this topic, so the PICOS framework was deemed appropriate for the review.

4.5.1   Electronic search

Pilot testing was performed, and the search string was adapted to improve the sensitivity and specificity of the search for a mental health services context. The base search was created using CINAHL and utilised free-text and subject heading searches (see Table 4.1). Search strings were adapted for additional healthcare databases, including Medline; EMBASE; PsycINFO; and HMIC (a list of adjusted search strings can be viewed in Appendix B). An Open Grey search was initially intended, but this search was omitted as HMIC also identifies grey literature.

4.5.2   Other resources

Of the records retrieved, a sample of reference lists were used to generate additional articles pertinent to the review question.

4.6   Study selection

Initially, duplicates identified during the search were removed. A title and abstract screening was subsequently performed, followed by a full-text screening phase to determine which articles would be included in the review synthesis. To minimise risk of bias, two reviewers (MS and LA) completed the title and abstract screening and the full-text screening, independently, and in duplicate. Reviewers met following each screening phase to discuss

cases of disagreement and achieve a consensus. A strategy was agreed for articles where

consensus was not achieved, whereby the lead reviewer would make the final decision;

however, this did not need to be actioned. An expert panel (GF and PT) was also available to

discuss the findings as the review progressed.

Table 4.1 Base search - CINAHL Plus via EBSCO

(CINAHL Plus via EBSCO) \<searched on 02/03/2017\>

1. (Professionali?m OR professionali?ation OR unprofessional*).m_titl.

2. ("professional competenc*" or "professional skill*" or "professional value*" or "professional role*" or "professional attitude*" or "professional identit*" or "professional practice*" or "professional communication*" or "professional standard*" or "professional accountab*" or "professional dissonanc*" or "professional impair*" or "professional dysfunction*" or "professional malpractice*" or "professional misconduct*" or "professional omission*").m_titl.

3. ((Professiona* ADJ3 (issue* OR behav* OR act* OR ethic* OR humanis*)) NOT (Professional ADJ3 activ*)).m_titl.

4. *professionalism/

5. 1 OR 2 OR 3 OR 4

6. ("mental health" or psychiatr* or "learning disabilit*" or "learning difficult*" or "learning disorder*" or "intellectual disabilit*").ti,ab.

7. (AMHP* or counsell* or RMN* or psychotherap* or therap*).ti,ab.

8. mental health/

9. psychiatry/

10. learning disorders/

11. intellectual disability/

12. 6 OR 7 OR 8 OR 9 OR 10 OR 11

13. ("physical therap*" OR "occupational therap*").ti,ab.

14. 12 NOT 13

15. 5 AND 14

16. limit 15 to (english language and yr="2006 -Current")

4.6.1 Titles and abstracts

Once all duplicates were removed by the lead reviewer, who used EndNote X8 for the

process, a spreadsheet was developed using Microsoft Excel (version 2016). The

spreadsheet incorporated 1) a list of titles and abstracts of the articles to be screened, 2) the review protocol, 3) inclusion and exclusion criteria, and 4) a list of high-income countries. Titles and abstracts were reviewed independently and in duplicate by LA and MS and screening responses were documented by each reviewer on an Excel spreadsheet. Screening responses were determined by whether an article met the inclusion criteria. Once all titles and abstracts had been screened, a meeting took place between the reviewers (LA & MS) to discuss cases of disagreement. Following the meeting, all articles achieved consensus regarding their inclusion in the next phase.

4.6.2   Full text articles

All articles that were to be included in the review were sought by the lead reviewer using credentials from the University of York and the University of Hull. Those that were freely accessible were retrieved and shared with the review team. A further Microsoft Excel spreadsheet (version 2016) was developed with the list of papers for full-text screening. For this phase, the spreadsheet incorporated an additional column, whereby reviewers were able to provide a rationale for choosing to exclude items; the column had various drop-down options that could be selected, which were developed using the exclusion criteria. Papers were again reviewed independently and in duplicate by both reviewers (LA and MS) against the PICOS criteria. A further meeting took place to discuss cases of disagreement related to this phase. Following discussion, consensus was achieved on all records and a list of articles to be included in the review was finalised.

4.7   Data extraction

A standardised template was developed for the extraction of data, which incorporated the following headings: the article type, any definition or description of professionalism provided, the journal of publication, the author's profession and country of residence, a summary of the article, a brief description of how the article related to the mental health service context, information on who's views were portrayed within the article, and any declaration of interest noted. Data was extracted by the first author (LA) and stored on an

electronic data collection sheet, again using Microsoft Excel. Due to the qualitative nature of the review question, the lead reviewer (LA) included an extra column for any additional information deemed relevant, or for themes that were observed in an article. LA extracted all data; MS reviewed the data extracted for the initial seven articles (10%).

4.8   Critical appraisal

A pragmatic decision was made to not perform a critical appraisal of the literature; this decision was made due to the limited time resource and the integrative nature of the review, which resulted in many theoretical papers.

4.9   Data analysis

A narrative synthesis was performed, as this approach is particularly useful for studies that are insufficiently similar (Popay et al., 2006). The four stages of narrative synthesis comprise of: 1. Developing a theoretical model; 2. Developing a Preliminary synthesis; 3. Exploring relationships in the data; and 4. Assessing the robustness of the synthesis product. Thematic analysis was performed on the data to help synthesise the findings. Thematic analysis (Braun and Clarke, 2006) was undertaken by the lead reviewer and commenced during the data extraction phase. At this stage, the lead reviewer (LA) began to code values and attributes of professionalism dependent on what appeared most relevant regarding the research question (using Microsoft Excel). As articles were screened before data extraction, the lead reviewer was already familiar with the dataset. It was helpful that the lead reviewer was involved in both of these stages as the rapid nature of the review constricted the amount of time available for continuously re-reading the data.

Once data extraction had been completed, the lead reviewer printed off all definitions and descriptions, and began to code these manually by placing them into similar groups. Following this, data was uploaded onto nVivo (version 11), which assisted with the management of data. Within nVivo, codes (nodes) were generated, and sections of the text were highlighted based on early analyses of the data. These were developed to cover the

whole dataset. Codes were revised upon reading and rereading the data, alongside drafting an initial report. This allowed the lead reviewer to become more fully immersed in the dataset. This is in line with recommendations by Braun and Clarke (2006), who note that writing is an integral part of thematic analysis. Hierarchies were established using nVivo, which incorporated themes and subthemes. Thematic maps were then generated by the lead reviewer, whilst comparing and contrasting themes. Themes were continually revised by the lead reviewer until the themes were considered to accurately capture the data extracted. To limit researcher bias (Bucci et al., 2015), two topic experts (PT, GF) were consulted and themes were revised and refined through negotiation.

4.10 Results

A total of 70 articles that defined or described professionalism in a mental health services context were included in the review. 1,184 articles were initially identified from the various resources, which included CINAHL (219), Medline (161), EMBASE (303), PsycINFO (474), HMIC (19), and eight additional sources identified via snowballing. After removing duplicates using bibliographic software (EndNote X8), 779 records remained. Title and abstract screening resulted in 573 articles being excluded, because they did not meet the PICOS criteria. From 206 full-text articles, an additional 136 were excluded, resulting in 70 articles being included in the review (see Figure 4.1 for PRISMA flow diagram; Moher et al., 2009).

Identification

Records identified through database searching
CINAHL (219)
Medline (161)
EMBASE (303)
PsycINFO (474)
HMIC (19)

Additional records identified through other sources (n=8)

Screening

Records after duplicates removed (n=779)

Titles and abstracts screened (n=779)

Records excluded (n=573)

Eligibility

Full-text records assessed for eligibility (n=206)

Full-text records excluded (n=136)
Ineligible population (n=3)
Not mental health services (n=32)
Ineligible MH services (n=3)
Lack of content or definition (n=42)
No attribution to concept (n=20)
Not freely accessible (n=24)
Book, book review, or thesis (n=9)
Correction paper (n=1)
Not in English (n=2)

Included

Records included in narrative synthesis (n=70)

Figure 4.1 PRISMA flow diagram

Articles included in the narrative synthesis incorporated 1 meta-ethnography, 24 discussion papers, 20 editorial/opinion pieces, and 25 empirical studies. Of the latter, there were 15 papers that reported on quantitative research, 7 reporting on qualitative findings, and 3 discussing a mixed methods approach (see Table 4.2).

Table 4.2 Table of characteristics

Article	Type of article	Journal / Source	Profession of author(s)	Country of authorship	Reference made to the mental health services context
Ljungberg, Denhov & Topor (2015) The Art of Helpful Relationships with Professionals: A Meta-ethnography of the Perspective of Persons with Severe Mental Illness	Review article	Psychiatric Quarterly	Social Work / Psychiatry	Sweden / Norway	Review focuses on helpful relationships for people with Serious Mental Illness
Bhugra (2008b) Renewing psychiatry's contract with society	Discussion paper	Psychiatric Bulletin	Psychiatry	UK	Focus on psychiatry's contract with society
Bhugra & Gupta (2011) Alienist in the 21st century	Discussion paper	Asian Journal of Psychiatry	Psychiatry	UK	Focus on the history of the psychiatric profession
Bouras & Ikkos (2013) Ideology, psychiatric practice and professionalism	Discussion paper	Psychiatriki	Psychiatry	UK	Focus on the psychiatric profession with regards to ideology
Brendel et al. (2007) The price of a gift: an approach to receiving gifts from patients in psychiatric practice	Discussion paper	Harvard Review of Psychiatry	Psychiatry	USA	Focus on the psychiatric profession and ethical dilemmas regarding gifts
Coverdale (2007) Virtues-based advice for beginning medical students	Discussion paper	Academic Psychiatry	Psychiatry	USA	Whilst discussing virtues, reports on a survey with the academic psychiatry editorial board

De Waal, Malik & Bhugra (2010) The psychiatric profession: an expertise under siege?	Discussion paper	International Journal of Social Psychiatry	Psychiatry	UK	Focus on threats to the psychiatric profession
Dingle & Stuber (2008) Ethics education	Discussion paper	Child and Adolescent Psychiatric Clinics of North America	Psychiatry	USA	Focus on ethics, specifically within child and adolescent mental health services
Elman & Forrest (2007) From trainee impairment to professional competence problems: Seeking new terminology that facilitates effective action	Discussion paper	Professional Psychology-Research and Practice	Counselling Psychology	USA	Focuses on issues with terminology, across counselling psychology
Fay (2013) The Baby and the Bathwater: An Unreserved Appreciation of Nick Totton's Critique of the Professionalisation of Psychotherapy	Discussion paper	Psychotherapy and Politics International	Clinical psychology / Psychotherapy	New Zealand	Focus on the professionalisation of psychotherapy
Gottlieb, Younggren & Murch (2009) Boundary Management for Cognitive Behavioral Therapies	Discussion paper	Cognitive and Behavioral Practice	Psychology/ Psychotherapy	USA	Focus on ethical issues with cognitive behavioural therapy
Haverkamp et al. (2011) Professional Issues in Canadian Counselling Psychology: Identity, Education, and Professional Practice	Discussion paper	Canadian Psychology-Psychologie Canadienne	Counselling Psychology	Canada	Focus on the identity of Canadian counselling psychologists
Ikkos & Mace (2009) Professionalising psychotherapy: Lessons from the development of psychiatry	Discussion paper	European Journal of Psychotherapy and Counselling	Psychiatry / Psychotherapy	UK	Focus on the professionalisation of psychotherapy
Jakovljevic (2012) Professionalism in psychiatry and medicine: a hot topic	Discussion paper	Psychiatria Danubina	Psychiatry	Croatia	Focus on professionalism in psychiatry
John et al. (2016) Training Psychiatry Residents in Professionalism in the Digital World	Discussion paper	Psychiatric Quarterly	Psychiatry	UK	Focus on digital media within the psychiatric profession

Mendelberg (2014) The integration of professional values and market demands: A practice model	Discussion paper	The Psychologist-Manager Journal	Clinical Psychology	USA	Talks about a private practice developed to serve those with mental illness
Paprocki (2014) When Personal and Professional Values Conflict: Trainee Perspectives on Tensions Between Religious Beliefs and Affirming Treatment of LGBT Clients	Discussion paper	Ethics & Behavior	Psychology	USA	Discusses ethical issues related to the delivery of psychological therapy for LGBT clients
Peek, H. S. et al. (2015) Blogging and Social Media for Mental Health Education and Advocacy: a Review for Psychiatrists	Discussion paper	Current Psychiatry Reports	Psychiatry	USA	Focus on digital media within the psychiatric profession
Randall & Kindiak (2008) Deprofessionalization or Postprofessionalization? Reflections on the State of Social Work as a Profession	Discussion paper	Social Work in Health Care	Social Work	Canada	Focus on the professionalisation of social work
Roberts & Termuehlen (2013) (Honest) letters of recommendation	Discussion paper	Academic Psychiatry	Psychiatry	USA	Focus on psychiatric issues
Robertson & Walter (2007) Overview of psychiatric ethics I: Professional ethics and psychiatry	Discussion paper	Academic Psychiatry	Psychiatry	Australia	Focus on ethics in psychiatry
Sanders, Servis & Boland (2014) The four general competencies	Discussion paper	Academic Psychiatry	Psychiatry	USA	Focus on competencies in the psychiatric profession
Schreiber et al. (2016) The Patient-Psychiatrist Relationship on the Axis of the Other and the Same	Discussion paper	Psychiatric Quarterly	Psychiatry	Israel	Discusses the patient / psychiatrist relationship

Schwartz, Kotwicki & McDonald (2009) Developing a modern standard to define and assess professionalism in trainees	Discussion paper	Academic Psychiatry	Psychiatry	USA	Minimal reference made to the field of mental health, but authors work in psychiatry and article published within a psychiatric journal
Young et al. (2013) The EAP Project to establish the professional competencies of a European psychotherapist	Discussion paper	International Journal of Psychotherapy	Psychotherapy	International	Focus on competencies in psychotherapy
Bhugra (2009) Professionalism and psychiatry: past, present, future	Editorial / Opinion piece	Australas Psychiatry	Psychiatry	UK	Focus on the psychiatric profession
Bhugra (2010) Editorial: Teaching Professionalism in Psychiatry	Editorial / Opinion piece	International Journal of Social Psychiatry	Psychiatry	UK	Focus on professionalism in psychiatry
Bhugra & Brown (2007) Editorial: psychiatry: de-professionalisation	Editorial / Opinion piece	International Journal of Social Psychiatry	Psychiatry	UK	Focus on threats to the psychiatric profession
Bhugra & Gupta (2010) Medical professionalism in psychiatry	Editorial / Opinion piece	Advances in Psychiatric Treatment	Psychiatry	UK	Focus on professionalism in psychiatry
Brown & Bhugra (2007) 'New' professionalism or professionalism derailed	Editorial / Opinion piece	Psychiatric Bulletin	Psychiatry	UK	Focus on professionalism in psychiatry
Coverdale, Balon & Roberts (2011) Cultivating the professional virtues in medical training and practice	Editorial / Opinion piece	Academic Psychiatry	Psychiatry	USA	Editorial for an issue in academic psychiatry

Gosselink & de Man (2012) The psychiatric scrapbook: fantasizing from the patient's perspective	Editorial / Opinion piece	Educ Health (Abingdon)	Psychiatry	Netherlands	Focus on a teaching programme in psychiatry
Grounds et al. (2010) Contemplating common ground in the professional ethics of forensic psychiatry	Editorial / Opinion piece	Criminal Behaviour and Mental Health	Forensic Psychiatry	USA / UK	Discusses ethics relating to forensic psychiatry
Happell (2006) Would the real mental health nurse please stand up? The relationship between identification and professional identity	Editorial / Opinion piece	International Journal of Mental Health Nursing	Nursing	UK	Discusses psychiatric nurses in comparison to nurses in other specialties
Ikkos, McQueen & St. John-Smith (2011) Psychiatry's contract with society: What is expected?	Editorial / Opinion piece	Acta Psychiatrica Scandinavica	Psychiatry	UK	Focus on psychiatry's contract with society
Lampshire (2012) Living the dream	Editorial / Opinion piece	Psychosis-Psychological Social and Integrative Approaches	Former service user	New Zealand	Talks about their personal experience as a former mental health service user
Malhi (2008) Professionalizing psychiatry: from 'amateur' psychiatry to 'a mature' profession	Editorial / Opinion piece	Acta Psychiatrica Scandinavica	Psychiatry	Australia	Focus on psychiatry and professionalisation
Malone (2012) Ethical professional practice: exploring the issues for health services to rural Aboriginal communities	Editorial / Opinion piece	Rural Remote Health	Psychology	Canada	Discusses ethical issues related to working in aboriginal communities
Peek (2014) Psychiatry and Professionalism in the Digital Age	Editorial / Opinion piece	Psychiatric Times	Psychiatry	USA	Focus on digital media within the psychiatric profession

Poole & Bhugra (2008) Editorial: Should psychiatry exist?	Editorial / Opinion piece	International Journal of Social Psychiatry	Psychiatry	UK	Focus on the profession of psychiatry
Roberts (2009) Professionalism in psychiatry: a very special collection	Editorial / Opinion piece	Academic Psychiatry	Psychiatry	USA	Editorial for an issue in academic psychiatry
Rogers (2009) Dare we do away with professionalism?	Editorial / Opinion piece	Therapy Today	Counselling	UK	Discusses how professionalisation would be detrimental to the counselling practice
Scott Johnson, Chiu & Czelusta (2015) For residents, technology can put professionalism and reputation at risk	Editorial / Opinion piece	Current Psychiatry	Psychiatry	USA	Focus on digital media within the psychiatric profession
Talbott & Mallott (2006) Professionalism, medical humanism, and clinical bioethics: The new wave-does psychiatry have a role?	Editorial / Opinion piece	Journal of psychiatric practice	Psychiatry	USA	Focus on psychiatry
Wise (2008) Competence and scope of practice: ethics and professional development	Editorial / Opinion piece	Journal of Clinical Psychology	Psychotherapy	USA	Focus on the practice of psychotherapy and mental health
Baer & Schwartz (2011) Teaching professionalism in the digital age on the psychiatric consultation-liaison service	Quantitative study	Psychosomatics	Psychiatry	USA	Focus on digital media within the psychiatric profession
Goodyear et al. (2016) A global portrait of counselling psychologists' characteristics, perspectives, and professional behaviors	Quantitative study	Counselling Psychology Quarterly	Counselling psychology	International	Focuses on counselling psychology internationally

Harris & Kurpius (2014) Social Networking and Professional Ethics: Client Searches, Informed Consent, and Disclosure	Quantitative study	Professional Psychology-Research and Practice	Counselling Psychology	USA	Reports on a survey with counselling and psychology graduate students
Jain et al. (2010) Psychiatry Residents' Attitudes on Ethics and Professionalism: Multisite Survey Results	Quantitative study	Ethics & Behavior	Psychiatry	USA	Reports on a survey with psychiatry residents
Jain, Lapid, et al. (2011) Psychiatric residents' needs for education about informed consent, principles of ethics and professionalism, and caring for vulnerable populations: results of a multisite survey	Quantitative study	Academic Psychiatry	Psychiatry	USA	Reports on a survey with psychiatry residents
Jain, Dunn, et al. (2011) Results of a multisite survey of U.S. psychiatry residents on education in professionalism and ethics	Quantitative study	Academic Psychiatry	Psychiatry	USA	Reports on a survey with psychiatry residents
Joiner et al. (2015) Medical professionalism education for psychiatry trainees: does it meet standards?	Quantitative study	Australas Psychiatry	Psychiatry	UK	Reports on an audit with psychiatry trainees
Komic, Marusic & Marusic (2015) Research Integrity and Research Ethics in Professional Codes of Ethics: Survey of Terminology Used by Professional Organizations across Research Disciplines	Quantitative study	Plos One	(unclear)	Croatia	Focus on research integrity and ethics codes across organisations, including mental health
Lapid et al. (2009) Professionalism and ethics education on relationships and boundaries: psychiatric residents' training preferences	Quantitative study	Academic Psychiatry	Psychiatry	USA	Reports on a survey with psychiatry residents

Marrero et al. (2013) Assessing professionalism and ethics knowledge and skills: preferences of psychiatry residents	Quantitative study	Academic Psychiatry	Psychiatry	USA	Reports on a survey with psychiatry trainees
Morreale, Balon & Arfken (2011) Survey of the importance of professional behaviors among medical students, residents, and attending physicians	Quantitative study	Academic Psychiatry	Psychiatry	USA	Reports on a survey with psychiatry residents, physicians and trainees
Roberts et al. (2006) Preferences of Alaska and New Mexico psychiatrists regarding professionalism and ethics training	Quantitative study	Academic Psychiatry	Psychiatry	USA	Reports on a survey with psychiatrists
Russinova et al. (2011) Recovery-promoting professional competencies: perspectives of mental health consumers, consumer-providers and providers	Quantitative study	Psychiatric Rehabilitation Journal	Psychiatry	USA	Reports on a survey with consumers, consumer providers and providers of mental health services
Symons et al. (2011) Allegations of serious professional misconduct: An analysis of the British Association for Counselling and Psychotherapy's Article 4.6 cases, 1998–2007	Quantitative study	Counselling and Psychotherapy Research	Counselling / Psychotherapy	UK	Focus on complaints made in counselling and Psychotherapy services
Wu et al. (2012) Professional values and attitude of psychiatric social workers toward involuntary hospitalization of psychiatric patients	Quantitative study	Journal of Social Work	Social Work	Taiwan	Reports on a survey with social workers about psychiatric detention
Bhugra (2008a) Professionalism and psychiatry: the profession speaks	Mixed methods	Acta Psychiatrica Scandinavica	Psychiatry	UK	Reports on a survey with psychiatrists re professionalism in psychiatry

Leppma et al. (2016) Working With Veterans and Military Families: An Assessment of Professional Competencies	Mixed methods	Professional Psychology-Research and Practice	Counselling	USA	Discusses competencies needed for working with veterans in mental health
Sims (2011) Reconstructing professional identity for professional and interprofessional practice: a mixed methods study of joint training programmes in learning disability nursing and social work	Mixed methods	Journal of Interprofessional Care	Health and Social Care	UK	Reports on a survey / interviews with dual trained learning disability nurses / social workers
Alves & Gazzola (2013) Perceived professional identity among experienced Canadian counsellors: A qualitative investigation	Qualitative study	International Journal for the Advancement of Counselling	Counselling	Canada	Reports on a study performed with counsellors that work in mental health
Blegeberg, Bloomberg & Hedelin (2008) Nurses' conceptions of the professional role of operation theatre and psychiatric nurses	Qualitative study	Nordic Journal of Nursing Research & Clinical Studies / Vård i Norden	Nursing	Sweden	Reports on interviews with nurses regarding the psychiatric nursing role
Coy, Lambert & Miller (2016) Stories of the Accused: A Phenomenological Inquiry of MFTs and Accusations of Unprofessional Conduct	Qualitative study	Journal of Marital and Family Therapy	Marriage and Family Therapy Counselling	USA	Reports on interviews with marriage and family therapists that were accused of misconduct
Crawford, Brown & Majomi (2008) Professional identity in community mental health nursing: a thematic analysis	Qualitative study	International Journal of Nursing Studies	Nursing / Psychology	UK	Interviewed community mental health nurses about their role
Gonyea, Wright & Earl-Kulkosky (2014) Navigating dual relationships in rural communities	Qualitative study	Journal of Marital and Family Therapy	Marriage and Family Therapy Counselling	USA	Focus on marriage and family therapy

Groves & Kerson (2011) The Influence of Professional Identity and the Private Practice Environment: Attitudes of Clinical Social Workers Toward Addressing the Social Support Needs of Clients	Qualitative study	Smith College Studies in Social Work	Social Work	USA	Reports on interviews and focus groups with social workers regarding social support for patients
Pelto-Piri, Engstrom & Engstrom (2012) The ethical landscape of professional care in everyday practice as perceived by staff: A qualitative content analysis of ethical diaries written by staff in child and adolescent psychiatric in-patient care	Qualitative study	Child and Adolescent Psychiatry and Mental Health,	Psychiatry	Sweden	Reports on a study with various occupational staff working in child and adolescent mental health services

Of the 70 papers included in the review, 44 were written by psychiatrists; these psychiatrists included Dinesh Bhugra and Laura Weiss Roberts, whom were the two most predominant authors identified by the review. Additional authors incorporated psychologists, counsellors, nurses, social workers, and therapists (see Table 4.2). It is worth noting that the nursing and social work professions were underrepresented in the literature; four papers were authored by social workers, and three papers were authored by nurses. A former service user had also authored an article providing an account of their own experience of using mental health services. Most articles were written in the United States of America (USA; n=32), or the United Kingdom (UK; n=21), with additional articles being authored by individuals from Canada, Australia, Croatia, New Zealand, Sweden, Norway, Taiwan, Israel, and Netherlands (see Table 4.2). Whilst the search strings incorporated keywords for learning disability services, none of the papers discussed this area and the results are thus presented for mental health services only.

4.10.1 Main findings

Various observations were made during the thematic analysis and are reported here.

4.10.1.1 *An abstract construct*

Professionalism was described as an abstract construct that is often misunderstood (John et al., 2016, Brown and Bhugra, 2007). Thematic analysis found that professionalism was conceptualised on two levels; first, at a societal level; and second, at an individual level (see Figure 4.2). On a societal level, professionalism was viewed as the basis of a *dynamic social contract* between professions and society, that: a) *requires purpose*, and (b) *evolves* as the needs and expectations of society change. On an individual level, professionalism was described as '*individuals representing the profession*,' possessing a) '*intrapersonal professionalism*,' b) '*interpersonal professionalism*,' and c) '*working professionalism*.' Whilst these are identified as separate themes, it is important to note that these are not mutually exclusive. Each subtheme has many associated elements as listed in Table 4.3.
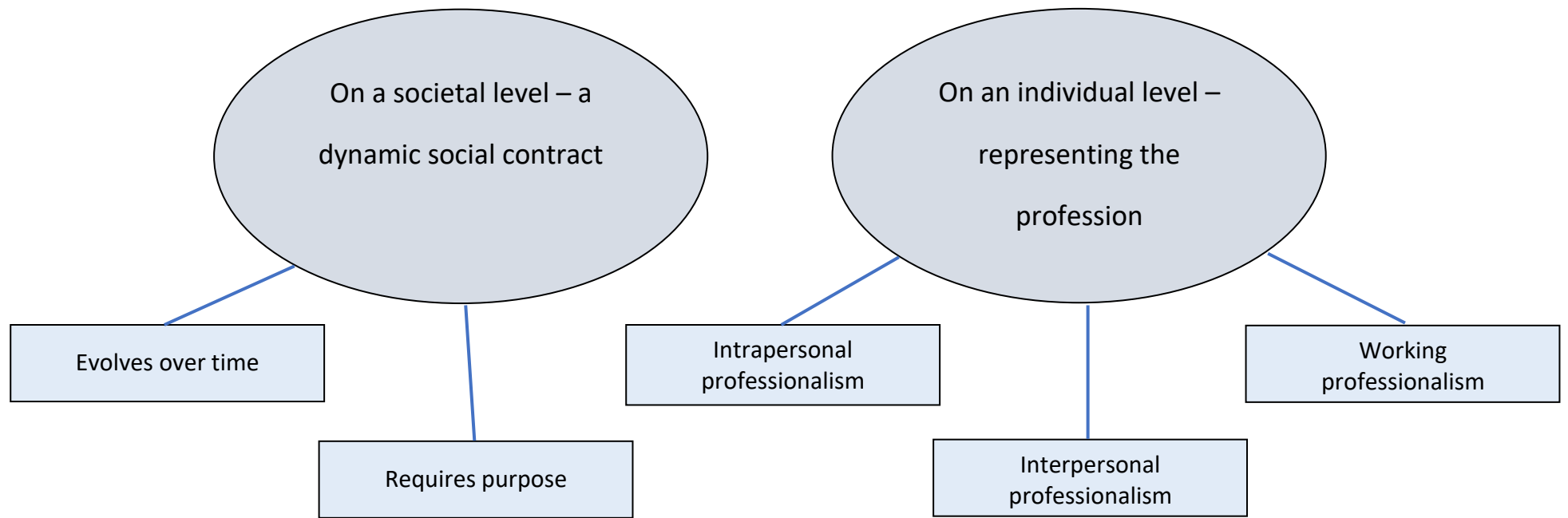
Figure 4.2 Thematic map of professionalism in mental health services

Table 4.3 Themes, subthemes, and elements of professionalism in a mental health service context

Themes	Subthemes	Associated elements
On a societal level- a dynamic social contract	Requires purpose	A profession's mission and core values are established; and professional identity is generated via knowledge, skills and expertise.
	Evolves over time	The social contract must be renegotiated regularly between professions and society
On an individual level- representatives of the profession	Intrapersonal professionalism	Expectations held of individuals in order to meet the expectations of their profession *Honours professional codes of conduct, self, and others, demonstrates a commitment to professional and ethical practice, acting professionally at all times, demonstrates core values including probity, objectivity, courage, and truthfulness, integrity, and self-sacrifice, possesses self-awareness and self-discipline, is responsible and accountable to self and the profession, possesses appropriate knowledge and skills for role and self-regulates, having a commitment to continuing professional development, monitors own wellbeing and possesses a secure, stable, calm and confident persona.*

Interpersonal professionalism	Possessing the necessary skills to relate to others in an appropriate manner.
	Acts in the patient's best interest at all times, facilitates but does not dictate treatment, provides patients with hope and positive feedback, portrays genuine respect, having trust, benevolence, honesty, altruism, respect, self-effacement, compassion and motivation to help others, values the worth and dignity of all and treats all patients equally, possesses concern for others' welfare, has a humane and personal nature with an empathic and diplomatic ability, places an emphasis on the practitioner-client relationship, forming effective therapeutic relationships, focuses on patients' strengths and assets, offering choices, facilitating self-worth and promoting empowerment, listens without judgement and believes in recovery, demonstrates cultural awareness and competence, understands the importance of communication and interpersonal skills, maintains healthy relationships with others and appropriate boundaries, promotes patients human rights, contains the anxieties of patients and colleagues, and shares expertise via training, research and policy development.
Working professionalism	Ability to form judgements and act accordingly, thinking critically and using reflection in action
	Demonstrates critical thinking skills and acts wisely, possessing practical wisdom and using professional judgement in situations of uncertainty, ambiguity, and/or instability, acts in the patient's best interest and adheres to professional and ethical consensus, using knowledge in action, reflection in action and reflection in learning, demonstrates an openness, willingness, and flexibility in practice, reflects critically on practice and is responsive, having social perceptiveness and emotional resonance, manages transference and counter-transference accordingly, demonstrates appropriate action when societal, cultural and ethical obligations diverge.

4.10.1.2 *Professionalism at a societal level – a dynamic social contract with society*

Thematic analysis of the literature determined that professionalism was viewed as the basis of a dynamic social contract between professions and society. This theme was initially considered to incorporate three subthemes, including 1) Power and purpose, 2) Bidirectional expectations, and 3) Change and variability; however, on further analysis, the subthemes 'Power and purpose' and 'Bidirectional expectations' were amalgamated into the subtheme 'Purpose.' Henceforth, the results demonstrated that a social contract a) requires *purpose*, and b) e*volves over time.* Thematic analysis itself is a process that develops over time (Braun and Clarke, 2006); given the rapid nature of the review it is not surprising that further analyses resulted in revisions to the original themes generated from the data.

Requires purpose

It is widely accepted that professionalism forms the basis of a social contract between professions and society, and this was manifest throughout the literature (Bhugra, 2008b; 2009; Bhugra and Gupta, 2010; 2011; Bouras and Ikkos, 2013; Brendel et al., 2007; Brown and Bhugra, 2007; Ikkos and Mace, 2009; Jain et al., 2011b; Komic et al., 2015; Lapid et al., 2009; Randall and Kindiak, 2008; Roberts, 2009; Robertson and Walter, 2007). The Charter on medical professionalism notes that the three fundamental principles of professionalism are the 1) primacy of patient welfare, 2) patient autonomy, and 3) social justice (Bhugra and Gupta, 2010; 2011; Brown and Bhugra, 2007; Jakovljević, 2012; John et al., 2016). Practitioners must be altruistic, serve for the benefit of others and promote public good (Bhugra, 2008b; 2009; Bhugra and Gupta, 2010; Ikkos and Mace, 2009; Randall and Kindiak, 2008; Roberts, 2009; Robertson and Walter, 2007). Practitioners should benefit the public at large through the provision of high quality care, development of policy, research, and training (De Waal et al., 2010). Professions are expected to self-regulate and expand their skills and knowledge (Bhugra, 2008b; Bhugra and Gupta, 2010; Roberts, 2009; Robertson and

Walter, 2007), but they must have trust (Bhugra, 2008b; Randall and Kindiak, 2008), gain autonomy (Bhugra, 2008b; Randall and Kindiak, 2008; Robertson and Walter, 2007), and be given the opportunity to do so (Bhugra, 2008b; Randall and Kindiak, 2008; Robertson and Walter, 2007).

To fulfil their side of the contract, professions expect that patients will strive to tell the truth about their illness and adhere to treatment (Brendel et al., 2007; Robertson and Walter, 2007). The profession expects participation in public policy and shared responsibility for health, as well as financial and nonfinancial rewards (Bhugra, 2008b). In addition, to enable a suitable service, the profession expects the right funding and a value-driven healthcare system (Bhugra, 2008b). This contract allows patients to receive relief from their suffering, whilst practitioners receive a wage and are rewarded for the outcomes of their work (Brendel et al., 2007).

The acquisition of skills is fundamental to developing professionalism (Malhi, 2008). For example, Lopez, as highlighted by Haverkamp et al., argues that the deliberate focus on individual strengths and assets in counselling psychology, has helped the specialty maintain its integrity and identity within professional psychology (Haverkamp et al., 2011). In contrast, there is a lack of clarity regarding the nursing role (Blegeberg et al., 2008; Happell, 2006), which poses the profession with difficulty in promoting and justifying their service for today's healthcare system (Crawford et al., 2008).

Evolves over time

The data highlighted that professionalism is a multidimensional (Sims, 2011), dynamic construct that evolves over time and across professions (Bhugra, 2008b; 2010; Bhugra and Gupta, 2010; Brown and Bhugra, 2007; Malhi, 2008). Professionals are having to adhere to more prescriptive policies, increased regulation, and varied expectations of their role, which results in a gradual reduction of their autonomy (Brown and Bhugra, 2007). Ethics codes are amended and updated (Gottlieb et al., 2009; Harris and

Kurpius, 2014), as variations in policy and regulations are influenced by subtle variations in individuals' beliefs and understanding (Malhi, 2008).

Roles originally performed by psychiatrists, such as medical prescribing, are now being filled by other professions (Malhi, 2008). Community mental health nurses, also, have been seen to develop skills from outside of their profession to '*make themselves more professional as nurses*' (Crawford et al., 2008, p.1060).

4.10.1.3 *Professionalism at an individual level - representing the profession*

At an individual level, professionalism was manifest as representing the profession, possessing; a) '*Intrapersonal professionalism,*' b) '*Interpersonal professionalism,*' and c) '*Working professionalism.*'

Intrapersonal professionalism

Intrapersonal professionalism refers to professionals having the necessary attributes to conduct themselves accordingly and act as 'representatives of the profession, its mission, and its core values' (Groves and Kerson, 2011). Coverdale argues the importance of personality in the practice of a professional, noting 'the most powerful tool you have is you' (Coverdale, 2007). Patients desire professionals that are secure, stable, and calm, as well as confident in times of distress (Ljungberg et al., 2015). Professionals must have integrity and morality (Bhugra, 2008b; 2009; Bhugra and Brown, 2007; Ikkos and Mace, 2009; Randall and Kindiak, 2008; Robertson and Walter, 2007; Talbott and Mallott, 2006) and be devoted and committed to their work (Bhugra, 2008b; Ikkos and Mace, 2009; Randall and Kindiak, 2008; Robertson and Walter, 2007). They must be accountable for their actions (Bhugra, 2008a; 2008b; 2010; Brown and Bhugra, 2007; Ikkos and Mace, 2009; John et al., 2016; Joiner et al., 2015; Malhi, 2008; Randall and Kindiak, 2008; Robertson and Walter, 2007; Sanders et

al., 2014; Schwartz et al., 2009) and are expected to adhere to standards (Baer and Schwartz, 2011; Bhugra, 2008b). Professionals must self-regulate and keep their continuing professional development up to date (De Waal et al., 2010), being adequately trained (Happell, 2006; Wise, 2008), and having the right knowledge and skills to undertake their role (Bhugra, 2008b; 2009; Bhugra and Brown, 2007; Bhugra and Gupta, 2010; Brendel et al., 2007; Ikkos and Mace, 2009, Randall and Kindiak, 2008, Robertson and Walter, 2007). In addition, practitioners must be culturally competent (Leppma et al., 2016; Malone, 2012) and gain the trust of the community by being in a professional role at all times (Australian Psychological Society, 2004; as cited in Malone, 2012; Schank and Skovholt, 2006).

Interpersonal professionalism

Interpersonal professionalism refers to professionals possessing the necessary skills to relate to others in an appropriate manner. A former service user highlighted that during her time in mental health services, she believed that professionals wanted a monument, to evidence their talents, and that they had been less interested in getting to know her. She argued that they didn't have time to get to know her because they had to complete paperwork, attend meetings, and fulfil their personal development requirements (Lampshire, 2012). Patients need mental health professionals whom they can have a relationship, be offered real choices, and who will facilitate, but not dictate treatment (De Waal et al., 2010).

Communication and interpersonal skills are important (Brendel et al., 2007; Dingle and Stuber, 2008; Elman and Forrest, 2007; Haverkamp et al., 2011; Ikkos and Mace, 2009; Roberts et al., 2006; Sanders et al., 2014; Talbott and Mallott, 2006; Wise, 2008), and clinicians must have the ability to engage patients (Malhi, 2008). Professionals must demonstrate respect for patients, carers and colleagues (Brendel et al., 2007; Talbott

and Mallott, 2006), and must also demonstrate moral and ethical behaviour (Bhugra, 2009; Bhugra and Brown, 2007; Talbott and Mallott, 2006), be honest (Bhugra and Brown, 2007; Talbott and Mallott, 2006) and act with compassion (Brendel et al., 2007; Talbott and Mallott, 2006). Humanity and personal nature are viewed as important as other intellectual strengths (Roberts and Termuehlen, 2013), given that the practitioner-patient relationship is fundamental to both professionalism and ethics (Schreiber et al., 2016). Another highly regarded skill in psychiatry is the ability to deal with transference and counter-transference (Ikkos et al., 2011).

Professionals must adhere to confidentiality guidelines (Baer and Schwartz, 2011; Talbott and Mallott, 2006), maintain appropriate boundaries with patients, and have healthy relationships with colleagues and trainees (Jain et al., 2010; 2011a; 2011b; Lapid et al., 2009; Sanders et al., 2014; Schwartz et al., 2009). Boundaries must extend beyond the clinical setting and include online behaviours also (Peek, 2014; Peek et al., 2015). Clinicians must treat patients, carers and colleagues equally, regardless of their age, gender, sexual orientation, ethnicity, cultural and socioeconomic background, or their religion (Ikkos and Mace, 2009; Schwartz et al., 2009; Talbott and Mallott, 2006). It is ethically inappropriate for mental health professionals to work with clients of diverse backgrounds without the appropriate training or competence (Sue et al., 1992). Professionals must respect and support the human rights of patients also (Jakovljević, 2012), focusing on their welfare and social inclusion (Ikkos, 2010; as cited in Bouras and Ikkos, 2013),

Various factors were noted in a meta-ethnography, by Ljungberg et al., to contribute to helpful relationships for people with serious mental illness, which include mutually trusting relationships, professionals having the knowledge, skills, and competence

suggestive of the role, and sharing relevant information with patients (e.g., information about medications or an individual's diagnosis). Patients were grateful of the opportunity to talk to professionals and share things, expecting them to understand more than others, and not fearing that the discussion would be repeated due to confidentiality. Whilst the term shared-decision making was not used, Ljungberg (2015) referred to research that suggested professionals and patients acting like allies, and working together, was also helpful. Professionals exerting their influence on other professionals or organisations often helped individuals get the support they needed.

The relationship between a professional and a client is frequently found to be more influential on positive outcomes than technical interventions or treatments (Elman and Forrest, 2007; Ikkos et al., 2011). Bhugra and Brown (2007) note that whilst diagnosis, investigation, and treatments are a way to a means, the professional values are instrumental to this and must be strengthened.

Working professionalism

Working professionalism refers to practitioners' ability to form appropriate judgements, act accordingly, think critically, and use reflection-in-action. Practical wisdom is the central virtue of professionalism, according to Racey, which he argues unify the moral and intellectual virtues (as cited by Bhugra, 2010). There is significant discussion regarding ethical dilemmas in the literature, which focus on issues such as boundary violations (Dingle and Stuber, 2008), involuntary treatment (Wu et al., 2012), receipt of gifts (Brendel et al., 2007), and the provision of certain therapies (Gottlieb et al., 2009). Mental health professionals often face unpredictable situations, and ethical codes may be ambiguous, thus requiring practitioners to use their clinical judgement (Malhi, 2008).

Acts of professionalism vary dependent on context (Brendel et al., 2007; Harris and Kurpius, 2014; Ikkos and Mace, 2009; Malhi, 2008; Wise, 2008). As technology progresses and digital media is increasingly used, further ethical dilemmas present and boundaries regarding what is professional and what is personal become blurred (Scott Johnson et al., 2015). Professionals' views may differ regarding what level of professional distance is appropriate (Pelto-Piri et al., 2012), and professionals are frequently exposed to situations that are ambiguous in nature and may call for one of many alternative responses. Professionals must be mindful of the limits to their competence, particularly when managing difficult or complex cases (Gottlieb et al., 2009). Practitioners are expected to reflect on their practice regularly, which would include having 'knowledge-in-action', 'reflection-in-action' and 'reflection-in-learning' (Schon, 1988; as cited in Bhugra, 2008b). Self-awareness and reflective practice help individuals identify areas for their future professional development.

A key trait of psychologists, according to graduate students, is a 'willingness' or 'openness'; one student noted that they would have concerns about the competence of a therapist who was particularly rigid and less flexible in their practice and therapeutic style (Paprocki, 2014). A meta-ethnography conducted by Ljungberg et al. (2015) suggests that whilst the professional role is helpful for individuals with severe mental illness, some helpful actions went beyond professional neutrality and distance. Whilst being rigorous in their ethical commitment, mental health professionals must be able to step outside of this role if it is in the patient's best interest (Brendel et al., 2007).

4.10.1.4 *Differences among the professions in their conceptualisation of professionalism*

It is difficult to derive any firm conclusions regarding how professions vary in their conceptualisations of professionalism, because the literature was dominated by the psychiatric, counselling, and psychological professions; limited papers expressed the opinion of social workers, nurses, and patients.

McQueen et al., previously defined the 7 E's of psychiatric professionalism, as: attention to *evidence, emotions, ethics, engagement, expertise, education* and research for future care, and a commitment to the *empowerment* of patients (as cited by Ikkos et al., 2011). Psychotherapy was argued to differ from psychiatry, and one of these differences included the emphasis placed on the psychotherapist in the practitioner-patient relationship (Ikkos and Mace, 2009). A study by Goodyear et al. (2016) highlighted that a key counselling psychology value is their focus on people's strengths and assets. Geller et al. (as cited by Elman and Forrest, 2007) argue that because interpersonal relatedness is core to the psychotherapies, then social perceptiveness, emotional resonance and responsiveness, compassion, motivation to help others, self-awareness, and self-discipline are integral aspects to psychotherapists' work with patients.

The nursing literature typically included studies on professional identity and highlighted a lack of clarity regarding the nursing role (Blegeberg et al., 2008; Crawford et al., 2008; Happell, 2006); the social work literature made note of the dilemmas that social workers face (Randall and Kindiak, 2008; Wu et al., 2012); the importance of interpersonal skills and less rigid practice was highlighted amongst the literature expressing patient opinions and psychological practice (Gottlieb et al., 2009; Haverkamp et al., 2011; Lampshire, 2012; Ljungberg et al., 2015; Paprocki, 2014; Wise, 2008).

4.11   Operational definition of professionalism

Having performed a thematic analysis of the literature, two operational definitions of professionalism are proposed. These definitions apply to all professions working in a mental health services context.

1.   Professionalism forms the basis of a dynamic social contract between professions and society. This contract (which can have both tacit and explicit elements) specifies that society will remunerate the members and permit the profession to self-regulate on the understanding that the profession use their skills for patient and public good.

2.   On an individual level, professionalism can be conceptualised as a latent trait, composed of elements of intrapersonal, interpersonal, and working professionalism. This trait may only be observed through manifest behaviours in certain situations. Such behaviours will be in keeping with society's expectations and demonstrate a commitment to ethical practice, cultural-sensitivity, self-awareness and reflection and self-discipline.

4.12 Discussion

A rapid systematic review was conducted to derive an operational definition of professionalism for a mental health services context. Seventy papers met the inclusion criteria, the majority of which focused on the psychiatric profession, with others discussing psychology, counselling, nursing, and social work. There was limited patient presence in the literature; that is, most articles were written by professionals and, to the current author's knowledge, one paper was written by a 'former' service user. Thematic analysis identified that professionalism was conceptualised on two levels: at a societal level, and at an individual level. At a societal level, professionalism was manifest as a dynamic social contract, which requires purpose and evolves over time. At an individual level, professionalism was viewed as being representatives of the

118

profession, having intrapersonal professionalism, interpersonal professionalism, and working professionalism.

4.12.1 Argument against professionalisation and professionalism

Whilst the discussion of professionalism typically highlighted the expectations on both professions and professionals, arguments were also posed that; 1) professions are in it for themselves; and 2) professionalisation (AKA professionalism) goes against the core values of care and is therefore detrimental to those who need it.

4.12.1.1 *In it for themselves*

Whilst altruism is frequently highlighted in the literature (Bhugra, 2010; Ikkos and Mace, 2009; Schwartz et al., 2009), it was argued that professions may act in their own interests to gain a monopoly of control over service provision, as well as other financial benefits (Freidson, 1970; Larson, 1984; as cited in Randall and Kindiak, 2008; McKinlay, 1973). Randall and Kindiak (2008) suggested that professions use their knowledge as a weapon to maintain their status; professions have argued, however, that it is they who have the knowledge to decide what services should be delivered, and to what standard they should be achieved (Randall and Kindiak, 2008).

4.12.1.2 *Goes against the core principles of care*

Rogers suggests that professionalism (AKA professionalisation) and the regulations attached to it, are at odds with psychotherapy's core values; and instead becomes a network of oppressive conditions that are likely to impact clients (Rogers, 2009). Whilst professionalism may elevate the status and power of a therapist, Rogers states that this it is at the expense of a client's empowerment. The paternalism of psychiatry is argued to be manifested as a doctor's actions without a patient's autonomy, which can at times be defined as a violation of an individual's human rights (Lolas, 2010; as cited in Jakovljević, 2012). Whilst professional regulations may advocate professional boundaries and professional distance, Zur (2007) has argued that excessively rigid practice may in itself be harmful to patients (as cited in Jakovljević, 2012). Given the

nature of external regulation however, it is likely that professionals will become increasingly rigid, for fear of repercussions or allegations of misconduct.

As cited by Ljungberg, a lack of congruity between professionals' conceptualisation of professionalism and what users' want has previously been demonstrated (Hem, 2003; Ljungberg et al., 2015; Moyle, 2003). Whilst the healthcare service is continually trying to make cost-savings, it is argued that efficiency and effectiveness tempts professions onto 'the rocks of a repressive and oppressive economic and social order' (Fay, 2013, p.30).

4.12.2 Time to revisit the contract

Most definitions observed in the literature originated in the medical profession, with a smaller proportion being applied and developed specifically for the mental health services context. This suggests that authors may not have a more pertinent definition available for the mental health setting. Referring to psychiatry, Bhugra and Gupta (2010) note that mental health professionals must be aware of the core attributes of professionalism, but they must amend these to align with changes in societal and patient expectations, as well as the National Health Service and current practice. Other articles also express the need to renegotiate the contract held between professions and society (Bhugra, 2008b; Bouras and Ikkos, 2013; Ikkos and Mace, 2009), which follows an earlier proposal by Cruess, cited by Bhugra (2008b), that if an implicit contract exists, then negotiation of this contract becomes a legitimate professional activity.

Bhugra (2008b) suggested that relationships with other disciplines, colleagues and stakeholders are crucial to the survival of the psychiatric profession, and now is the perfect time for psychiatrists to fight back and to work alongside nurses, with patients,

carers, and other healthcare sectors in asserting their role, enhancing their autonomy and promoting the importance of professionalism (Poole and Bhugra, 2008).

During renegotiations, professions must collaborate with patients and carers, as it is these individuals that professions are there to serve (Brown and Bhugra, 2007). Discussions must look at mutual expectations (Bhugra, 2008b), and acknowledge choice as well as the current personalisation of mental health services (Bouras and Ikkos, 2013; Ikkos et al., 2011; McQueen et al., 2009).

4.12.3 Strengths and limitations

As the review was conducted in a rapid manner, there was limited time available to analyse the data. As noted by Braun and Clarke, thematic analysis is an ongoing process and requires reading and re-reading the data to become fully immersed in the dataset (Braun and Clarke, 2006). The lead reviewer screened articles over the review period and extracted all data; this allowed the reviewer to familiarise themselves with this data before analysis fully commenced. Having a second reviewer also allowed discussion between screening phases, which supported the later analysis and minimised the risk of error and bias,

To facilitate others in conducting a similar review, the methodology and findings are presented in a transparent manner having followed the PRISMA checklist for reporting (see appendix C; Moher et al., 2009). Whilst a rapid approach was deemed most appropriate for the review, this brings its own limitations and a more comprehensive systematic literature review may be needed in the future. Despite this, rapid reviews have been found to generate similar conclusions to full systematic reviews (Watt et al., 2008). Due to the limited time resource, a pragmatic decision was made not to undertake a critical appraisal of the literature. Articles were excluded if not written in the English language or not freely available at time of retrieval. This restricted the

available evidence but was deemed a justifiable decision to perform the review in a time limited manner.

4.12.4 Conclusion

The rapid review was performed to derive an operational definition of professionalism for a mental health services context. Two review questions were posed; (1) how is professionalism conceptualised within mental health services; and (2) does this definition vary across the mental health professions. Common themes were identified across the professions. Professionalism was conceptualised on two levels. First, on a societal level – a dynamic social contract; and second, on an individual level – representing the profession. Patients expect professionals to have the appropriate skills and knowledge to undertake their role. Professionals are frequently faced with ethical dilemmas in a mental health setting, and they must use their judgement, thinking critically and acting accordingly. There was a significant focus on interpersonal skills in the literature, and the relationship between a practitioner and patient was found to have more influence on positive outcomes than technical interventions.

The importance of interpersonal skills and the requirement for staff to think critically, whilst responding to ethical dilemmas, lend support to the idea of developing an SJT that would assess one's knowledge of professionalism for a mental health services context. However, a lack of patient presence was observed in the literature, and the nursing and social work professions were underrepresented. The operational definitions derived from this review will facilitate future research with key stakeholders in this field allowing renegotiations of the contract. Indeed, the following chapter reports on a qualitative study that sought to further conceptualise professionalism from the perspective of patients, carers and staff members from a range of professional disciplines working in mental health services.

4.13 Addendum to the original review findings

The review reported here has been cited in several international papers since the time of its publication. Citing papers/theses have focused on a range of professional disciplines and settings. Whilst the content of some papers relate to mental health settings specifically, including a psychiatry emergency department (Yahyavi et al., 2021) and community-based mental health services (Rioli et al., 2020), another related to a specific healthcare role, namely non-medical Responsible Clinicians (Oates et al., 2020). The topic of medical professionalism (Song et al., 2021), radiography (Hale and Wright, 2021), and home-based family therapy (Fitzgerald, 2019) were also covered, and a citing thesis reported on mental health help-seeking behaviour (Cheesmond, 2020). None of the citing documents refuted the findings of the original review.

The literature search was repeated on 27th July 2022. The CINAHL database was chosen for the search because it is stated to be "*the world's most comprehensive source of full-text for nursing & allied health journals*." This was deemed particularly relevant given the findings of the quantitative study, reported on in chapter 6. Search strings were kept as identical as possible to the original search, however some minor edits were needed (see Appendix B); for example, the proximity operator ADJ3 was replaced with N3 as the former is no longer identified by CINAHL. In addition, instead of searching for articles that were published between 2006 to 2nd March 2017, criteria were added to capture articles that have been published since March 2017. It was hoped this would identify articles that were absent from the original review findings.

Using the search string detailed in Appendix B, 204 papers were generated during the search. A title and abstract screening conducted solely by the author resulted in 21 papers for full-text screening; articles were predominantly excluded at the title and abstract screening stage because they did not refer to specialist mental health services. Of the 21 articles that were retained for full-text screening, nine met the inclusion criteria that are noted in section 4.4 (Aylott et al., 2019; Brunn et al., 2020; Day-Calder, 2021; Fish, 2022; Gabbard, 2019; Glas, 2017; Glauser, 2020; Sabin and

Harland, 2017; Stacey and Pearson, 2020). It is important to note that one of these articles was written by the current author (Aylott et al., 2019). This article reports on the findings of the review discussed in this chapter and is therefore not expanded upon here. At the full-text screening stage, 12 articles were excluded because they either did not refer to a mental health services context (n=3), had a lack of content or definition (n=8), or were not freely accessible at the time of retrieval (n=1). Similar to the original review findings, over half of the remaining eight articles were written by psychiatrists despite the search being undertaken on the CINAHL database. The articles focused on the importance of relationships with both patients (Day-Calder, 2021; Glas, 2017), and other professionals (Brunn et al., 2020), as well as how the professional relates to their own role (the existential core of professionalism; Glas, 2017). There were two articles that focused on digital professionalism, which was somewhat expected given the increasing use of social media more recently (Gabbard, 2019; Sabin and Harland, 2017).

Professionalism is referred to as an 'occupational value' Evetts 2011 (as cited by Fish, 2022) that "*requires self-discipline and fidelity to ethical norms*" (Sabin and Harland, 2017). Practitioners must maintain clear boundaries to maintain their professionalism (Day-Calder, 2021). Breaches of professionalism were highlighted in the literature regarding internet activity (Gabbard, 2019), which included 'posting scenes of intoxication', 'profanity' or 'sexually suggestive material' (Sabin and Harland, 2017); on the other hand, softer skills which promote engagement and facilitate cooperative endings, as well as the ability to offer containment and safety were considered indicators of professionalism (Stacey and Pearson, 2020). One author commented on how education on professionalism during undergraduate training stressed different rules for women in comparison to men (Glauser, 2020). An example the author provided was how much time was used to explain what women should and should not wear, such as covering one's shoulders, which contrasted to the short amount of time that was applied to men. Of concern, a discussion was provided regarding how

clinicians may have formal complaints made about them if they call out the practice of others in their profession (Glauser, 2020). Glauser cites Dr Nav. Persaud who says that questioning the status quo can be "*interpreted as a criticism of the profession in general, or some of your colleagues in particular, even if the goal is to improve care for your patients*" (Glauser, 2020, p.E1647).

In addition to the above discussion, it is argued that professionalism should be perceived "*as more than a bunch of executive functions requiring excellence*" as this provides a "*decontextualised, rational and technical view of the professional role*" (Glas, 2017). Glas (2017) instead refers back to professionalism being legitimacy and entitlement on the basis of a social contract. According to the social contract, there are core values, which should be *internalised* as a professional. Glas (2017, p.541) suggests that if professionals adhere to a technical, instrumentalist view of professionalism and detach their personal self from their professional self, then they will communicate this indirectly to a patient as "*I am here for you as someone with certain knowledge and skills; however, as a person I do not have anything to do with you*". Glas cites his prior work in noting that professionalism has an existential component. He subsequently proposes that dependent on how professionals relate to their role, important existential messages may be communicated to others.

Chapter 5: Understanding and Exploring Professionalism in Mental Health Services: A Qualitative Study using Stakeholder Focus

Chapter 4 reports on the findings of a systematic review and presents two operational definitions of professionalism for a mental health services context. The reviewers observed a lack of patient presence in the literature and the views of nurses and social workers were also underrepresented. This chapter discusses the conduct and findings of a qualitative study that sought to further conceptualise professionalism from the views of key stakeholders in mental health services, including the experience of patients, carers, and staff members across multiple disciplines.

5.1   Background

Professionalism is a context dependent (Rees and Knight, 2007), multidimensional concept (van de Camp et al., 2004). The systematic review, reported on in chapter 4, resulted in two operational definitions of professionalism that apply to all professionals working in mental health services (see chapter 4, section 4.11). These being:

1.   Professionalism forms the basis of a dynamic social contract between professions and society. This contract (which can have both tacit and explicit elements) specifies that society will remunerate the members and permit the profession to self-regulate on the understanding that the profession use their skills for patient and public good.

2.   On an individual level, professionalism can be conceptualised as a latent trait, composed of elements of intrapersonal, interpersonal, and working professionalism. This trait may only be observed through manifest behaviours in certain situations. Such behaviours will be in keeping with society's expectations and demonstrate a commitment to ethical practice, cultural-sensitivity, self-awareness and reflection and self-discipline.

Whilst the domains of professionalism in mental health services may be similar to other healthcare settings, it is more a matter of emphasis on certain elements of professionalism that differ (see chapter 4). Interpersonal relationships, empathy, and communication, among other traits, are paramount in healthcare services. However, there are specific challenges and more nuanced skills often required when working with individuals affected by mental health difficulties or developmental disabilities. These include the prevalence of social and communication issues in these groups, as well as the occurrence of behaviours that can challenge professionals. Professionals must possess practical wisdom, an element of working professionalism, which allows professionals to behave accordingly in situations where uncertainty and ambiguity arise (see section 4.10.1.3).

Whilst many qualitative studies have sought to explore the concept of 'professionalism' in medicine (Birden, 2012; Brownell and Côté, 2001; Jha et al., 2006; Monrouxe et al., 2011; van de Camp et al., 2004; Wagner et al., 2007), there is a dearth of studies focusing on the concept in mental health services, where patients' needs differ. The systematic review reported on in chapter 4 observed that the professionalism literature is mostly written by psychiatrists (45 of 70 records), and there is minimal literature reporting on the perceptions of patients and carers, or the other healthcare professions (see section 4.10). Knowledge of the desired professional

attributes for mental health services may be used to facilitate the teaching and assessment of professionalism in this field. Knowledge regarding the behaviours that are perceived as demonstrating these attributes would also help facilitate the development of item content for assessment tools, including situational judgement tests (SJTs). Given that patients and carers should be the primary focus of mental health services, we must explore their views, alongside the perceptions of various professional groups that work in this setting. Therefore, this study sought to explore the views of patients, carers and professionals on professional attitudes and behaviours in a mental health services context. The researchers decided to analyse the data against the Professional Attributes Framework. This framework originated as part of the 'Improving Selection to the Foundation Programme' project, which was commissioned by the Medical Schools Council (2011). As most of the professionalism literature regarding mental health services focuses on psychiatry (see chapter 4, section 4.10), the Professional Attributes Framework was considered a suitable framework. The Professional Attributes Framework also helped inform the development of the interview guide.

5.2   Research questions

The research questions for the study were:

1) What are patients', carers' and professionals' experiences of, and perceptions of professionalism in a mental health setting?

2) How does this experience align, if at all, with medically defined, generic, professional standards and attributes?

3) Do conceptualisations of professionalism, resulting from this study, coincide with the findings of an earlier systematic review on the topic?

5.3   Methods

All procedures were approved by the University of York Research Governance Ethics Committee (Appendix D), the Health Research Authority, and the local NHSFT. A

favourable ethical opinion was also obtained from London - Camden & Kings Cross Research Ethics Committee (REC reference: 18/LO/0630; Appendix E).

5.3.1 Study design

This cross-sectional qualitative study was conducted in a mental health services setting. Arksey and Knight (1999) argue that research methods should not be chosen on the basis that one method is superior to the other, but on what method is most suitable for the research purpose. Interviews and focus groups have been used together successfully in prior research studies (Lambert and Loiselle, 2008; Steffen et al., 2020). The application of both interviews and focus groups was deemed a suitable approach for this study, because individual interviews can facilitate an in-depth discussion of social and personal matters (DiCicco-Bloom and Crabtree, 2006), and focus groups are an effective method for exploring the attitudes and needs of staff (Kitzinger, 1995). Therefore, semi-structured interviews were undertaken with service users, and focus groups, as opposed to group interviews, were performed with carers and professionals to explore the perceived attributes of professionalism in mental health services. To support the development of the SJT, the critical incident technique (Flanagan, 1954) was utilised to generate examples of professional and unprofessional behaviour within this setting.

5.3.1.1 *Interviews*

Following a meeting with the Research Ethics Committee, one to one interviews were judged to be most appropriate for patients, due to issues of confidentiality; first, patients may have disclosed the names of staff members, and second, the researchers were concerned about the confidentiality of the patients' medical histories being preserved. This consideration is in keeping with the World Medical Association's Declaration of Helsinki (2013), which states that "*every precaution must be taken to protect the privacy of research subjects and the confidentiality of their personal information*". Whilst it is possible to ensure confidentiality on behalf of the researcher, the researcher cannot promise that other participants will maintain confidentiality too.

5.3.1.2 *Focus groups*

Focus groups were performed with carers and professionals as this was deemed the most appropriate method to generate shared meaning; this allowed the researchers to capture as many facets and experiences of professionalism as possible. To increase candour and avoid power relationships amongst staff, focus groups were conducted separately for each of the professional groups; i.e., there was a separate focus group for psychiatrists, psychologists, nurses, allied health professionals (namely, occupational therapists), and carers. Some focus groups consisted of professionals that knew one another, whereas others included individuals that were not familiar with one another.

Confidentiality was not considered to be a problem for focus groups, as professionals are bound by confidentiality practices in the trust; professionals were reminded of this in the participant information sheet. Whilst carers are not bound by the same rules, confidentiality was not felt to be an issue with this group, because carers would be less likely to know the same members of staff as one another.

5.3.2 Reflexivity

In line with a social constructionist perspective, there was an awareness that the data generated during the study was a result of the interaction between the researcher and the researched. Reflexivity was therefore utilised over the course of the study. Reflexivity promotes the validity of qualitative research (Pillow, 2003), because researchers consider how their background, characteristics, and engagement with the research process may influence the data generated and their subsequent interpretation of findings (Berger, 2015). For example, the researcher identifies as a gay woman, and it is possible that this perceived commonality with a patient may have influenced a discussion that took place during one of the interviews. The researchers were all aware that they had preconceived ideas, not only from having been a former patient, carer and/or staff member working in mental health services, but also having undertaken a literature review on the topic; indeed, it was this initial understanding

130

that helped formulate the interview guide. To increase the validity of the research, two researchers analysed the qualitative data. The author has also been transparent within this thesis regarding their position in the research process.

5.3.3 Ethics

The researchers were aware that the interviews with patients could potentially bring up distressing memories for them. A clinician was therefore available should there have been any concerns of this nature; contact with the clinician was not required throughout the study period. The researcher reminded participants at the commencement of each interview that they were free to withdraw at any time. Interview participants (i.e., patients) were informed that they could take a break if needed and that there was no need to disclose the nature of any mental illness that affected them, or that they had received care for. In addition, interview participants were provided a sheet with the contact details of various third-party support organisations, in case they wanted to obtain support, external to the NHS Trust, following their participation. It is not known whether participants used these contact numbers following the study. At the end of each interview and focus group, a post-briefing took place. The next steps of the research project were reiterated, and participants were thanked for their time.

5.3.4 Recruitment

The study was advertised across Tees, Esk and Wear Valleys NHSFT, and a purposive sample was obtained. Interviews and focus groups were timetabled and located to optimise participation; all interviews and focus groups were held on premises where NHS staff worked. Participation in the interviews and focus groups was voluntary. Water was available during all interviews, and light refreshments were made available during focus groups. Patients and carers were provided a high-street shopping voucher, to the value of £20, to compensate for their time and travel.

5.3.4.1 *Recruitment of interview participants*

Flyers advertising the study were posted in various community mental health centres. Flyers were also distributed via a patient and public involvement newsletter and the NHS Trust's social media site. Participants were asked if they knew other individuals that may be interested using a 'snowballing' method to recruitment. In one instance, a patient who participated in an interview contacted many fellow patients by email, noting their perceived value of the research and their experience of taking part; this resulted in additional individuals coming forward to participate. The study initially sought 6-10 participants for interview; this figure increased however, due to the number of expressions of interest received. The inclusion and exclusion criteria for interview participants can be viewed in Table 5.1.

Table 5.1 Inclusion and exclusion criteria for interview participants

Inclusion Criteria	• Current or recent (last 24 months) user of mental health services whose mental health is stable (i.e., not a current in-patient, open to the crisis team or in urgent need of a treatment review) • Sufficient English language fluency that permits comprehension of written information relating to the study • Sufficient capacity to be able to consent to participate • Aged 16 years or above
Exclusion Criteria	• Lacking capacity to provide informed consent to participate • Less than 16 years of age • Experiencing an acute episode of mental ill health • Having a significant or serious act of self-harm within three months prior to the study necessitating medical and/or psychiatric assessment or intervention • Having experienced a period of hospitalisation for mental health problems six months prior to the study • Experiencing literacy problems that would impair the ability to read and understand written material relating to the study • Patients in Forensic services that are supervised by the probation service • Individuals that do not consent to all statements on the consent form

5.3.4.2 *Recruitment of focus group participants*

It was recognised that focus group discussions can be inhibited when there is too much heterogeneity among participants (Freeman, 2006). Therefore, a decision was made to hold separate focus groups for each of the professions. Social media was utilised to recruit participants. Flyers that advertised the study were again posted in various community mental health centres. The study was advertised on the trust intranet (internal website) and in an electronic newsletter that is sent to staff across the NHS Trust on a weekly basis. In addition, the lead researcher met with team managers and advertised the study during team meetings. To maximise participation, service leads were contacted to determine whether there were any Trust events that were taking place where a focus group could be conducted, this led to the focus groups with psychiatrists, as well as the occupational therapists. Professional leads in the trust were also contacted and asked if they would be happy to email professionals in their discipline; the professional lead for psychology and the allied health professions subsequently distributed an email to all professionals working in these disciplines whilst asking for expressions of interest.

To recruit carers, flyers advertising the study were distributed via a patient and public involvement newsletter. The author knew some carers personally, and these individuals were contacted to identify forums for recruitment. This led to the lead researcher attending a carers' network meeting where the study was discussed; at this time, attendees were invited to contact the researcher if they were interested in participating.

The study sought to identify between 6-10 participants from each stakeholder category, including carers, psychologists, nurses, allied health professionals, and psychiatrists. In addition, attempts were made to recruit from all the main service areas, including adult mental health services, children and adolescent mental health services, learning disability services, forensic services, and mental health services for

older people. The inclusion and exclusion criteria for carers and professionals can be viewed in Table 5.2.

Table 5.2 Inclusion and exclusion criteria for focus group participants

Staff members	Inclusion Criteria	• Registered professionals working in a mental health and learning disability NHSFT
	Exclusion Criteria	• Individuals that do not consent to all statements on the consent form • Non-registered staff (e.g., nursing assistants)
Carers	Inclusion Criteria	• Unpaid carers of patients using mental health and/or learning disability services (i.e., relatives and/or friends)
	Exclusion Criteria	• Individuals that do not consent to all statements on the consent form

5.3.5 Informed consent

With regards to obtaining consent, the procedures differed for individuals participating in interviews to those participating in focus groups. All individuals that expressed an interest in the study were provided a participant information sheet (see Appendix F for an example), and consent form (Appendix G), and were advised to contact the researcher if they had any questions or wanted any further information. There were separate information sheets for each of the professional groups, as well as for patients and carers. Individuals were given at least two weeks after being sent an information sheet before any further contact was made. After two weeks, individuals were contacted again to discuss the study and determine whether they would like to take part. Individuals were reminded that their participation was voluntary and that they could withdraw at any time. Eligibility was ascertained at this time, by referring to the study's exclusion criteria.

5.3.5.1 *Consent process for interviews*

Interview dates and locations were arranged according to the participant's preference. Upon meeting face to face, the researcher (LA) discussed the participant information

sheet with the participant, whilst evaluating the individual's comprehension of this. Participants were next asked to sign a consent form. All individuals were deemed to have the capacity to provide informed consent, and all went on to sign the consent form and participate in the interview.

5.3.5.2 *Consent process for focus groups*

Individuals that had expressed an interest to participate in the focus groups were contacted two weeks after being sent a participant information sheet, to discuss the study and determine whether they would like to take part. Once a focus group date and location had been arranged, individuals were forwarded a further copy of the participant information sheet, as well as a consent form; individuals were asked to review the information and sign the consent form, where possible, before the day of the focus group. At the start of each focus group, the researcher (LA) discussed the participant information sheet with the group, evaluating participants' comprehension of this. Participants were next asked to complete the consent form, had they not already done so. All participants were determined to have the capacity to provide informed consent. All individuals decided to participate and signed the consent form prior to their participation in the study.

5.3.6 Data collection

All data were collected face-to-face and were audio recorded. The tape-recorder was tested at the start of each interview and rooms were arranged to minimise external noise and distractions. Interviews and focus groups were facilitated by the author (LA). Four focus groups were co-facilitated by a clinical researcher (PT), or a researcher that specialised in qualitative methods (GF). Interview questions were established based on the critical incident technique (Flanagan, 1954), as well as earlier research on professionalism in other healthcare settings (Burford et al., 2014; Medical Schools Council, 2011). A pilot focus group confirmed that the interview questions were fit for purpose and did not need editing. Interview stems remained the same for both interviews and focus groups (see Table 5.3); however, not all prompts were necessary,

during focus groups, to elicit the information. Nevertheless, the topic guide facilitated a semi-structured approach and minimised the likelihood of interviewer bias.

At the start of each interview, participants were provided a questionnaire, which asked their age, gender, and ethnic group. Those attending a focus group were also asked to identify whether they were a carer or a professional, and professionals were subsequently asked to document their profession. Participants were asked to indicate on this same questionnaire whether they were happy to be contacted via email or phone for the follow-up section of the study. The follow up was conducted to facilitate the development of SJT items and a provisional scoring key; this is discussed further in chapter 6.

Table 5.3 Topic guide used to generate discussion on professionalism in mental health services

Topic stem	Prompts
Professionalism	What does professionalism, in general, mean to you?
	What does good 'professionalism' look like, in general?
	What does 'unprofessional' behaviour look like, in general?
	What are the most important aspects of 'professionalism' for staff working in 'mental health services'?
	What skills are important when working in mental health services?
	What interpersonal attributes are important when working in mental health services?
	What values are expected of staff working in mental health services?
Professional Behaviours	Is anyone able to discuss an example of professional or unprofessional behaviour, that they have recently observed, especially good or bad practice?
	What were the circumstances leading up to the incident?
	Would you please describe the professional's reaction?
	What did the person(s) do that was (un)professional?
	- has anyone else got any other views regarding this behaviour(s)?
	- how else could the professional have responded?
Professional Dilemmas	Has anyone observed any dilemmas that a professional has recently encountered?
	What were the circumstances?
	Did the professional(s) behave in a professional or unprofessional manner?
	What did the professional do?
	What was it about the behaviour that was (un)professional?

The facilitator introduced themselves at the start of each interview and noted the purpose of the study; co-facilitators also introduced themselves, where present. Some ground rules were discussed at the start of each focus group, which touched on the following matters:

- professionals were asked to place their mobile phone on silent and politely asked if they would leave the room if they needed to take a call
- professionals were advised to help themselves to drinks that were provided
- the importance of confidentiality was reiterated, whilst noting that issues related to safeguarding would need to be raised with the appropriate authorities
- individuals were advised that there would be a post group briefing should they wish to discuss any concerns following the group
- individuals were asked to respect one another's views. It was noted that there were no right or wrong answers and that we were not looking for agreement among the group
- Individuals were asked to talk one at a time, but were also encouraged to interact with one another
- Individuals were subsequently asked to introduce themselves

With regards to the focus groups, the researchers aimed to generate discussion and interaction among the participants and this was encouraged from the start; the researchers also allowed the participants to lead on the discussion, where possible. Fewer prompts were used during focus groups, which appeared to facilitate the interaction among participants. Interviews and focus groups formed two parts; first, participants were asked to define professionalism for a mental health services context; second, the interviews and focus groups were used to generate scenarios that would manifest professional or unprofessional behaviours. At the end of each interview and focus group, the researcher summarised the discussion and asked the co-facilitator,

where applicable, whether they had observed anything other than that summarised. The participants were then asked whether this accurately captured the group discussion.

With regards to the individual interviews, in two instances patients had asked the researcher, in advance of the interview, what the questions would be. The researcher was aware that this may influence the data collected and provided a broad explanation regarding the questions, reemphasising that mentioned on the participant information sheet (i.e., that the participant would be asked to explain professionalism from the participant's perspective and would be asked for examples of professional and unprofessional behaviour). Participants were advised that more specific questions would be asked on the day. In one instance, a patient brought pre-prepared notes along to the interview with them; this participant was informed that the information would need discussing during the interview if they wished for it to be included in the study. Various probes were used during the focus groups to build rapport and encourage the involvement of others, including; can you tell me a bit more about that, that's interesting, what does everyone else think, does anyone else have any other examples. Alongside these verbal prompts, the researcher was mindful of their non-verbal communication. Contextual notes were made immediately following each focus group, whilst the event was clearly in mind.

Interviews and focus groups were transcribed verbatim by a member of staff in the NHS Trust's Research and Development department; a third-party organisation took over the transcription of audio data once the Trust approved this. Participants were informed about the use of a transcription firm in the participant information sheet. All transcripts were quality checked, and revised accordingly, by the author. The ethics committee requested that audio recordings were destroyed following transcription, which contrasts with the guidance of many ethics committees, who demand that tapes are kept for ten years (Holloway and Galvin, 2017). Data protection and storage

guidelines were adhered to, including the use of industry standard encryption. Transcripts were not returned to participants, but regular member checking was used during the focus groups and interviews to confirm the researchers' interpretation of the data. Transcripts were anonymised so that they did not identify participants.

5.3.7 Data analysis

Data were managed using nVivo (version 12). Having received mental health services previously, as well as conducted the review reported on in chapter 4, the researchers had *a priori* themes in mind and framework analysis (Ritchie and Spencer, 1994; Ritchie et al., 2003) was therefore chosen as the approach to analyse the data; the process of which can be viewed in Figure 5.1. Grounded Theory, as a qualitative analysis framework, was ruled out, because the study commenced from a well-defined starting position. That is, that the researchers had recently undertaken, and published, a rapid systematic review on the topic (see chapter 4).

Framework analysis involved a five-step process, including: familiarisation of the data, developing a theoretical framework, indexing and charting, summarising data within the analytical framework, and synthesising the data (Ritchie and Spencer, 1994; Ward et al., 2013). During the familiarisation stage, open coding was used to grasp a better understanding of the data. In doing so, it was determined that the data fit the Professional Attributes Framework (Medical Schools Council, 2011). The Professional Attributes Framework was originally developed as part of the 'Improving Selection to the Foundation Programme' project, commissioned by the Medical Schools Council (2011). The framework consists of nine professional attributes, including: Commitment to Professionalism, Coping with Pressure, Effective Communication, Learning and Professional Development, Organisation and Planning, Patient Focus, Problem Solving and Decision Making, Self-Awareness and Insight, and Working Effectively as Part of a Team. Both inductive and theoretical approaches can be used to analyse research data (Braun and Clarke, 2006); for this study, a deductive approach was used to index data against all nine professional attributes, however, an inductive approach was also

utilised, as the research team recognised that inductive approaches can provide a more expansive analysis of the entire body of data (Kiger and Varpio, 2020). Using an inductive approach facilitated the creation of additional codes and themes pertinent to mental health services.

Familiarisation
Gain an overview of the range and diversity of data

Identifying a thematic framework
Devise an index, which can be used to sift and sort data

Indexing
Systematically apply the index to the research data

Charting
Rearrange the data according to thematic references

Mapping and interpretation
Analyse the data according to the key research objectives
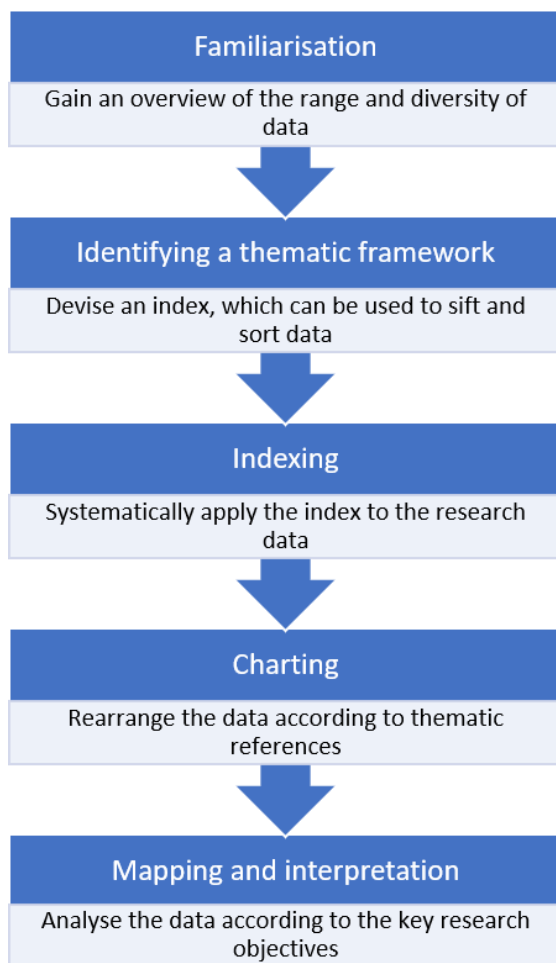
Figure 5.1 Stages of Framework Analysis (adapted from Ritchie and Spencer, 1994)

It is recognised that researchers bring their own bias when interpreting research findings, particularly if they are familiar with the setting and literature on the topic (Holloway and Galvin, 2017). Therefore, data analysis was primarily undertaken by two researchers; both the author (LA) and Dr Sally Brown (SB), whom has particular expertise in qualitative research. At the time of the study, SB was a lecturer at Edinburgh Napier University. SB used an iterative process, whilst reviewing a transcript

against the codes created by LA. Codes and themes were subsequently amended through discussion and negotiation. To enhance the validity of the analysis, Professor Finn and Professor Tiffin reviewed a sample of the coding. For clarity, the themes will now be referred to as the professional attributes; also, the codes are the expectations placed on professionals that work in mental health services.

5.3.8   Reliability and validity

With regard to the reliability and validity of research, some researchers posit that qualitative studies should be evaluated by different criteria than that used for quantitative studies (Bryman, 2016). The validity of qualitative research can be improved through various procedures, including triangulation, respondent validation, clear exposition of methods, reflexivity, attention to negative cases, and fair dealing (Mays and Pope, 2000). Respondent validation, which is otherwise referred to as informant feedback, member checking, member validation, or dependability checking (Varpio et al., 2017), was used throughout the current study. Transcripts and initial analyses were not provided to participants, however LA summarised the discussion during interviews and focus groups, noting their interpretation of the data, to verify whether the participants agreed with the researcher's account. Results were also disseminated to participants following the study; none of the recipients expressed any disagreement with the findings.

5.4   Results

A total of 13 interviews and six focus groups were conducted between 15th August 2018 and 15th January 2019. There were 56 participants in total, which incorporated 10 nurses, 10 occupational therapists, 7 psychiatrists, 7 psychologists, 7 carers and 15 patients. Whilst it was intended that the focus groups were split by discipline, on the day of the focus group, a psychological wellbeing practitioner turned up to the nursing focus group and a student nurse attended the psychology focus group. The pilot focus group also incorporated two patients and two carers. The remaining 13 patients

participated in one-on-one interviews. With regards to staff participation, all allied

health professionals were invited to take part, including dieticians, physiotherapists,

pharmacists, and occupational therapists; however, only occupational therapists were

able to participate due to scheduling. The duration of each focus group can be seen in

Table 5.4; the mean duration of a focus group was 1 hour and 15 minutes. Interviews

lasted between 36 and 78 minutes; the mean duration of an interview was 1 hour.

Table 5.4 Focus group size and duration

Profession	Number of participants	Time elapsed (hours and minutes)
Nurses	10	1 hour and 10 minutes (70 minutes)
Occupational Therapists	10	59 minutes
Psychiatrists	7	1 hour and 19 minutes (79 minutes)
Psychologists	7	1 hour and 28 minutes (88 minutes)
Carers	5	1 hour and 22 minutes (82 minutes)
Pilot	4	1 hour and 10 minutes (70 minutes)

In one instance, a team manager provided a room and agreed to have staff members

participate in a focus group to facilitate their professional development. On this

occasion, most individuals that participated in the focus group (eight of ten) worked in

the same building, albeit in one of three separate teams. The two other individuals

that participated worked elsewhere in the organisation. A detailed breakdown of

participants' demographic details is not provided here, as this may compromise the

anonymity of individuals; however, the overall sample included 36 females, 18 males,

and 2 individuals that described their gender as other. The age of participants ranged

from 21 to 86 years; with a mean age of 47 (having excluded two individuals that did

not give their exact age). Most participants described their ethnicity as white;

however, two individuals described their ethnicity as black, and two described their ethnicity as Asian. Professionals worked in a range of specialties, including adult mental health services, children and adolescent mental health services, learning disability services, forensic services, and mental health services for older people. In addition, professionals worked across inpatient, community and corporate settings.

Interview and focus group data were indexed together using the professional attributes framework (Medical Schools Council, 2011). The charting phase of the analysis allowed data to be summarised for each case (i.e., interview or focus group), as well as by each category (i.e. attribute). Data generated during the study aligned to all nine attributes of the Professional Attributes Framework, however the codes within each attribute varied from that originally cited by the Medical Schools Council (2011). The analysis found that the data required a tenth attribute for professionals working in mental health services, which incorporated 'Working with Carers.' Each of the resulting ten attributes are discussed here with a brief description provided for each.

5.4.1 Commitment to professionalism

Summary: Individuals should be committed to honouring their profession by adhering to guidelines and challenging poor practice. Professionals must have integrity and be a responsible practitioner.

The professional attribute 'Commitment to Professionalism' generated 13 subcodes (as demonstrated in Table 5.5). Participants expect professionals to adhere to their professional guidelines and codes of conduct; otherwise, clinicians have a duty to challenge, thus requiring confidence and courage. Focus groups highlighted that social media could be a challenge to one's professionalism, specifically with regard to the use of online dating and social media platforms, such as Facebook. To overcome these challenges, participants were found to use two separate online accounts to keep their work and personal life separate.

...it's about having clear boundaries as well isn't it, so I've got a Twitter account for purely professional stuff … but then Facebook, I have strict settings on that... (OT – Focus group)

I am a single person, but I have got an issue for example with internet dating as well because of that status you have got [as a psychiatrist], I don't do it.... I think there are some unwritten rules you are supposed to follow. (Psychiatrist – Focus group)

Table 5.5 Corresponding codes for the attribute Commitment to Professionalism

Subcode	Number of interviews and focus groups	Number of references
Adheres to professional, ethical and legal guidelines	9	36
Behaviour standard – behaves according to expectations	9	30
Challenges the system	3	4
Challenges unacceptable behaviour	6	19
Displays a commitment to the role	11	28
Is trustworthy and has integrity	12	26
Maintains confidentiality	8	25
Owns up to mistakes	5	11
Possesses confidence and courage	6	9
Takes responsibility for own actions	5	7
Understands and demonstrates an awareness of ethical issues	8	18
Upholds the profession's and organisation's reputation	5	6
Uses social media appropriately	3	10

5.4.2 Coping with pressure

Summary: Practitioners must utilise their clinical judgement, particularly in times of uncertainty and ambiguity. Professionals must have resilience and be able to de-escalate situations when others are experiencing distress.

The professional attribute 'Coping with Pressure' generated seven subcodes (as demonstrated in Table 5.6). It became apparent during the study that professionals need to separate themselves from the work that they do, in order to not let this impinge upon their own wellbeing.

> *Professionalism for me is walking out of here and onto the next thing with a fresh mind, a fresh view...you have got to detach. (Patient – Interview)*

During focus groups, a nurse commented "*we don't shout back; we just have to stay professional*". In stark contrast however, a patient mentioned how they had calmed down, having been restrained on the floor, yet a care assistant continued to pull the patient's thumb back. Similarly, a carer highlighted that some staff de-escalate situations better than others.

> *I have seen physical restraint used detrimentally to the point even when people were saying 'oh should we do your hair' while they are restraining this person on the floor, whereas other staff would be able to talk to that person and dissolve the situation (Carer – Focus group)*

Table 5.6 Corresponding codes for the attribute Coping with Pressure

Subcode	Number of interviews and focus groups	Number of references
Being overworked	2	2
Checking the facts	3	4
Clinical judgement	4	6
De-escalation	3	12
Is resilient and employs effective coping strategies	7	23
Manages uncertainty and ambiguity	2	4
Remains calm and in control of situations	5	9

Professionals must be aware of the potential for malpractice and deal with this accordingly, whilst managing their own wellbeing. Humour was observed to be a means of making light, defusing stress and normalising situations that occur in such challenging environments. A pertinent example was provided during a focus group, with regards to a suicidal patient frequently swallowing a specific form of vegetation. After numerous incidents, a staff member commented, '*I can't wait for spring, when there are no [vegetation] around.*' Such remarks were viewed as a positive coping mechanism for some, however they were viewed negatively in certain circumstances, such as when one is critical of a patient. A professional expressed discomfort whilst making the following comment.

> *"I heard staff being a bit critical of the patient…It was put down, it's how the staff cope, they are working in a really stressful situation, these individuals can be really challenging, it's how they offload." (OT – Focus group)*

Some individuals expressed using humour when discussing difficult topics, however it was recognised that this may be judged negatively by an external person present.

> *"we use humour to get us through that. ... all of us use that humour protectively. However, another person comes in, we don't know how that's going to be judged." (Psychologist – Focus group)*

5.4.3 Effective communication

Summary: Practitioners must communicate effectively, using both verbal and non-verbal communication. Professionals should have the ability to build rapport with patients, and validate the thoughts and feelings of others.

The professional attribute 'Effective Communication' generated 13 subcodes (as demonstrated in Table 5.7). Referring to a staff member and patient using derogatory language in jest, a participant commented in an interview that professionals should not use bad language with patients.

"If you are using regular bad language and especially if you personalise it by calling a patient that, then that… [is] really unprofessional." (Patient – Interview)

Banter was recognised as a means of building rapport; however, during a focus group, participants noted that this could go wrong very quickly, particularly "*if someone feels targeted, … where they've got trauma histories, where they've been humiliated, bullied, and all the rest.*" (Psychologist – Focus group)

Table 5.7 Corresponding codes for the attribute Effective Communication

Subcode	Number of interviews and focus groups	Number of references
Adjusts style of communication	5	12
Advocates appropriately for patients	3	3
Builds rapport with patients	6	10
Communicates effectively with team members	3	5
Demonstrates open and honest communication	9	24
Demonstrates sensitive use of language	7	16
Ensures surroundings are appropriate when communicating	3	7
Is personable and able to relate with others	10	21
Listens effectively	14	24
Has good observational skills	2	2
Understands and responds to non-verbal cues appropriately	2	4
Uses non-verbal communication effectively	6	11
Validates the thoughts and feelings of others	10	17

5.4.4 Learning and professional development

Summary: Practitioners must possess the appropriate knowledge and skills for their role, utilising professional development opportunities. Professionals should accept feedback and utilise supervision accordingly.

The professional attribute 'Learning and Professional Development' generated seven

subcodes (as demonstrated in Table 5.8). Participants highlighted that they would

discuss the difficulties they had with patients during supervision; to air these views

elsewhere was deemed unprofessional.

> *"You could say the same thing, but in different contexts. And some are*
> *professional and some are not professional." (Psychologist – Focus group)*

Table 5.8 Corresponding codes for the attribute Learning and Professional Development

Subcode	Number of interviews and focus groups	Number of references
Accepting of feedback	2	2
Application of knowledge and learning to practice	2	6
Appropriately makes use of supervision and formulation meetings	7	21
Possesses the relevant knowledge and skills for the role	13	32
Learns on the job (continuing professional development)	6	7
Possesses appropriate qualifications for the role	4	4
Is involved with research	1	1

5.4.5   Organisation and planning

Summary: Practitioners must maintain accurate records and read case notes

attentively. Professionals must effectively manage limited resources, and their time

accordingly.

The professional attribute 'Organisation and Planning' generated seven subcodes (see

Table 5.9). At times, participants views differed. This was the case with regards to

documentation; a professional felt that documentation was a waste of clinician time,

whereas a patient stated that at times documentation must take precedence.

"All the notes we keep really are litigation proof that's all it is you know (Nurse – Focus group)

"The planning of my care could depend on that, … If it isn't recorded properly I might lose out on the appropriate care." (Patient – Pilot focus group)

Professionals must familiarise themselves with the notes, yet a patient disclosed how this did not seem to be the case in their experience.

"I feel as if they didn't know anything about me … they had never read any notes on me." (Patient – Interview)

Table 5.9 Corresponding codes for the attribute Organisation and Planning

Subcode	Number of interviews and focus groups	Number of references
Is efficient - Goes the extra mile	5	8
Ensures that systems are in order	2	2
Maintains accurate records	5	9
Manages limited resources effectively	4	19
Reads case notes	5	9
Utilises effective time management	7	15
Wears suitable attire for the role	4	9

5.4.6 Patient focus

Summary: Practitioners must possess qualities that enable them to build therapeutic relationships with patients, such as altruism and humility. Professionals, also, must maintain an appropriate professional distance and not impose their own values on patients, delivering person-centred care.

The professional attribute 'Patient Focus' generated 18 subcodes (as demonstrated in Table 5.10). Whilst it was agreed that professionals should be approachable and friendly, boundaries must also be instilled.

> *"when they think that you are their friend … you then need to remind possibly remind them that it's a professional relationship." (Nurse – Focus group)*

Table 5.10 Corresponding codes for the attribute Patient Focus

Subcode	Number of interviews and focus groups	Number of references
Attends to patients' physical health needs	3	4
Is a human interface for the organisation	3	7
Builds therapeutic relationships	4	6
Is caring and possesses kindness and compassion	13	45
Contains the emotions of patients and carers	2	2
Does not impose own values on patients	3	5
Has a cheerful and friendly nature, yet maintains an appropriate professional distance	9	22
Is altruistic and possess humility	10	21
Is approachable and makes oneself available for patients	11	28
Is empathic and understands the illness that mental illness can have	13	38
Is genuine, honest and fulfils promises	10	33
Is non-judgemental	10	32
Provides reassurance	9	27
Provides equal access of services to all patients, regardless of challenges	2	3
Maintains safety	12	26
Treats patients with respect	15	40
Utilises a person-centred approach	13	53
Utilises a recovery focused approach with meaningful activity	10	31

Clinicians must respect patients. However, during their discussion of two opposing experiences of receiving treatment from mental health teams, a patient highlighted in an interview that respect is not always evident. This finding was supported by the comments of an occupational therapist.

> *"Just the way someone speaks to you honest they speak to you like you're on their shoe …, I honestly don't believe at any point you need to speak down to, or insulting to someone with mental health issues." (Patient – Interview)*

> *"You know when people open doors without knocking, or open curtains without getting permission… It's low-grade stuff, but it's poor practice." (OT – Focus group)*

During an interview, one patient discussed their anxiety at making telephone calls. The patient highlighted that their nurse challenged them to make calls, which the patient appreciated; the patient stipulated however, that they must only be challenged on something that they want or need to do.

> *It's finding … something I need to do or want to do and pushing me to do that, as opposed to pushing me to do whatever they think most people who are depressed or hear voices or whatever do. (Patient – Interview)*

It is therefore important that professionals deliver person-centred care, whilst attending to the differing needs and wishes of service users.

5.4.7   Problem solving

Summary: Practitioners must be able to reason with abstract information. Professionals also must understand problems from a wider perspective, having the ability to adapt their practice.

The professional attribute 'Problem Solving' generated five subcodes (as demonstrated in Table 5.11). The nature of mental illness means that you cannot just look at someone and diagnose them; peoples' experiences are not clearly visible to the naked

eye. Professionals must therefore conceptualise with abstract information, as highlighted by the following comment.

> *"The word schizophrenia does not describe people with schizophrenia, you know everybody's different and I think you have to have that ability to conceptualise and think outside the box a bit in mental health that you maybe don't have to do so much in sort of medical care." (Nurse – Focus group)*

Professionals must adapt their practice, dependent on the current circumstances.

> *"There is no point being professional if somebody's crying." (Nurse – Focus group)*

Table 5.11 Corresponding codes for the attribute Problem Solving

Subcode	Number of interviews and focus groups	Number of references
Is able to use abstract reasoning	3	3
Helps patients problem solve	2	2
Makes appropriate decisions based on all the relevant information	3	5
Understands problems from a wider perspective, taking account of the whole picture	5	11
Uses initiative and adapts practice to meet people's needs	7	13

5.4.8 Self-awareness and insight

Summary: Practitioners must possess self-awareness and acknowledge the limits to their competence. It is appropriate to disclose some information about self, but professionals must recognise and maintain boundaries when doing so.

The professional attribute 'Self-Awareness and Insight' generated four subcodes (as demonstrated in Table 5.12). Self-disclosure is known to break down some of the

boundaries that exist in a staff - patient relationship; the level of disclosure must be appropriate for each patient and must cater for an individual's needs at the time.

"It also seems a bit of a balance sometimes because you expect a patient to share so much of themselves… And then you kind of go, but I won't tell you anything about me." (Psychologist – Focus group)

It was highlighted during a focus group that professionals may send blunts emails when emotions run high, even though they would not typically do so. At these times, it is important that professionals reflect-in-action.

"I've seen a fair few emails that maybe come across as disrespectful and unprofessional, because they're emotive emails about a topic that people have strong opinions on… I think professionalism is about taking a step back and thinking right, I'm quite annoyed about this, maybe I should send this later or ask someone else to look over it." (Psychologist – Focus group)

Table 5.12 Corresponding codes for the attribute Self-Awareness and Insight

Subcode	Number of interviews and focus groups	Number of references
Acknowledges the limits to one's competence	4	6
Discloses appropriate amounts of information about self	4	7
Possess self-awareness and reflects on practice	5	12
Recognises and maintains appropriate boundaries	10	25

5.4.9 Working effectively as part of a team

Summary: Practitioners must work alongside colleagues effectively, acknowledging people's strengths and capabilities. Professionals must also be a positive role model.

The professional attribute 'Working Effectively as Part of a Team' generated seven subcodes (as demonstrated in Table 5.13). Discussion arose during a focus group

regarding how professionals are a role model for students; the detrimental impact negative role modelling can have on students, as well as newly qualified professionals was highlighted.

> *"When they're qualified, you can't go home and have a cup of tea and I think that sometimes with students, … we set the wrong examples." (OT – Focus group)*

Professionals form part of a team and must work together to achieve common goals.

> *"I think that it is very important, working as part of a team, everybody's working in the same direction, with the same values." (Patient – Interview)*

> *"you know we work with everyone around us and I can only work well if I have a really good team around me." (Psychiatrist – Focus group)*

Table 5.13 Corresponding codes for the attribute Working Effectively as Part of a Team

Subcode	Number of interviews and focus groups	Number of references
Is able to identify and utilise the most appropriate person for a task	3	5
Is a positive role-model for others	6	9
Maintains appropriate relationships with colleagues	2	6
Maintains consistency with colleagues	5	7
Supports colleagues	3	5
Works effectively with other teams	8	20
Works effectively with colleagues	11	23

5.4.10 Working with carers

Summary: Professionals must involve carers where possible, whilst adhering to the bounds of confidentiality. Carers generally want to be involved and should feel supported in their role.

The professional attribute 'Working with Carers' generated three subcodes (as demonstrated in Table 5.14). It is recognised that not all carers are caring, and professionals must remain aware of this.

> *"We shouldn't imagine that all carers fit into that box of being desperately caring, … I think carers as some people said can be part of the problem."*
> *(Carer 1 – Focus group)*

> *"Yes, yes, exactly" (Carer 2 – Focus group)*

> *"Well carers can also be one of the causes of the caree's problem altogether and I can understand the professionals… have to assess you very quickly as to whether you're actually the one that is causing the problem or if you are genuinely trying to help." (Carer 3 – Focus group)*

Table 5.14 Corresponding codes for the attribute Working with Carers

Subcode	Number of interviews and focus groups	Number of references
Assesses a carer's motives	1	1
Involves carers and makes use of their expertise	3	9
Supports and validates the carer's perspective	1	8

During a focus group, it was expressed that professionals at times hide behind policy. It was explicitly stated that most carers want to be involved, yet they feel left to the wayside. Professionals must work with carers, where possible, to improve patient care.

> *"the majority of carers want to be involved and want to help and get very fed up of being left in a corner." (Carer – Focus group)*

> *"it should be exploited and it isn't and that's what I find is amazing." (Carer – Focus group)*

"my experience is you have to bang on doors, you have to make a nuisance of yourself to actually be listened to." (Carer – Focus group)

To facilitate cross referencing, all attributes are listed in Table 5.15, alongside the corresponding expectations of professionals working in mental health services.

5.5   Discussion

It is widely cited that professionalism forms the basis of a social contract between professions and society (see chapter 4, section 4.10.1.2). A profession's main activities are protected and governed by licensing laws, which in turn, allow its members to practice (Cruess, 2006). Professions are expected to self-regulate (Cruess, 2006), but they must incorporate the views of patients they serve, who place their trust in the profession, when doing so (Irvine, 1997). As Bhugra and colleagues (2008b; 2010) suggest, the social contract must be renegotiated regularly, however, the views of society must be incorporated during this activity.

The aim of the current study was to explore the attributes desired of professionals working in mental health and learning disability services. During the analysis, it became apparent that the attributes desired are similar to that reported in the Professional Attributes Framework (Medical Schools Council, 2011). As the analysis progressed, a further research question arose: 'how does the experience of participants align to medically defined, generic, professional standards and attributes?.'

5.5.1   Main findings

Whilst the findings of this study resonate with that listed in the Professional Attributes Framework, there are differing emphases, and carers must also be factored into the equation. Self-awareness and insight were observed to be key attributes for mental health professionals as they help professionals maintain a professional stance with patients. Another professional attribute observed over the course of the study was effective communication. Whilst the importance of effective communication is evident across professions, and healthcare settings, this is particularly important in mental

156

health services where patients may be severely unwell and lack mental capacity, resulting in them having additional difficulties communicating.

Table 5.15 Professional attributes and expectations of staff working in mental health services (adapted from Medical Schools Council, 2011).

Professional attribute	Associated expectations
1. Commitment to Professionalism	*Adheres to guidelines; Behaves according to expectations; Challenges the system accordingly; Challenges poor practice; Displays a commitment to the role; Is trustworthy and has integrity; Maintains confidentiality; Is a responsible practitioner; Possesses confidence and courage; Demonstrates awareness of ethical issues; Upholds the profession's and organisation's reputation; Uses Social Media appropriately.*
2. Coping with Pressure	*Checks the facts of a case; Utilises clinical judgement accordingly; Utilises de-escalation techniques appropriately; Possesses resilience and manages own wellbeing; Remains calm and in control of situations.*
3. Effective Communication, with Patients, Carers and Colleagues	*Utilises an appropriate style of communication; Advocates for patients when needed; Builds rapport with patients and is personable; Communicates effectively with colleagues; Is open and honest, whilst communicating in a proactive manner; Communicates sensitively, taking context into consideration; Listens effectively; Observes accordingly; Understands non-verbal communication; Utilises appropriate non-verbal communication; Validates the thoughts and feelings of others.*
4. Learning and Professional Development	*Is accepting of feedback; Applies knowledge and learning to practice; Utilises supervision; Has the appropriate knowledge and skills for the role; Undertakes continuing professional development and learns from practice; Possesses the relevant qualifications for the role; Undertakes research.*
5. Organisation and Planning	*Is efficient in the role; Maintains accurate records; Reads patient's case notes; Is able to make appropriate use of limited resources; Manages time accordingly; Wears appropriate attire at work.*
6. Patient Focus	*Attends to patients' physical healthcare needs; Acts as a human interface for the organisation; Builds therapeutic relationships; Demonstrates compassion; Contains the emotions of others; Does not impose own values on patients; Is friendly, but maintains an appropriate professional distance; Is altruistic and possesses humility; Is approachable; Is empathic and understands the impact of mental illness; Is genuine, honest and fulfils promises, Is non-judgemental, Provides reassurance, Maintains Safety, Treats Patients with respect, Utilises both a person-centered and recovery-focused approach*
7. Problem Solving and Decision Making	*Is able to reason in an abstract manner, Helps patients to problem solve, Makes appropriate decisions based on all the relevant information, Understands problems from a wider perspective, Uses their initiative and adapts practice, to meet a patient's needs*
8. Self-Awareness and Insight	*Acknowledges one limits, Discloses appropriate amounts of information about self, Has self-awareness and reflects on practice, Recognises and maintains appropriate boundaries*
9. Working Effectively as Part of a Team	*Able to identify and utilise the most appropriate person for a task, Acts as a positive role model, Maintains appropriate relationships with colleagues, Maintains consistency with colleagues, Supports colleagues, Works with other teams, Works effectively with colleagues*
10. Working with Carers	*Assesses a carer's motives accordingly, Involves carers and makes use of their expertise, Supports and validates the carer's perspective*

It is acknowledged that carers aren't unique to mental health settings, however they clearly play a pivotal role in the lives of many patients with mental health problems. Carers highlighted the impact that mental illness can have on them personally, - "*the whole family changes, you're traumatised, you are not sleeping...then also you are faced with a professional who regards you as a problem*." As previously noted by Cleary, Freeman, and Walter (2006), an essential part of modern mental health service delivery includes clinicians interacting with carers and providing information and support to them. Nevertheless, carers report being unsupported in their role. When a patient is unable to advocate for themselves and their mental capacity is impaired, the importance of involving carers is paramount. As a carer clearly articulated, "*if you are really doing the best for your patient, you have to be able to look at their social network, particularly in mental health*" (Carer).

5.5.2 Comparisons with the existing literature

The study found that, amidst the 10 professional attributes, many expectations are placed on professionals (see Table 5.15): professionals must have integrity and challenge poor practice; they must use their clinical judgement in times of uncertainty and ambiguity; they must communicate effectively; they must have organisational skills, managing their time and limited resources effectively; they must establish therapeutic relationships with patients; they must be able to adapt their practice, taking account of the bigger picture; they must possess self-awareness; and work effectively with colleagues; whilst also supporting carers. Meanwhile, professionals face many challenges to their professionalism, such as negative role modelling, social media, limited resources, and being overworked. During a discussion regarding the impact of ever-increasing funding cuts, a psychiatrist commented "*I think we are regarded at a level that we are, because we don't let bad things happen, because we do paper over the cracks all the time*." To overcome the challenges that professionals face, humour, including occupational and gallow's humour is used. Camaraderie has previously been condoned by medical undergraduates as a legitimate means of

diffusing stress (Finn et al., 2010, p.819). Freud has previously theorised that humour is a 'mature defence mechanism' (Freud, 1905); according to Freud's theory, the mind creates humour to decrease anxiety and increase coping in a situation that may feel threatening (as cited by Averitt, 2005). Acknowledging the complexity of humour, it has been recommended previously that the topic of humour is included in the social work curriculum (Moran and Hughes, 2006). Professionals argue that humour is vital for their wellbeing, yet, in the current study, many referred to its context dependence and sensitivity, noting how detrimental this can be in front of a patient, especially if the patient has a trauma history.

Whilst the data have been reported on according to the ten professional attributes, similarities can be observed between the data and the definition of professionalism previously proposed in chapter 4 (see section 4.11), thus highlighting the suitability of the definition for a mental health services context. For example, the attribute 'Commitment to Professionalism' is congruous with 'Intrapersonal Professionalism' – expectations held of individuals in order to meet the expectations of their profession; 'Patient Focus' and other attributes align to the concept of 'Interpersonal Professionalism' – possessing the necessary skills to relate to others in an appropriate manner; and 'Coping with Pressure' is harmonious with 'Working Professionalism' – the ability to form judgements and act accordingly, thinking critically and using reflection in action.

The context dependent nature of professionalism and the need for situational judgement has previously been highlighted by Burford et al., (2014). Whilst the study by Burford et al., was not conducted in mental health services, their findings resonate with several themes identified over the course of the study: for example, as highlighted by the attribute 'Learning and Professional Development,' where you say something determines whether it is professional or not; also, as observed with the attribute 'Effective Communication,' humour and banter may not be appropriate when

used in front of patients with trauma histories. Such behaviours are therefore dependent on contextual, patient related factors.

The importance of carer involvement is reported in the Triangle of Care guidance document, which provides standards and resources for mental health professionals to assist them with including and supporting carers (Worthington et al., 2013). Relatives want information, involvement in decision-making and supportive staff with whom they can communicate (Noble and Douglas, 2004), yet carers in the current study report being shunned by professionals, who won't even talk to them, because '*the patient has not given consent for the carer to be involved*'. These results add further support to the earlier findings of Cleary et al. (2005), that over 50% of carers report that information, such as that regarding medication, illnesses and community resources, is not provided to them. Green et al. (2009) previously found that patients and practitioners did not always place the same regard on certain professional behaviours. This finding was echoed in the current study; patients did not discuss the relevance of professional standards or the appropriate application of social media, there were also differing views between a patient and professional regarding documentation.

Practitioners are expected to be professional around the clock, which means that they can never be seen to be unprofessional. The study found that professionals have been observed to use a pseudonym on Facebook in order that they could express views that they would otherwise feel unable to. The experiences of professionalism on social media platforms described in this study resonate with the idea of a 'virtual mask' previously proposed by Finn et al. (2010). It was recognised that whilst professionals are a role model for students, they at times may teach students bad practice, such as going home early, or going shopping whilst at work. This highlights the nature of the 'hidden curriculum'; that is, that there are differences among what students are taught in the classroom as opposed to what they may pick up on placement (Hafferty and

Franks, 1994). This raises alarm, as the observation of and participation in unethical conduct has been found to result in ethical erosion overtime (Satterwhite et al., 2000). It has also previously been identified that problematic behaviour during training can predict disciplinary action in later clinical practice (Papadakis et al., 2004).

The data highlighted that professionalism was more noticeable by its absence, than its presence, with professionalism typically being construed as about 'what *not* to do.' Similarly, it has previously been reported that there may be greater consensus regarding what is considered inappropriate, as opposed to appropriate behaviour (De Leng et al., 2018). Whilst professionals appeared open in the content of their discussions during focus groups, it is possible also that some professional behaviours may not have been disclosed, particularly those of an unprofessional nature. The regulation of the professions may influence such censorship, which aligns with the theory proposed by McLachlan (2017) regarding 'Pious Platitudes' about professionalism; McLachlan highlights that when practitioners are asked to define professionalism, practitioners may respond with what they think they ought to say, rather than with what they have actually observed. Focus groups took place on NHS premises, and during work hours; thus, some participants may have been on their 'best behaviour.' This is evidenced by a professional stating *"we can all be unprofessional now"* as a focus group was coming to an end.

5.5.3 Interpretation of findings

Jones, McCullough and Richman (2006) suggest that the Physician charter does not consider many elements of surgery practice, which may pose specific ethical and professional dilemmas for the specialty. Indeed, various medical specialties have highlighted that the concept of professionalism must be refined in order for this to be pertinent to their practice (Woodruff et al., 2008). Given the nature of mental health services, it is not surprising that the current study identified a tenth professional attribute - Working with Carers. It is important to note that carers were not involved, directly, in the development of the 'Improving Selection to the Foundation

Programmes' Professional Attributes Framework. However, there are nuances in mental health services that are not present in other healthcare specialties: for example, detained patients lose their patient autonomy, one of the three fundamental principles of medical professionalism (Project of the ABIM Foundation, 2002); and patients using mental health services are vulnerable (Department of Health, 2000, updated in 2015), often relying on carers to protect and advocate for them. Carers may spend significant amounts of time with the patient and may thus be able to provide insights that a professional would not be aware of otherwise. The Triangle of Care guidance document highlights this point well, "*Carers are usually the first to be aware of a developing crisis – often at times when professional help has not yet been established or is unavailable. They are often best placed to notice subtle changes in the person for whom they care, and usually the first to notice the early warning signs of a relapse*" (Worthington, et al., 2013, p.7).

There are many challenges for professionals working in mental health services. Discussing the challenges professionals face, a psychiatrist remarked "*isn't our status because we do, we make good a bad system … I think we are regarded at the level that we are regarded because we don't let bad things happen*." Patients may be criminals, including sexual offenders; yet the professional has a duty of care and must do their best by the patient, despite their own feelings. Patients, also, may commit suicide, which likely affects a professional's own wellbeing. As the results demonstrate, many expectations are placed on professionals working in mental health services; indeed, the current study sought to highlight these. Nevertheless, the requirements of the professions themselves must also be considered, including a properly funded and value driven healthcare system (Bhugra, 2008b; Cruess, 2006). The social contract requires patients, also, to play their part, having a 'shared responsibility' for their own health and wellbeing (Bhugra, 2008b; Cruess, 2006). To facilitate the level of care that's desired, patients and carers must work with professionals; a couple of examples of

which could be turning up to their appointments on time and being polite to professionals.

The lack of discussion with patients regarding professional standards and social media does not mean that these topics are not important to service users. It is possible that the topic guide, or the use of interviews instead of focus groups influenced the lack of discussion in these areas. Rather than imposing their own perspective, the researcher attempted to understand the data from the participants' point of view; thus utilising 'empathetic understanding' (Holloway and Galvin, 2017, p.6). The researcher also reflected on their own views during the process of analysis. With an insider perspective, from the standpoint of a patient, it is possible that the researcher related more to the experiences of these participants, in comparison to the views of professionals. The insider perspective has its own advantages; for example, it can give greater insights as the researcher is already familiar with the topic in question; however, the researcher may also have preconceptions, which can impact upon the opportunity for generating new knowledge (Holloway and Galvin, 2017). Recognising that the researcher is an active participant in co-constructing meaning during interviews (Holloway and Galvin, 2017), two individuals were involved in the analysis of data in order to minimise bias. Two additional researchers also reviewed the codes.

Conducting the study in an NHS setting meant that the researcher could observe interaction within the setting of interest. Rather than being interviewed in their own homes, patients met the researcher in NHS premises, thus minimising the likelihood of distractions, as well as the level of risk imposed. The data obtained during interviews needs to be considered with regard to the social context where the interview took place (Manderson et al. 2006). In some instances, patients had become aware of the study via their professional worker. It is recognised, by the researcher, that even if participants were advised that their participation is voluntary, they may have felt obliged to take part, due to their relationship with the professional (Holloway and

Galvin, 2017). Patients frequently portrayed an extremely positive image of the professionals that they worked with; whilst it is possible that this reflected their experience of care in this setting, it is possible that participants felt obliged to respond in this manner to avoid endangering their relationship with the professional in question. This concern has previously been highlighted by Holloway and Galvin (2017). Professionals may also have been selective with regards to which patients they recommended the study to. The researcher, consciously, dressed down during interviews to minimise, where possible, any perceived power imbalance. Whilst the researchers sought to conduct focus groups and interviews of a homogenous nature, intentionally separating patients, carers, and the professional groups, it became evident, in some instances, that patients had also worked as professional members of staff, and some professionals were also carers. This reflects the true nature of healthcare services; that healthcare professionals are not immune to mental health problems, and in many cases bring their own lived experience.

5.5.4 Strengths and limitations

This is the first study to explore professionalism from the experience of patients, carers, and professionals across a mental health services context. The findings are enhanced by the involvement of patients and carers, voices underrepresented in the professionalism literature. Focus groups and interviews were combined during the study for practical scheduling purposes, as well as the ethical issues noted in the methods section. Some researchers criticise the combining of these two approaches, however, this issue remains contested. Lambert and Loiselle (2008) argue that a combination of individual interviews alongside focus groups can contribute to a productive iterative process and actually enhance the trustworthiness of findings. An interview guide was used during both interviews and focus groups. Like Lambert and Loiselle's (2008) study, not all prompts were required during the focus groups to elicit the information desired; however, the guide enabled a semi-structured approach and

minimised the likelihood of interviewer bias. Four focus groups were cofacilitated by a second researcher that subsequently reviewed the codes generated during analysis.

Individuals could only participate if they were cared for in the community and were not current inpatients. Holloway and Galvin (2017) have previously highlighted that participants may, at times, be too emotionally involved to make rational decisions regarding taking part and continuing with a study. This was felt to be the case with this study, particularly due to the nature of the questions being asked. It is worth noting that a service evaluation was undertaken, and attempts were made to collect data from NHS complaints records that could have fed into the overarching results; the data were not sufficient to add to the findings however, due to the level of detail that was removed during the Trust's anonymisation procedure.

When convenience sampling is utilised, caution needs to be paid regarding the generalisability of the findings (Bryman, 2016). Nevertheless, a purposive sample was obtained, and all participants met the desired criteria; that is, they were either patients, carers, or registered professionals working in mental health services. Whilst the study was conducted in one NHS Trust only, this was a large organisation. It became apparent during the focus groups and interviews that there were geographical variances between where professionals and patients had trained, worked, and/or received healthcare previously. Professionals worked across a range of settings, including services for children and adolescents, adults, the elderly, people with learning disabilities and forensic services. The research team argue that these findings are therefore representative of, and transferable to, mental health services nationally. It is important to add that a reflexive approach was taken throughout, acknowledging the researchers' context, potential biases, and presuppositions (Berger, 2015).

5.5.5 Implications for policy and practice

Whilst the confidentiality of patients must be upheld, the findings demonstrate the importance of clinicians supporting, and working with carers to best deliver patient

care. Lloyd and King (2003) have previously proposed various strategies to facilitate staff working with carers, which the current author advocates. Such strategies include managers openly asking staff about their involvement with carers, collaboration being a key feature of performance appraisals, and staff receiving appropriate education on the topic. The findings of this study could support the development and delivery of training curricula for mental health professionals, or those seeking a career in this field.

Some unprofessional behaviours were highlighted in discussions about staff being negative role models to students; these included a range of behaviours, such as being critical of a patient or sending students home early. Previous research highlighted that approximately a quarter of students found the school environment 'not very conducive' or 'not at all conducive' to the open discussion of ethical concerns (Satterwhite et al., 2000). Henceforth, it is important that more sense-making opportunities are made available to students as part of the formal curriculum, thus allowing students to engage in open dialogue and transformation, whilst under the supervision of clinical educators (Monrouxe et al., 2011).

The findings of this study would help facilitate the development of assessment tools for the selection of staff into mental health services. Such assessments may include SJTs, which was the purpose of this study. Health care regulators, too, would benefit from considering these findings when revising and updating current standards and inspection guides.

5.5.6 Implications for future research

Based on the data, it is difficult ascertain what patients' views are regarding professional standards and social media. Researchers may wish to address these topics directly with service users in future studies. The current study sought to identify the key professional attributes across all professions working in mental health services; any attempt to distinguish the differences among professions was beyond the scope of the current study. Further research could use the current findings in order to

determine to what extent the individual professions endorse each of the professional attributes and their associated expectations.

Professionals rely on patients to engage with them, in order that the professional can provide the level of care required. If professionals are expected to adhere to a social contract, what might the patient's side of this entail? For example, are patients expected to adhere to all treatment recommendations? Future research could explore the professionalism of patients using mental health services, whilst also exploring the impact that this has on the clinical care delivered. Such understanding could facilitate the development of more tailored packages for patients and more educational interventions for staff.

5.5.7 Conclusion

The study found that professionals need to fulfil the expectations across 10 professional attributes in mental health services, which includes one additional attribute to that specified in the original 'Improving Selection to the Foundation Programmes' Professional Attributes Framework. In mental health services, professionals must have the ability to 'work with carers'. Often, patients rely on carers to protect them, and advocate on the patient's behalf. Carers often spend significant amounts of time with the patient and may offer insights that a professional would not be aware of otherwise. Professionals are therefore urged to support carers and provide a forum for them to air their views. In a mental health crisis, a carer's views could be lifesaving.

Chapter 6: Evaluating the Criterion Validity of a Situational Judgement Test for Assessing Professional Judgement in Mental Health Services

Situational judgement tests (SJTs) are used extensively in medical selection, as well as other disciplines and occupations to assess an understanding of interpersonal effectiveness in a work context. It is unknown, however, whether SJTs would be effective for the selection of staff into mental health services as a means to facilitate NHS 'values-based recruitment' (VBR).

The findings of a systematic review and qualitative study that explored the concept of professionalism in mental health services are reported on in chapters 4 and 5 of this thesis. Utilising earlier findings, this chapter reports on the development and validation of an SJT to assess 'professional judgement' for a mental health and learning disability services context, thus meeting objective (3) of the overarching research project.

6.1   Background

SJTs are an assessment method, often used in high stakes testing situations, such as personnel selection and professional examinations. SJT developers thus need to provide evidence of their validity to justify their use in such settings. Whilst many different aspects of validity exist (Mohajan, 2017), it is mainly criterion-related validity that developers and end-users of SJTs are interested in. That is, do the scores produced by the SJT predict criterion related to workplace effectiveness? Prior research demonstrates that SJT scores generally predict aspects of actual job performance (Mcdaniel et al., 2007). This is also true where SJTs are used as part of selection into the various stages of medical training (Webster et al., 2020). Furthermore, SJT scores may often, though not always, provide incremental validity to scores derived from assessments of cognitive ability and personality traits when predicting job performance (Chan and Schmitt, 2002; Clevenger et al., 2001). SJTs are

also attractive to selectors in that they are relatively cheap to deliver at scale as a component of 'values-based recruitment.'

6.2   Aims and objectives

6.2.1 Aims

An SJT would be a potentially valid method for assessing knowledge of professionalism amongst staff working in mental health services. Therefore, the primary aim of the study was to evaluate the criterion-related validity of scores resulting from an SJT that was developed to assess professional judgement in this setting. Aspects of convergent and divergent validity were also evaluated. The feasibility of any test is also important. That is, can the test be used to obtain enough information to discriminate accurately between candidates in an acceptable time frame? The study thus aimed to, also, determine the total number of items required for an SJT that would suitably predict staff members' *professionalism* and *effectiveness*.

6.2.2 Objectives

1.   To evaluate the criterion-related validity of the SJT scores using a variety of commonly employed scoring systems, including: raw Subject Matter Expert (SME) polytomous scoring (range 1 to 4), raw SME binary scoring (0 or 1), dichotomous modal consensus scoring (0 or 1) and proportion modal consensus scoring (a continuous proportion between 0 and 1).

2.   To explore the relationship between SJT scores and professional and demographic characteristics.

3.   To evaluate the convergent and divergent validity of the SJT scores by exploring relationships with the scores from The Big Five Inventory–2 Short Form (BFI-2-S) personality assessment.

4.   To assess the acceptability and feasibility of the SJT, and in turn, obtain information to guide selection of the items for the final version of the SJT assessment.

For this study, the criterion-related validity of the SJT's scores were assessed by evaluating their relationship with colleague ratings of job performance; specifically, colleagues were asked to provide ratings of perceived '*professionalism*' and '*effectiveness*.' It was assumed that these ratings of *professionalism* and *effectiveness* would be positively correlated with one another, given the importance of professionalism in the effective delivery of mental healthcare services. The criterion-related validity of the assessment would be evidenced, by a statistically significant, and substantively meaningful, positive association between SJT scores and ratings of both *professionalism* and *effectiveness*. That is, an observed relationship that is unlikely to be due to chance alone, and similar to that for other, similar SJTs and selection tools. For example, prior meta-analytic research that sought to evaluate the validity of SJT scores for personnel selection observed an estimated population validity of .34 across a range of measures and samples (McDaniel et al., 2001).

SJT scores have been found to correlate, at least modestly, with various personality traits, including *emotional stability*, *conscientiousness*, and *agreeableness* (McDaniel and Nguyen, 2001). As the current SJT was developed to assess ones' knowledge of professional behaviours for mental health services, it was anticipated that there would be significant positive associations observed between SJT scores and ratings of *conscientiousness* and *agreeableness*, as assessed by the BFI-2-S. These traits are desirable in practitioners working in mental health services, where patients expect professionals to have compassion and be responsible and respectful to colleagues, patients, and carers. It is possible that there would also be a significant negative association between SJT scores and *negative emotionality*; one would expect individuals that are more stable and emotionally resilient to be more professional and effective in mental health services, especially given the emotional challenges that staff face on the job. Analyses also explored whether SJT scores had incremental validity over scores obtained on a personality assessment in explaining supervisor ratings of perceived *professionalism* and *effectiveness*. A secondary aim of the study was to

evaluate for evidence of test bias related to gender and ethnicity among other characteristics.

As noted earlier in this thesis, various scoring approaches have been used for SJTs previously, including; modal scoring, 'lenient mode' scoring, dichotomous scoring, distance scoring, and adjusted distance scoring (MacCann et al., 2004; Weng et al., 2018). De Leng et al. (2017) compared 28 different scoring approaches for a rating scale format SJT and concluded that the final scoring method used for an SJT must be subject to a thorough examination. Therefore, the current study also explored whether various scoring approaches influenced the relationship between individuals' SJT scores and ratings of job performance. It is recognised that there is significant variation in the duration, content and culture of training and practise for each professional group, and the expectations placed on each of the professions vary (see chapter 4, section 4.10.1.4). It was unknown, to what extent, if any, that the SJT's scores would be valid across different mental health disciplines. However, this was an issue that was able to be explored, to some extent, as part of this study.

6.3   Methods

This chapter primarily reports on the piloting of an SJT that was developed for a mental health services context. Whilst reference is made to the development of the assessment during this chapter, the ethics statement here applies to the pilot study only. An ethics statement pertaining to the involvement of participants in the development of the SJT can be found in section 5.3. The pilot study was performed with staff; as patients were not involved in this part of the project, approval was not required from an independent NHS ethics committee. Approval was, however, obtained from Hull York Medical School (Appendix H), the Health Research Authority (REC ref: 19/HRA/6403; Appendix I) and Tees, Esk and Wear Valleys NHS Foundation Trust (TEWV NHSFT; R&D reference: 0619/19).

There were originally plans to pilot the SJT with core psychiatric trainees that were participating in the Independent Assessment of Clinical Skills programme (IACS). The IACS is a role play 'high fidelity' simulation. Such a study was hoped to provide additional validity evidence as it would help determine whether the SJT scores were associated with ratings of clinical skills (e.g., communication) in psychiatric trainees in this setting. These plans could not be conducted however as the Covid-19 pandemic prevented the IACs from going ahead on the dates originally scheduled. Thus, supervisor and personality ratings provided support for the convergent, divergent, and criterion-related validity of the SJT scores.

6.3.1 Study design and context

This observational, concurrent, criterion-related validity study was conducted in TEWV NHSFT. At the time of the study, TEWV provided mental health and learning disability services across County Durham, the Tees Valley, York and North Yorkshire (Tees Esk and Wear Valleys NHS Foundation Trust, 2021).

6.3.2 Procedure

The study aimed to develop and evaluate the validity of an SJT to assess professional judgement for mental health services. This aspect of validity would be crucial to establish whether the SJT were to be implemented in a high-stakes personnel selection setting in the future. The study was open to all professionals that worked in a mental health and learning disability NHSFT, as long as they were registered with a professional, clinical regulatory body. In a couple of instances, however, there were senior, experienced staff that took part who were not technically registered with a professional body (e.g., a chaplain). Participants that completed the SJT were asked to leave contact details for their manager or supervisor and up to three colleagues so that workplace related feedback could be obtained for them. The primary outcome measure was evaluations of '*professionalism'* and '*effectiveness'* as rated by colleagues and supervisors/managers. A brief personality assessment, collecting both self- and third-party ratings was included as a secondary outcome. A variety of recognised,

previously described scoring systems were explored in relation to 'reliability' and validity.

6.3.3 Materials and measures

Participants that completed the SJT were required to complete a survey (see Appendix J), which incorporated three main elements. First, participants were asked for information regarding their demographic and professional characteristics; second, individuals were required to complete SJT items; and third, individuals were asked to provide self-ratings on a personality inventory before leaving contact details for their colleagues, including their manager and/or supervisor. Please note that only one version of the survey is incorporated in the appendix and SJT items that are scored in the final test have been omitted from this.

6.3.3.1 *Demographics*

The Director of Human Resources in the participating research site wanted to understand more about any potential discrimination against individuals according to their equality characteristics. Many demographic details were therefore captured during the study so that subsequent analyses could explore any potential discrimination against certain groups; the demographic details collected at the start of the survey included age, gender, sexual orientation, whether individuals considered themselves to have a disability, ethnicity, religion, and whether English was their first language.

Professional characteristics were also requested during data collection because it was vital to establish for what staff groups the SJT scores were likely to be most valid. The characteristics requested included the years individuals had been registered to the profession, their grade (according to Agenda for Change; AFC), and both the setting and specialty they worked. This was important to capture as, previously, experience has been both positively and negatively related to SJT scores (Groothuizen, 2020; Wagner and Sternberg, 1985; Weekley and Jones, 1999; Weekley and Ployhart, 2005).

It should be stressed that those rating individuals' *professionalism* and *effectiveness* were specifically instructed not to take experience of the participant into account when scoring. This was because it was absolute, rather than relative, perceived *professionalism* and *effectiveness* that was required for the validation study. Thus, third party raters who 'made allowance' for relative inexperience etc would add unwanted bias or noise into the data.

6.3.3.2 *Development of an SJT to assess professional judgement in mental health services*

Chapter 5 reports on a qualitative study that sought to explore the concept of professionalism from the views of key stakeholders. The study formed two parts: first, individuals were asked to describe what professionalism means to them; and second, individuals were asked to provide examples of professional and unprofessional behaviour. The latter questions were used, primarily, to facilitate the development of item content for the SJT. Having determined that an SJT would suitably assess essential constructs that are required for working professionally in mental health services, the literature was reviewed to explore the design options for this assessment method.

Design Specification

To develop appropriate SJTs, it is important to consider the institution's philosophy, the community that the organisation serves, and the system in which the assessment will be used (Patterson et al., 2018). During SJT development, multiple meetings were held with the Director of Human Resources and Organisational Development in the participating mental health Trust to discuss the development of the SJT, noting the organisation's values, principles, business objectives and how they anticipated using an SJT as part of personnel selection. These meetings facilitated the development of a test blueprint and, in turn, enhanced the content validity of the test (Bridge et al., 2003).

SJTs typically have two types of response instruction; there are questions that ask what one 'should do,' and questions that ask what one 'would do,' in a given situation.

Responses to the former typically correlate with scores obtained on knowledge tests and cognitive ability, whereas responses to the latter are more associated with personality traits (Mcdaniel et al., 2007). With regards to selection and recruitment, it is important to know how individuals would respond in a given interpersonal situation that could plausibly arise in the workplace; however, SJTs that pose the question 'what would you do' are prone to faking or social desirability bias in high stakes settings (Peeters and Lievens, 2005). For example, when asked 'what one would do,' it may be relatively easy in some cases to determine which response would make the best impression. Studies have demonstrated that in high stakes selection contexts individuals are more likely to amend their answers to reflect the most desirable response and would therefore be more likely to respond as to what one 'should do', even though they may be asked 'what would you do' (Motowidlo et al., 2006b). To ensure fairness, a decision was made to ask knowledge-based questions within the current SJT, recognising that this is more likely to tap into individuals' *maximal* performance, rather than their typical performance. This was deemed a suitable approach as it is unlikely that a professional would respond appropriately, in a given clinical situation, if they did not know what the appropriate course of action was. Individuals that know what behaviours are effective in a given role are understandably, more likely to perform effectively (Motowidlo et al., 2006b). Thus, in this sense, the SJT could be considered a special case of a knowledge test about appropriate or effective workplace behaviours. Such knowledge could be considered a necessary, though not sufficient condition for actually exhibiting such behaviours in actual practice (Tiffin et al., 2020).

While it is important to elaborate on scenarios in an SJT for a respondent to have enough information to decide how best to act, it is also recognised that too much information can result in test-taker fatigue (Brooks and Highhouse, 2006). Efforts were therefore made to limit the amount of detail provided within each SJT item. Prior analyses have revealed that psychometrically, using Item Response Theory, it is often

not possible to discriminate between more than three or four points on an SJT item (Tiffin and Carter, 2015); a decision was therefore made to utilise a four-point response scale. The latter approach would also help minimise the impact of bias due to extreme response style. That is, a preference to select categories at the extreme ends of a rating scale (e.g., selecting very appropriate as opposed to appropriate, but not ideal). Indeed, extreme response style is recognised as a source of potential bias in relation to ethnic minority candidates (Batchelor and Miao, 2016).

Motowidlo and colleagues have previously developed single response SJTs and found these to be a valid assessment of procedural knowledge that predicts job performance (Crook et al., 2011; Motowidlo et al., 2009). Single response SJTs provide the respondent only one response at a time to consider. This format can be a useful alternative to the multiple response format, where respondents may have to select several answers from a list (e.g., best three, or best and worst response etc). Thus, a single-response format was used for the current SJT in order to facilitate the development of a higher number of SJT items within a limited timeframe. It is also easier to mathematically model single response format questions, for example, using Item Response Theory. Video format SJTs have sometimes demonstrated increased predictive validity in comparison to 'paper and pencil tests' (Christian et al., 2010; Lievens and Sackett, 2006). However, the latter method (in an electronic format) was used for the current study to minimise development costs. It is worthy to note that scores from paper and pencil SJTs generally demonstrate high levels of predictive validity also. For example, Motowidlo et al. (1990) found that scores on SJT-like paper and pencil test had correlation coefficients that ranged from .28 ($p < .01$) to .37 ($p < .01$) with supervisory ratings of performance. Thus, the increased costs of video based SJTs may be difficult to justify in many contexts.

A decision was made to utilise the response instructions used in the current SJT employed by the University Clinical Aptitude test (UCAT- formerly the UKCAT)

176

The UCAT is a computer-based test that is used by universities across the United Kingdom (UK), and in parts of Australasia to help select applicants onto medical and dental undergraduate degrees. The UCAT SJT was piloted in 2012 and first introduced into selection in 2013. The context domains are labelled as 'integrity,' 'team working' and 'perspective taking.' The scores demonstrate an ability to predict tutor ratings on relevant constructs (Patterson et al., 2017), and also the risk of disciplinary action being subsequently taken against an applicant entering medical school (Tiffin et al., 2022).

An example SJT item that was created as part of the present project, for use in a mental health services context, can be viewed in Table 6.1. In this example, having been shown a scenario (item stem), the respondent is asked '*How important is it to take into account the following consideration when deciding how to respond to the situation?*' Alternatively, the respondent could be asked how *appropriate* a given response is with regards to the item stem. The former 'importance' instructions require an individual to understand what factors are relevant, to roughly what degree, in making a potential behavioural response professional or unprofessional. In contrast, the latter 'appropriateness' items require individuals to understand to what extent potential behavioural responses are professional or unprofessional given the scenario portrayed. As suggested previously, it is possible that the different response formats could enable the evaluation of different traits (Tiffin and Carter, 2015).

The SJT initially aimed to assess the knowledge of all registered, clinical professionals working in mental health services, regardless of the profession that they aligned to; however, it has previously been suggested that SJT's should be developed to target specific staff groups (Patterson et al., 2016d; 2018). Thus, later analyses were performed to explore what professional groups the SJT scores were most valid for. In particular, there was an awareness that the respondents would not be expected to encounter every scenario included in the SJT, due to the differing nature of their roles.

To lower the cognitive load placed on test-takers, a decision was made to write scenarios in the first person. It was hoped that by writing scenarios in the first person, test-takers would be more likely to emotionally connect with the situation. It was also expected that such wording would 'draw' the respondent into the scenario and therefore their response would be more likely to reflect what they would actually do in practice. Nonetheless, the SJT was designed to be a knowledge-type test and individuals were thus asked to place themselves in the situation and comment on what 'they' should do, or how appropriate they thought a particular behaviour was.

Table 6.1 An example SJT item for use in mental health and learning disability services

SJT element	Example SJT item	Knowledge of professional attribute evaluated	Setting
Scenario / Item stem	SCENARIO: You have recently started working on an inpatient ward. A patient, who has a diagnosis of schizophrenia, informs you that a fellow staff member on the ward has made plans to meet up with a patient, Sheila, in a non-professional capacity, once they are discharged from hospital. You are aware that Sheila has a discharge meeting arranged for next week.	Patient Focus	Inpatient ward
Response instruction	How ***important*** is it to take into account the following consideration when deciding how to respond to the situation?		
Response option	CONSIDERATION: That there are previous concerns that the staff member has not always maintained appropriate boundaries with patients - Very important - Important - Of minor importance - Not important at all		

Stages of development

The development of the content of the SJT items took place in four stages: first, the Critical Incident Technique was used to generate scenarios for the test; second, items were written using an inductive and deductive approach; third, item content was assessed by a lead item reviewer (PT); and fourth, item content was reviewed by SMEs. During the latter step, SMEs were asked to score items also, in order to generate a provisional scoring key. The item development, review, SME scoring, and prioritisation process are described below; outcomes derived from this process are subsequently reported.

Item development

SJT items should be the product of a thorough job analysis to assess the key attributes and competencies required of a role (Patterson et al., 2016d; 2018). To generate a blueprint for the assessment, the author sought to determine how key stakeholders viewed the attribute of interest; namely, 'professionalism' in mental health services. A cross-sectional qualitative study was undertaken (see chapter 5), which incorporated interviews (n=13) and six focus groups with patients, carers and professionals working in mental health services. In addition to asking participants 'what professionalism means to them', the Critical Incident Technique (Flanagan, 1954) was utilised to generate item content for the SJT. Participants were asked to discuss '*an example of professional or unprofessional behaviour that they had recently observed, especially good or bad practice.*' In order to establish the series of events leading up to, and succeeding the incident, participants were asked to elaborate on 1) the circumstances leading up to the incident, 2) what the professional responding to the incident did, 3) what was professional or unprofessional about the professional's response, and 4) other possible desirable or undesirable behavioural responses that could have been experienced or observed. Once a participant described the event, other participants were invited to share their views or discuss any similar events that they had

encountered. The aim of such questioning was to elicit critical incidents that could be modified into item stems and response options. Participants were also asked whether they had encountered any professional dilemmas.

It is common for SJT items to be selected by job analysis, or the judgement of SMEs, however items can also be selected on the basis of theory (Weekley and Ployhart, 2006). During item writing, items were generated using inductive and deductive approaches. For example, using an inductive approach, critical incidents noted by the participants were first drafted into scenarios (i.e., item stems) by the lead researcher. Having reviewed these scenarios, responses were also developed using 1) the qualitative data, and 2) prior knowledge of professionalism in mental healthcare services, thus using a deductive approach also. Prior knowledge was gained whilst undertaking a systematic review on the topic (reported on in chapter 4), as well as the aforementioned qualitative study (see chapter 5).

Item review

Scenarios were developed by a lead item writer (LA) in the first instance; the lead item reviewer (PT) subsequently evaluated scenarios with regards to their *clarity* and *pertinence*. During review, the lead item reviewer revised items where they felt appropriate. The lead item reviewer was a Reader in 'psychometric epidemiology' and an honorary consultant psychiatrist at the time of the study. Following review of the scenarios, response options were developed by the lead researcher; the lead reviewer evaluated response options too, editing these where they deemed appropriate given their clinical and academic expertise. At this time, the lead reviewer also provisionally scored each item according to how 'appropriate' or 'important' they believed the response option was. As stated earlier, a 4-point Likert scale format was used for this purpose. The lead reviewer added additional behavioural responses to be depicted, where they thought these were suitable for inclusion in the assessment.

During development and review, scenarios were mapped to one of six professional attribute content domains that were identified via a thematic analysis during the qualitative study. These attributes included: '*Commitment to Professionalism,*' '*Coping with Pressure,*' '*Effective Communication,*' '*Patient Focus,*' '*Working Effectively as Part of a Team,*' and '*Working with Carers.*' Scenarios were not mapped to the professional attributes '*Organisation and Planning*' and '*Problem Solving*', because these attributes are, as reported previously, implicit to the SJT process (Medical Schools Council, 2011). Scenarios were also not mapped to '*Learning and Professional Development*' and '*Self-Awareness and Insight*'; the former attribute is an element of one's commitment to professionalism, whereas the latter attribute facilitates one's ability to cope with pressure.* Having utilised the research findings to map the desired personal qualities onto the created test content, the test items could then be distributed in a balanced way ensuring that all domains were covered equally among the SJT test forms. To enable a fair representation of settings for the depiction of situations in the assessment, scenarios were also mapped to inpatient and community settings, as well as miscellaneous and off-duty scenarios.

Once all scenarios and responses had been mapped to their related professional attribute and setting, it was necessary to reduce the number of items to facilitate their review and provisional scoring by SMEs; there were initially too many items to be practical to send to the SME panel. Draft items were first prioritised on the basis that a scenario had at least four *importance* and four *appropriateness* response options each, because this would potentially allow for the evaluation of 'method effects' for the two response types. That is, it was hoped to be able to evaluate any potential influence of the two instruction formats (*importance* vs *appropriateness*) on the observed scoring patterns. Items were next prioritised on their perceived difficulty and complexity. Following prioritisation, items were sent to the SME panel for review and provisional scoring.

Devising a draft scoring key

SME panels are often used to develop scoring keys for SJTs (Bergman et al., 2006). For the current study, the SME panel incorporated some individuals that had participated in the initial interviews and focus groups who agreed at the time of their interview or focus group that they would be happy to partake in a follow up study. It was hoped that SME panels of between 12 and 15 members would be composed, with representation, where possible, from a range of the professions. Attention was paid to ensure that there was fair representation across each of the professional groups in the SME panel, and that patients and carers were also included.

Items that were chosen to be reviewed and provisionally scored by the SME panel were uploaded into a survey using Qualtrics; this procedure allowed participants in the SME group to rate items electronically. Due to the number of items that required review, the items were split into two separate Qualtrics forms with 99 items each. These forms were then distributed via a survey link to one of two SME panels. The SME panels were asked to rate each scenario for 'relevance' and 'clarity', and response options for plausibility. They were also asked to score items according to how 'important' or 'appropriate' they thought the response options were (see Appendix K for a sample of the survey sent to each SME panel; the sample includes some instructions as well as practice items requiring review). This development stage resulted in a provisional scoring key for the draft SJT. Participants in the SME group were offered a £30 gift voucher for taking part at this stage. Whilst patients and carers are not typically included in an SME panel, it was felt important to receive their contributions during this stage; to facilitate their involvement, the researcher met with patients and carers individually allowing them to complete the survey on paper.

Patient and carer responses were inputted digitally, by the author, alongside the other participants' responses.

Item Prioritisation

In order to prioritise items for the pilot SJT, items demonstrating poor 1) expert consensus, 2) relevance, or 3) clarity, were discarded or reworded accordingly. SME responses were evaluated for consensus by calculating: dichotomous consensus, which determined the percentage of raters selecting the same side of the rating scale (i.e. appropriate versus inappropriate or important versus unimportant); polytomous consensus, which was calculated by determining the percentage of raters that chose the most common response; and corrected Krippendorff's alpha (Krippendorff, 1970) to determine the level of agreement between raters selecting the same option on the rating scale. The latter method is a more stringent approach, as it adjusts for agreement by chance by producing a coincidence matrix.

Results from the SJT development process

Item Development

Utilising the focus group and interview data, the lead researcher developed SJT scenarios (item stems) and response options. Of these, 98 scenarios and 663 response options were selected and sent to the lead item reviewer. As noted previously, scenarios had been mapped to one of six professional attributes, including *Commitment to Professionalism* (n=34), *Coping with Pressure* (n=15), *Effective Communication* (n=12), *Patient Focus* (n=22), *Working Effectively as Part of a Team* (n=11), and *Working with Carers* (n=4; see Table 6.2).

Item Review

The lead item reviewer examined scenarios (n=98) for both clarity and pertinence, revising these where they deemed appropriate. The lead item reviewer provisionally

scored response options also to help develop a scoring key. Given the high number of

scenarios and response options, items had to be shortlisted prior to review by the SME

panel (see Figure 6.1). Shortlisting resulted in 24 scenarios and 198 final items. An

overview of the scenarios that were prioritised for review can be seen in Table 6.3.

Table 6.2 Breakdown of scenarios by professional attribute and setting, prior to prioritisation for the concordance panel

	Community	Inpatient	Off Duty	Miscellaneous	Overall
Commitment to Professionalism	9	8	7	10	**34**
Coping with Pressure	9	6	0	0	**15**
Effective Communication	6	4	2	0	**12**
Patient Focus	11	7	2	2	**22**
Working Effectively as Part of a Team	6	4	0	0	**11**
Working with Carers	3	1	0	0	**4**

Table 6.3 Scenarios prioritised for SME review according to professional attribute and setting

	Community	Inpatient	Off Duty	Miscellaneous	Overall
Commitment to Professionalism	3	2	4	1	10
Coping with Pressure	3	0	0	0	3
Effective Communication	1	1	0	0	2
Patient Focus	1	3	0	0	4
Working Effectively as Part of a Team	2	2	0	0	4
Working with Carers	1	0	0	0	1
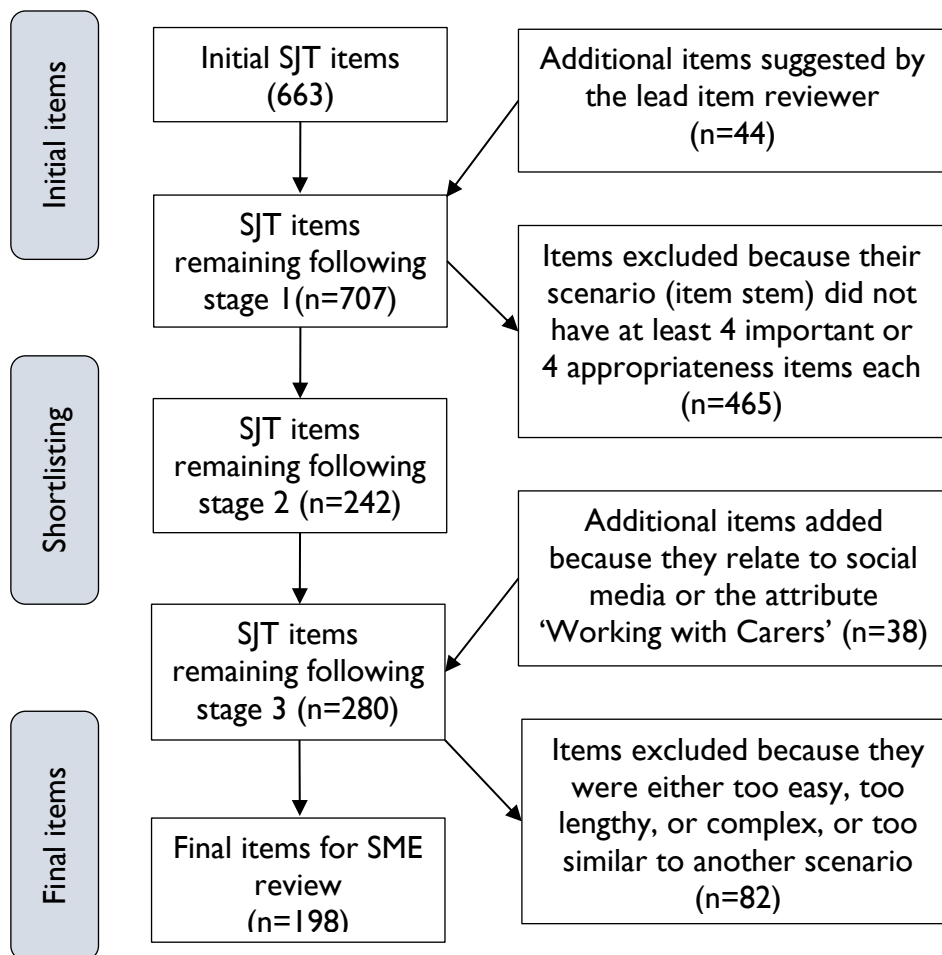
Figure 6.1 Shortlisting of items for review by the SME panel

Devising a draft scoring key

Shortlisted items were distributed among two separate Qualtrics surveys, which incorporated 99 items each. During distribution, attention was paid to ensuring that the alternate forms contained a similar number of items in each setting, and their related professional attributes, thus creating balanced blueprints in terms of content. Individuals that consented to being part of the follow up were split into one of two SME panels and received one version of the survey each. Each SME panel included staff from a range of disciplines, as well as patients and carers (see Table 6.4). Due to some individuals not completing the survey, there was no representation from psychologists in the first SME panel.

Table 6.4 Concordance panel participants

Concordance Panel 1 (gender)	Concordance Panel 2 (gender)
Lauren Aylott (f, patient)	Lauren Aylott (f, patient)
Paul Tiffin (m; psychiatrist)	Paul Tiffin (m; psychiatrist)
Patient (f)	2x Patient (m & m)
2x Carers (f & f)	Carer (f)
2x Psychiatrist (f & m)	2x Psychiatrist (f & m)
3x Occupational Therapist (f & f & m)	2x Occupational Therapist (f & f)
2x Nurse (f & m)	2x Nurse (f & m)
	2x Psychologist (f & f)

Item prioritisation

Following review by the SME panel, it was hoped that from the 198 items, 90-100 of the most 'reliable' items would be shortlisted for piloting across one of two forms; the reliability of items was predominantly determined by their level of agreement among the SME panel. Items were shortlisted as follows (see Figure 6.2). First, items were omitted if two or more individuals commented that they were either not plausible, ambiguous, or not realistic. In total, 31 items were omitted for one of these reasons, resulting in 167 items for further consideration. Next, items were omitted based on additional feedback provided by the participants; for example, in one instance, a participant commented that the scenario was job dependent, and did not apply to all workers/professions. The omission of these additional items (*n*=20) resulted in 147 remaining items for shortlisting. A decision was made to next reduce the number of items based on their level of agreement among the SME panel. Prior to doing so however, 24 items were retained to facilitate the analysis of method effects. The retained items met the following criteria: dichotomous consensus = >.50; polytomous

consensus = >.38. Having kept the latter 24 items, and previously discarded 51, 123 remaining items required shortlisting for the pilot SJT. Of the remaining 123 items, 22 were omitted on the basis that they did not have enough agreement among the SME panel. In order to remove the desired number of items from the item pool, the following criteria were initially applied to indicate 'not enough agreement' across the SME panels: dichotomous scoring < .65; polytomous scoring < .45; and Krippendorff's alpha < .31. To ensure that we had a suitable number of items for pilot testing (i.e., not too many), the Krippendorff's alpha cut-off was lowered and a further 18 items were removed on the basis that their Krippendorff's alpha was less than 0.25, resulting in 83 items for shortlisting. The Krippendorff's alpha value was prioritised over other measures of agreement at this time, because of it's use of a 'coincidence' matrix to correct for the play of chance.

Next, five items were omitted on the basis that they had perfect agreement (Dichotomous scoring = 1; Polytomous scoring = 1; Krippendorff's Alpha = 1), because these items were deemed to be too easy to discriminate between test-takers. Seven items were omitted on the (practical) basis that there was only one *appropriateness*, or one *importance* item for the item stem (scenario), because this would have made modelling method effects unfeasible (see earlier). One further item was omitted, because it was very similar to another item, resulting in 70 remaining items for shortlisting. A final four items were omitted on the basis that one individual had marked the scenario as either unclear or ambiguous. The remaining 66 items were selected for the pilot SJT, alongside the 24 items retained previously, resulting in a final 90 items for piloting.

SME panel review

SJT items reviewed by the SME panel
Panel 1 (99)
Panel 2 (99)

→ Items excluded (n=31*)
Item response was not plausible (n=9)
Scenario was not realistic or relevant (n=14)
Scenario was unclear or ambiguous (n=9)

↓

Items remaining after stage 1 (n=167)

→ Items excluded due to additional comments (n=20)

↓

Screening

Items remaining after stage 2 (n=147)

→ Items retained for exploration of method effects (n=24)

↓

Items requiring further consideration (n=123)

→ Items excluded (n=57)
Not enough agreement (n=40)
Perfect agreement (n=5)
Only one of either instruction type per scenario (n=7)
Similar item (n=1)
Scenario is unclear or ambiguous according to one SME member (n=4)

↓

Items retained following further consideration (n=66)

↓

Included

Items included in pilot study (n=90)

* 1 item was marked as both 'not plausible' and 'not realistic' during stage 1; hence, the total of these values is 32, yet only 31 items were omitted

Figure 6.2 Item prioritisation flow diagram

Final Pilot Assessment

The 90 prioritised items were distributed across two forms in order to pilot as many items as possible within a limited testing window. A decision was made to share 10 of the 90 items across both surveys to facilitate test equating of the two forms of the SJT.

The shared items were also anticipated to be useful as a check that the two groups of respondents, randomised to each form of the test, were unlikely to differ in their overall ability to answer the SJT. To create equivalence in content across the alternate forms, a blueprinting exercise was conducted (Coderre et al., 2009; Lievens and Sackett, 2007; Whetzel et al., 2020); efforts were made to distribute the remaining items evenly across the forms, according to their related professional attribute, setting and response instruction. The final two forms had 50 SJT items each and the composition of these surveys can be viewed Table 6.5.

Table 6.5 Breakdown of attributes by survey to be piloted

Field	Version 1	Version 2	Shared items	Overall
Commitment to Professionalism	13	7	10	30
Coping with Pressure	5	6	0	11
Effective Communication	5	3	0	8
Patient Focus	4	2	0	6
Teamwork	13	16	0	29
Working with Carers	0	6	0	6
Community	18	25	0	43
Inpatient	20	10	6	36
Off Duty	2	5	4	11
Social Media	3	3	4	10
Scenarios	9	9	2	20
Important	23	20	6	49
Appropriate	17	20	4	41

6.3.3.3 *Development of a workplace behaviours rating tool (WBRT)*

Previous studies evaluating the validity of SJT scores for personnel selection have commonly used supervisor ratings as a means to estimate 'on the job' performance. Many studies observe a statistically significant and substantively meaningful relationship between such SJT scores and supervisor ratings (e.g. Crook et al., 2011).

These 'validity coefficients' for SJT scores are often close to, or comparable in magnitude to those for face-to-face interview processes (McDaniel et al., 1994; 2001).

For the outcome measure, a bespoke workplace behaviours rating tool (WBRT) was developed to assess participants' perceived *professionalism* and *effectiveness* in their job role (Appendix L); the rating tool was developed to take approximately 5 to 10 minutes to complete. Alongside some initial questions regarding the colleague's gender, the relationship they held with the participant, and the length of time they had known the responding participant, the rating tool asked colleagues to rate the participants, on a scale out of 100, on both their *professionalism* and *effectiveness* using the relative percentile method (see Figure 6.3; Goffin et al., 1996; 2009). There was an interest in obtaining absolute ratings of *professionalism* and *effectiveness*, regardless of the individual's grade or level of experience; thus, colleagues were asked, specifically, '*In comparison to other staff members of the same profession, irrespective of their grade and experience, please estimate the level of professionalism of this individual*', and '*In comparison to other staff members of the same profession, irrespective of their grade and experience, please estimate how effective in their role, this staff member is*'.

Whilst providing their rating, colleagues were asked to provide a brief example of a behaviour that characterised the individual in the related domain. This approach is an adaption of other methods that have been used previously to help obtain more accurate estimates of performance (for example, see Borman et al., 2001; Smith and Kendall, 1963). Finally, colleagues were asked to score the participant on the BFI-2-S, providing peer as opposed to self-ratings (see below; Soto and John, 2017).

Please will you now think very carefully about where this staff member sits on the following scales?

In comparison to other staff members of the same profession, <u>irrespective of their grade and experience</u>, please estimate the level of professionalism of this individual.

Please use the following definition to base your judgement; please indicate your decision, by moving the below marker to the desired percentage.

Professionalism allows practitioners to make appropriate judgements in times of need, applying critical thinking, reflection and situational judgement.

	Below Average	Average	Above Average
	0 10 20 30 40 50 60 70 80 90 100		
Professionalism of this individual. ()	▬▬▬▮▬▬▬		
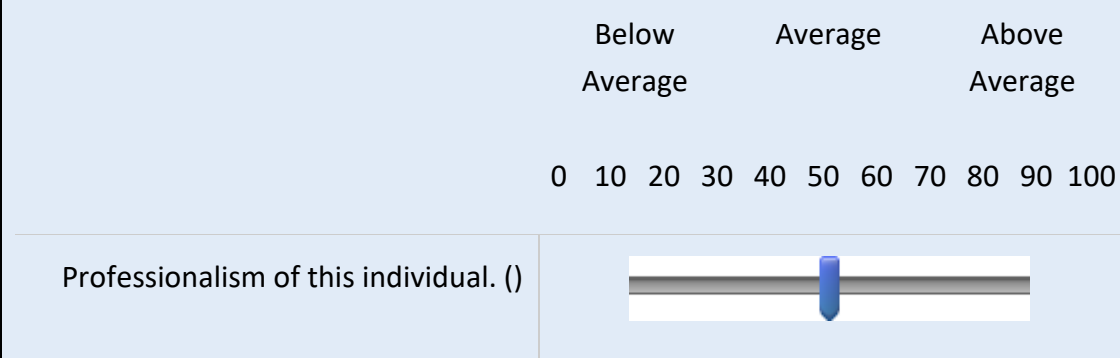
Figure 6.3 WBRT professionalism question using the peer percentile method (adapted from Goffin et al., 1996, Goffin et al., 2009)

6.3.3.4 *The Big Five Inventory–2 Short Form (BFI-2-S)*

The BFI-2-S was used to assess the 'big five' personality traits (Extraversion, Agreeableness, Conscientiousness, Negative Emotionality, and Open-Mindedness; Soto and John, 2017). The BFI-2-S is an abbreviated version of the BFI-2; the BFI-2 has 60 items, whereas the BFI-2-S has 30 items and takes between 3 and 5 minutes to complete. The BFI-2-S was developed specifically for research studies that posed significant constraints on the amount of time that a participant has to complete an assessment; the predictive validity of the BFI-2-S is somewhat less that the BFI-2, yet, the shortened form has still retained 93% (mean R^2 = 0.236) of the original BFI-2's predictive power (Soto and John, 2017). The BFI-2-S was deemed a more suitable assessment for the study, as participants already had many SJT items to complete and the research team wanted to minimise the burden on participants, and, therefore, any

subsequent drop out from the study. The BFI-2-S requires participants to rate how well certain statements describe them (e.g., 'Is outgoing, sociable', or 'Worries a lot') on a 5-point Likert scale where 1 = Disagree strongly and 5 = Agree strongly. The BFI-2-S has been developed for both self and peer ratings. For the current study, participants were asked to provide self-ratings on the BFI-2-S, whereas colleagues were asked to provide peer-ratings regarding the participant. The rationale for obtaining peer-ratings too, was that individuals are known to complete personality assessments in a socially desirable manner and the results are therefore, not always, a true reflection of an individual's behavioural disposition.

6.3.4 Recruitment and sampling for the pilot study

6.3.4.1 *Initial participants*

It is recommended that SJT pilot studies use participants that are similar in characteristics to the population intended to be selected using the final operational SJT; examples provided include a similar level of job experience, ethnicity and educational level (Weekley and Ployhart, 2006). All professionals working for TEWVs NHSFT were invited to participate in the study, so long as they were registered with a professional clinical body (e.g., the Nursing and Midwifery Council).

The study was initially advertised in the Trust e-bulletin, which is distributed, via email, to all staff members working in TEWV NHSFT. Alongside a brief summary of the study, the advert had a link to the Trust's intranet site where a flyer was provided. Contact details for the lead researcher were provided in both the e-bulletin advert and on the Trust intranet. Participation was voluntary, though staff were offered a £10 gift voucher to reimburse them for their time completing the study survey. Individuals were advised that they would be sent a copy of their personality assessment results if they took part, and that they would be sent a certificate of participation on request. Additional recruitment involved asking the professional leads to cascade an email to staff within their discipline regarding the study. Existing participants were also asked to discuss the study with their colleagues who may also want to take part.

Individuals that expressed an interest in the study were sent a participant information sheet (PIS; Appendix M) and privacy notice detailing the study (Appendix N). The PIS stated clearly that participants were required to leave contact details for their manager or supervisor so that the research team could obtain feedback regarding the participant's job performance. Initial participants had the opportunity to leave contact details for up to three of their peers also, although this was not mandatory. A link to the survey was sent the next working day. The online survey included a summary of the participant information sheet and several consent statements; participants were advised that by completing the survey they had read all the information provided and agreed to take part. A further statement included 'if you do not wish to take part in the research study, please end the survey now.' Therefore, having had the objective of the study clearly explained, informed consent was implied by completion of the survey.

6.3.4.2 *Colleague and manager/supervisor ratings*

All colleagues, including line-managers and supervisors, where contact details were provided, were contacted by email for workplace related feedback. The email incorporated a link for the privacy notice, as well as a PIS, developed for colleagues specifically. A link to an online survey allowing them to provide feedback was also included. Following liaison with the NHS site regarding the study, it was clearly stipulated on the study flyer that participants should inform their supervisor and colleagues that they would be approached for these ratings before providing their names for feedback. Colleagues were offered a £5 gift voucher to reimburse them for their time and both participants and colleagues were made aware of this incentive within their respective PIS. Colleagues were advised that their taking part, and the ratings provided, would be kept absolutely confidential and would be used for research purposes only. As with the initial participants, informed consent was implied by completion of the survey.

6.3.5 Data collection

Data were collected between January and October 2020. Surveys were distributed to both participants and colleagues using the survey platform Qualtrics. The survey sent to initial participants was designed to take approximately 30 minutes to complete. At the end of the survey, participants were asked to leave contact details for their manager or supervisor and up to three colleagues; this requirement was explicitly stated in the PIS. As the survey was delivered online, an internet connection was required. Having recognised that clinicians may not be able to complete the survey in one go, the survey was designed so that it would save automatically and could be completed in more than one sitting. Colleague ratings were requested one week following completion of the SJT. Up to two reminders were sent to participants, at two-weekly intervals, had they not completed the SJT, or provided feedback, within that timeframe.

6.3.5.1 *Demographics and professional characteristics*

Various demographic details were captured at the start of the survey so that subsequent analyses could be performed and differential scoring (i.e., potential bias), by certain groups on the SJT could be explored; this included age, gender, sexual orientation, whether individuals considered themselves to have a disability, ethnicity, religion, and whether English was their first language. Additional information requested at the start of the survey included an individual's profession, the years they had been registered to the profession, their grade, and both the setting and specialty they worked.

6.3.5.2 *Situational judgement test*

Initial participants were next required to complete 50 SJT items. As detailed earlier, there were two versions (forms) of the SJT with each form having 10 overlapping SJT items. The remaining 40 items differed according to the form presented to the participant. Each form was randomly distributed to participants using the RAND function on Excel. In this instance, RAND computed the number 1 or 2 for each

participant who was then assigned that version of the SJT. All SJT items were mandatory; this was incorporated into the survey design so that participants could not accidentally skip questions. Once participants had completed all SJT items, a few additional questions, with a free-text response format, were posed regarding the participant's experience of the SJT. This was to obtain information on the perceived acceptability of the SJT from the test-taker's perspective. Individuals were first asked how long it had taken them to sit the test; participants were then asked whether the test was relevant to their role, appropriate for their grade, suitable for the recruitment of staff into mental health services, and fair to all applicants regardless of their profession, gender, race, and other characteristics. Participants were asked also to provide a rationale if they answered 'no' to any of these questions. Finally, participants were asked if they had any further comments to make regarding the test.

6.3.5.3 *Personality measures*

Having completed all the SJT items, participants were asked to complete the BFI-2-S. Participants were informed that this would take between three and five minutes and that they would be sent a copy of their results once everyone had completed the survey.

6.3.5.4 *Job performance*

Colleagues, including managers and supervisors, were asked to complete the bespoke WBRT as a means of providing their feedback regarding the initial participant's *professionalism* and *effectiveness*. Following this, colleagues were asked to rate the initial respondents on the BFI-2-S, providing peer as opposed to self-ratings. All phases in the recruitment and data collection process are displayed in Figure 6.4.

Recruitment: Step 1
- advert placed
- potential participant contacts researcher
- participant information sheet and privacy notice distributed
- email sent with a link to the online survey

Part 1: SJT
- participants complete the survey, providing:
  - their demographic, and professional details,
  - responses and feedback regarding the SJT,
  - self-ratings on the BFI-2-S

Recruitment: Step 2
- email colleagues with link to workplace behaviours rating tool

Part 2: job feedback
- colleagues provide feedback via an online survey, regarding:
  - the participant's professionalism and effectiveness
  - the participant's personality attributes, via the BFI-2-S

Figure 6.4 Recruitment and data collection process

6.3.6 Data analysis

The primary aim of the data analysis was to assess the psychometric properties of the SJT, as well as its feasibility, in order to inform the scoring method, items and target population for the final test. Three objectives were set, in order to fulfil this aim. First, the psychometric properties, and especially the criterion-related validity of scores resulting from the SJT were evaluated according to a variety of plausible, and commonly employed scoring systems. The extent that different scoring approaches influenced the convergent and criterion-related validity of the resultant test was explored according to the pre-defined outcomes of interest, namely ratings of perceived *professionalism* and *effectiveness* and scores on the personality assessment. Second, the relationships between scores on the SJT and the professional and demographic characteristics of the participants were evaluated. Where appropriate, the relative scores achieved by different groups according to the main alternative

196

scoring systems were assessed for the following reasons: 1) to identify the professional group/s where the final test system is likely to be most valid, and 2) to identify the potential of the test to have an adverse impact on any particular demographic or professional groups. Third, the acceptability and feasibility of the SJT was evaluated. Once the final SJT items had been determined, the internal consistency of these items was examined. Factor analysis was also performed to assess the dimensionality of the responses to the final SJT items.

6.3.6.1 *Cleaning and management of data*

Data were anonymised and imported from Excel into Stata where the main data analysis was performed. The data were first checked for out-of-range values and values were recoded as appropriate. For example, in two instances, nurses had selected 'other' as their profession; both individuals were recoded as nurses to conduct subsequent analyses. Next, string variables, such as age, were turned into numerically encoded variables. All questions on the survey were mandatory and there were therefore no missing data. There were however two versions of the form. Aside from the overlapping items, participants that sat form 1 did not sit the remaining items on form 2, and vice versa. Thus, data on these mutually exclusive items were 'missing by design.' In some instances, participants selected the option 'prefer not to say' with regards to their demographic characteristics. These individuals were therefore excluded from the group analyses where their data were missing.

6.3.6.2 *Descriptive statistics*

Descriptive statistics were generated to explore the demographic and professional characteristics of the sample. These data were compared to the overall staff clinical workforce to determine to what extent the sample was representative of the overall workforce. Descriptive statistics were also generated for the duration of completion, the proportion of the sample that believed the test was relevant to their role, appropriate to their grade, suitable for the selection of staff into mental health services, and fair to all applicants. Finally, the distribution of outcome variables,

including scores on the SJT, the BFI-2-S, and ratings of *professionalism* and *effectiveness* were described.

6.3.6.3 *Scoring*

The distribution of scores on the SJT were explored using various scoring approaches, including: raw SME polytomous scoring, raw SME binary scoring, dichotomous modal consensus scoring, and proportion modal consensus scoring (see Table 6.6 for a description). It was recognised that raw SME polytomous scoring (0,1,2,3) would more likely be influenced by extreme response style. Therefore, the research team explored binary and dichotomous scores on the SJT, as well as polytomous scores. An alternative form of scoring, McDaniel shape scoring, has been found to reduce elevation and scatter and improve item validity for some SJTs where a relatively large number of Likert scale response points are used (e.g. 7 points; McDaniel et al., 2011). As there were only four options (Likert points) to choose from on the current SJT, McDaniel's method was not considered as useful as it would have been, had there been more points on the Likert scale. Therefore, the McDaniel method was not utilised during the study. The distribution of variables was assessed using histograms and Q-Q plots. Following this, judgement was used regarding what test was most appropriate for each analysis.

Using the mean and standard deviation obtained by the sample for each scoring approach, SJT z scores (mean 0, sd 1) were calculated for each form separately; Z scores were calculated by using the raw score minus the mean for that subgroup, divided by the standard deviation for the same population. Z scores were calculated as a means of approximately equating the SJT scores so that the forms could be analysed alongside one another; this amalgamation increased the available sample size for the analyses. The assumption was that this standardisation at least crudely equated the scores from the two forms of the pilot SJT. The success of this equating was evaluated as part of the analyses (see later).

Table 6.6 Scoring approaches utilised in the current study

Scoring method	Description
Raw SME polytomous scoring	A test-taker is allocated a score between 0 and 3 for each item, dependent on how highly that item was endorsed by the SME panel. A score of three indicates that the item was most preferred by the SME group.
Raw SME binary scoring	A test-taker is allocated a score of one, if their response is the first or second most commonly endorsed response according to the SME panel. Otherwise, the test-taker receives a score of zero.
Dichotomous modal consensus scoring	A test-taker is allocated a score of one for an item, if their response is the most common response provided by other test-takers in the pilot study; otherwise, they receive a score of zero.
Proportion modal consensus scoring	A test-taker is allocated a score for each SJT item dependent on the proportion of other test-takers in the pilot study that endorsed that response.

6.3.6.4 *Convergent and divergent validity*

The convergent and divergent validity of the SJT scores was assessed for each of the scoring approaches regarding each of the Big Five personality traits, as assessed by the BFI-2-S.

6.3.6.5 *Criterion-related validity*

The criterion-related validity of the SJT was explored by evaluating the relationship between SJT scores and ratings of perceived *professionalism* and *effectiveness* provided by colleagues, including supervisors and managers. It was expected that there would be a positive relationship between SJT scores and ratings of both *professionalism* and *effectiveness*; this was assumed because those that know about professional and unprofessional behavioural responses to challenging interpersonal situations are more likely to perform effectively in the actual workplace. Based on the findings of previous research, differing scoring approaches were expected to be associated with varying magnitudes of validity coefficients observed. The criterion-related validity of the SJT scores was therefore assessed for each of the scoring approaches.

6.3.6.6 *Adverse impact / group differences*

An earlier systematic review reported that there are somewhat differing expectations for each of the professions (see chapter 4, section 4.10.1.4); for example, a key value in counselling psychology is the focus the therapist places on people's strengths and assets (Goodyear et al., 2016). Thus, the SJT may have been more suited to some professions than others. Consequently, the predictive validity analyses explored whether the SJT scores demonstrated more validity for certain professional groups. The potential for adverse impact was also explored by examining the relationship between SJT scores and gender; due to the limited number of participants identifying as being from Black, Asian, and Minority Ethnic (BAME) groups, statistical analyses were not conducted to explore group differences with regards to ethnicity. Once the final SJT items were determined, further analyses explored whether there were differences in SJT scores according to age, sexual orientation, whether individuals considered themselves to have a disability, religion, the years they had been registered to the profession, and their grade (according to Agenda for Change; AFC).

6.3.6.7 *Acceptability*

The face validity and acceptability of the SJT was also evaluated and is reported on for the recommended target population.

6.3.6.8 *Determining the final items*

Odds ratios were calculated to determine which items were most promising for the final test (see section 6.4.9). Once the final items were chosen, a simulation was conducted to establish the criterion-related validity of these items.

6.3.6.9 *Incremental validity*

Next, the incremental validity of the SJT scores to predict ratings of perceived *professionalism* and *effectiveness*, over and above the Big Five personality traits, as assessed by the BFI-2-S, was explored using the most promising scoring approach.

6.3.6.10   *Dimensionality*

Once the final items, most promising scoring system and target professional group for the SJT had been determined, the dimensionality of scores was evaluated using binary factor analyses, as implemented in the software FACTOR 14.1 (Lorenzo-Seva and Ferrando, 2006; 2013) and Mplus v 8.7.

6.3.6.11   *Method effects and reliability*

Additional analyses explored whether the scores obtained for appropriateness items were related to the scores obtained for importance items, and whether either instruction format evidenced increased criterion related validity for predicting job performance.

SJT items are heterogeneous in nature; thus, assessing the reliability of SJT scores using Cronbach's alpha is problematic (Catano et al., 2012). Nevertheless, many authors continue to document Cronbach's alpha's when reporting on SJT studies. Henceforth, the reliability of both versions of the SJT was calculated once the final items, most promising scoring system, and target population had been determined. The Kuder-Richardson Formula 20 (KR-20) was used to calculate the reliability coefficients for item level data; KR-20 is a reliability index for binary data and was therefore most suited to the final scoring approach (Kuder and Richardson, 1937). The reliability of the ordinal scenario summed scores were also evaluated.

6.4   Results

6.4.1   Participants

In total, 170 Participants completed the SJT; 92 completed form 1 (response rate: 73.6%) and 78 completed form 2 (response rate: 72.2%; see Figure 6.5).

Participants expressed an interest in the research study and were subsequently sent a PIS (*n*=233)	
125 participants were sent a link to form 1	108 participants were sent a link to form 2

The SJT was completed (*n*=170)	
92 participants completed form 1 (73.6%)	78 participants completed form 2 (72.2%)

Workplace related feedback was requested for all respondents (*n*=170)	
524 feedback requests were sent to colleagues, managers and/or supervisors; of these, 405 separate individuals were contacted (some individuals were sent multiple requests)	

Workplace related feedback was returned (*n*=152)	
331 requests were returned; 3 responses were completed for the wrong individual, and thus deleted; the 328 final responses pertained to 152 participants (form 1, *n*=78; form 2, *n*=74)	

Figure 6.5 Outcomes of recruitment and data collection

6.4.1.1 *Demographics*

The demographic characteristics of individuals that were given form 1 and form 2 were similar (see Table 6.7). The participants' mean age was 40.6 years, 136 participants were female (80%) and 159 participants had English as their primary language (93.5%). This latter figure is identical to the number of individuals that were of white ethnicity. The gender split in the study sample matched the gender difference of professionals employed by TEWV (81% of study participants were female; 81% of TEWV professionals were female as of March 2021). Also, 8% of TEWV's professional staff reported being disabled, which is similar to that reported in the current study sample. Finally, 94% of TEWV's staff identified themselves as White ethnicity, which is similar to that reported in the current study.

6.4.1.2 *Professional characteristics*

Information regarding participants' professional discipline, the number of years they had been registered with their profession, and their agenda for change band (grade) can be viewed in Table 6.8. Data regarding the setting and the specialty participants

202

worked are provided in Table 6.9. In some instances, participants worked in more than

one setting and/or specialty.

Table 6.7 Demographics of participants

Demographics	Overall (n=170) N (%)	Form 1 (n=92) N (%)	Form 2 (n=78) N (%)
Age, years (mean, SD)	40.6 (SD - 9.89)	40.8 (SD – 9.94)	40.5 (SD – 9.89)
Gender			
- Female	136 (80.0)	76 (82.6)	60 (76.9)
- Male	34 (20.0)	16 (17.4)	18 (23.1)
Ethnicity			
- Asian / Asian British	4 (2.4)	3 (3.3)	1 (1.3)
- Black / African / Caribbean / Black British	2 (1.2)	1 (1.1)	1 (1.3)
- Mixed / Multiple ethnic groups	2 (1.2)	0 (0.0)	2 (2.6)
- White	159 (93.5)	86 (93.5)	73 (93.6)
- Other	3 (1.8)	2 (2.2)	1 (1.3)
Sexuality			
- Bisexual	9 (5.3)	4 (4.4)	5 (6.4)
- Gay or Lesbian	3 (1.8)	1 (1.1)	2 (2.6)
- Heterosexual or Straight	153 (90)	85 (92.4)	68 (87.2)
- Other	2 (1.2)	0 (0.0)	2 (2.6)
- Prefer not to say	3 (1.8)	2 (2.2)	1 (1.3)
Religion			
- Christian	72 (42.4)	41 (44.6)	31 (39.7)
- Hindu	1 (0.6)	0 (0.0)	1 (1.3)
- Muslim	3 (1.8)	3 (3.3)	0 (0.0)
- No religion	86 (50.6)	42 (45.7)	44 (56.4)
- Sikh	1 (0.6)	1 (1.1)	0 (0.0)
- Other	5 (2.9)	4 (4.4)	1 (1.3)
- Prefer not to say	2 (1.2)	1 (1.1)	1 (1.3)
Disability			
- No	154 (90.6)	81 (88.0)	73 (93.6)
- Yes	13 (7.7)	9 (9.8)	4 (5.1)
- Prefer not to say	3 (1.8)	2 (2.2)	1 (1.3)
Primary language is English			
- No	11 (6.5)	7 (7.6)	4 (5.1)
- Yes	159 (93.5)	85 (92.4)	74 (94.9)

Table 6.8 Professional characteristics of participants

	Overall (n=170) N (%)	Form 1 (n=92) N (%)	Form 2 (n=78) N (%)
Professional discipline			
- Allied health professional	36 (21.2)	22 (23.9)	14 (18.0)
- Nurse	73 (42.9)	37 (40.2)	36 (46.2)
- Pharmacist	7 (4.1)	4 (4.4)	3 (3.9)
- Psychiatrist	15 (8.8)	6 (6.5)	9 (11.5)
- Psychologist	22 (12.9)	13 (14.1)	9 (11.5)
- Social Worker	9 (5.3)	5 (5.4)	4 (5.1)
- Other	8 (4.7)	5 (5.4)	3 (3.9)
Years registered			
- Less than 1 year	4 (2.4)	2 (2.2)	2 (2.6)
- 1 – 5 years	36 (21.2)	22 (23.9)	14 (18.0)
- 6 – 10 years	38 (22.4)	22 (23.9)	16 (20.5)
- 11 – 15 years	31 (18.2)	22 (23.9)	9 (11.5)
- 16 – 20 years	24 (14.1)	6 (6.5)	18 (23.1)
- 21 – 25 years	13 (7.7)	5 (5.4)	8 (10.3)
- 26 – 30 years	10 (5.9)	7 (7.6)	3 (3.9)
- 31 – 35 years	9 (5.3)	4 (4.4)	5 (6.4)
- 36 – 40 years	2 (1.2)	2 (2.2)	0 (0.0)
- 41 – 45 years	1 (0.6)	0 (0.0)	1 (1.3)
- Not documented	2 (1.2)	0 (0.0)	2 (2.6)
AFC Band			
- 4	1 (0.6)	1 (1.1)	0 (0.0)
- 5	19 (11.2)	9 (9.8)	10 (12.8)
- 6	62 (36.5)	35 (38.0)	27 (34.6)
- 7	32 (18.9)	19 (20.7)	13 (16.7)
- 8	37 (21.8)	20 (21.7)	17 (21.8)
- Other	19 (11.2)	8 (8.7)	11 (14.1)

Table 6.9 Setting and specialty worked

	Overall (n=170) N (%)	Form 1 (n=92) N (%)	Form 2 (n=78) N (%)
Setting			
- Community services	134 (78.8)	70 (76.1)	64 (82.1)
- Inpatient services	44 (25.9)	30 (32.6)	14 (18.0)
- Corporate services	13 (7.7)	6 (6.5)	7 (9.0)
Specialty			
- All services, Trustwide	18 (10.6)	9 (9.8)	9 (11.5)
- Adult Mental Health services	64 (37.7)	39 (42.4)	25 (32.1)
- Child and Adolescent services	37 (21.8)	17 (18.5)	20 (25.6)
- Forensic Learning Disability services	8 (4.7)	6 (6.5)	2 (2.6)
- Forensic Mental Health services	20 (11.8)	11 (12.0)	9 (11.5)
- Learning Disability services (adults)	13 (7.7)	9 (9.8)	4 (5.1)
- Learning Disability services (children)	5 (2.9)	3 (3.3)	2 (2.6)
- Mental Health Services for Older People	36 (21.2)	20 (21.8)	16 (20.5)

NB: some individuals worked in more than one specialty and/or setting, thus, the figures noted above tally to more than the number of participants

6.4.2 Duration of completion

Individuals reported that the SJT took them between 5 minutes and 60 minutes to complete; the mean self-reported duration of the test was 20.7 minutes (SD = 9.75). 104 individuals (61%) had the Qualtrics survey open for less than 1 hour, indicating that these individuals, likely completed the survey in a single sitting. The survey was designed so that staff could complete this in more than one sitting if needed; in the longest instance, the survey was open to a participant for 5158583 seconds (i.e., 60 days).

6.4.3 Job performance

All individuals, whose contact details were left, were followed up with a request for workplace related feedback. For analysis purposes, colleagues that provided feedback were allocated to one of two groups, these being 'colleagues' or 'supervisors'. The latter category incorporated managers, supervisors, and professional leads, however the term 'supervisors' has been utilised due to the supervisory nature of these roles.

6.4.3.1 *Job performance ratings*

Some respondents received just one rating for their professionalism and effectiveness, whereas other participants received up to four ratings for each. Table 6.10 and Table 6.11 provide the average rating participants obtained for these attributes according to 'all raters', their 'supervisors' alone, and their 'colleagues' only. As demonstrated, the average rating received from supervisors for both professionalism and effectiveness was lower than the average rating received from colleagues.

Table 6.10 Descriptive statistics of the professionalism ratings, by participant

Scale	*N*	Minimum	Maximum	Mean	SD
Overall sample					
- All ratings	152	25	100	84.9	11.8
- Supervisor ratings	108	25	100	81.0	14.4
- Colleague ratings	124	48	100	87.6	10.0
Form 1					
- All ratings	78	44	100	84.7	10.4
- Supervisor ratings	49	44	100	81.1	12.4
- Colleague ratings	64	63	100	87.0	8.9
Form 2					
- All ratings	74	25	100	85.1	13.2
- Supervisor ratings	59	25	100	80.9	15.9
- Colleague ratings	60	48	100	88.3	11.2

Table 6.11 Descriptive statistics of the effectiveness ratings, by participant

Scale	*N*	Minimum	Maximum	Mean	SD
Overall sample					
- All ratings	152	36.3	100	83.4	11.5
- Supervisor ratings	108	26	100	79.5	14.2
- Colleague ratings	124	41.5	100	86.2	10.9
Form 1					
- All ratings	78	54	100	83.7	11.1
- Supervisor ratings	49	54	100	80.1	13.0
- Colleague ratings	64	60	100	86.5	10.2
Form 2					
- All ratings	74	36.3	100	83.0	12.0
- Supervisor ratings	59	26	100	79.1	15.2
- Colleague ratings	60	41.5	100	86.0	11.7

Supervisors typically provided lower ratings of *professionalism* and *effectiveness* than that provided by other colleagues. Supervisor ratings also had a larger variance than colleague ratings. It is possible that participants provided the contact details for colleagues they get on well with, and these colleagues may have, in turn, not wanted to provide negative feedback. A decision was subsequently made to exclude colleague ratings from the analyses when exploring relationships between SJT scores, peer ratings on the BFI-2-S, and ratings of job performance. A t-test observed no difference between participants that did and did not receive supervisor ratings with regards to their age, $t(168) = -0.21$, p = 0.83. Similarly, a Chi-Square test observed no difference between participants that did and did not receive supervisor ratings with regards to their gender, x^2 (N = 108) = 0.31, p = 0.58. Supervisor ratings of both *professionalism* and *effectiveness* were similar and had a negatively skewed distribution (see Figure 6.6). As can be observed, most ratings fell at the upper end.
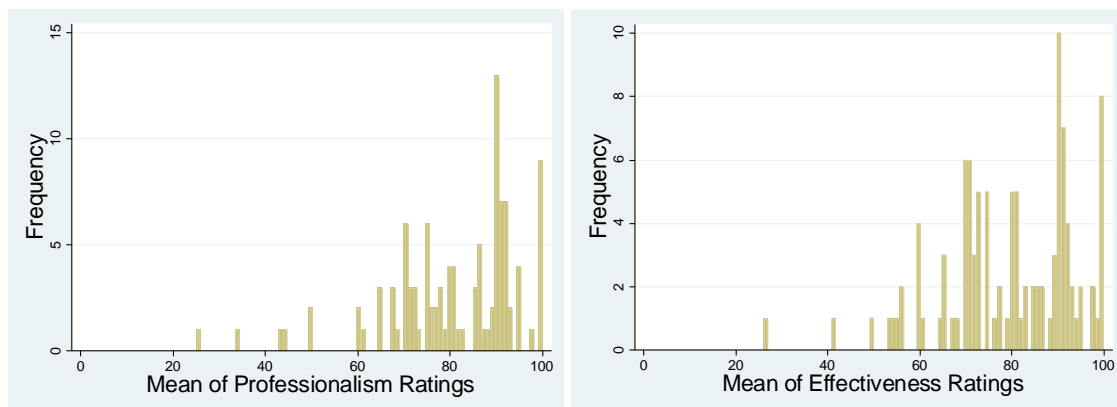
Figure 6.6 Distribution of job performance ratings provided by supervisors

6.4.3.2 *Relationship between ratings of job performance and demographic characteristics*

During a linear regression, no relationship was observed between gender or BAME status with regards to mean supervisor ratings of *professionalism* (p=.255, 95% CI -2.17-8.20; p=.970, 95% CI -7.85-7.55) or *effectiveness* (p=.572, 95% CI -4.34-7.86; p=.545, 95% CI -15.77-8.34), respectively. Due to the distribution of *professionalism*

and *effectiveness* ratings being skewed, the standard errors and confidence intervals were derived through bootstrapping (using 2000 replications). Spearman's Rho correlation coefficient was generated to assess the relationship between ratings of perceived *professionalism* and *effectiveness*. As expected, there was a statistically significant positive correlation between the two variables (r_s=0.79, p=0.000, N=108).

6.4.4 Personality ratings

Self-ratings on the BFI-2-S are hereby provided for all respondents. Peer-ratings are also reported for those individuals that received feedback from supervisors. Where individuals received more than one peer rating (i.e., they had more than one supervisor) mean scores are provided.

Table 6.12 displays the scores obtained for each of the five personality traits according to self-ratings on the BFI-2-S. Self-ratings of *extraversion* and *open-mindedness* aligned to a normal distribution (see Figure 6.7). Ratings of *agreeableness*, *conscientiousness* and *negative emotionality* were relatively normally distributed given the small sample size; however, this sample had a slightly right skew for *negative emotionality* and a slightly left skew for *agreeableness* and *conscientiousness* with most staff falling at the upper end for the latter two traits.

Table 6.12 Self-ratings provided on the BFI-2-S during the pilot study

Scale (BFI-2-S)	*N*	Minimum	Maximum	Mean	SD
Extraversion	170	1.50	5.00	3.42	0.74
Agreeableness	170	2.17	5.00	4.18	0.57
Conscientiousness	170	1.83	5.00	3.94	0.66
Negative Emotionality	170	1.00	5.00	2.49	0.88
Open-Mindedness	170	2.00	5.00	3.76	0.66
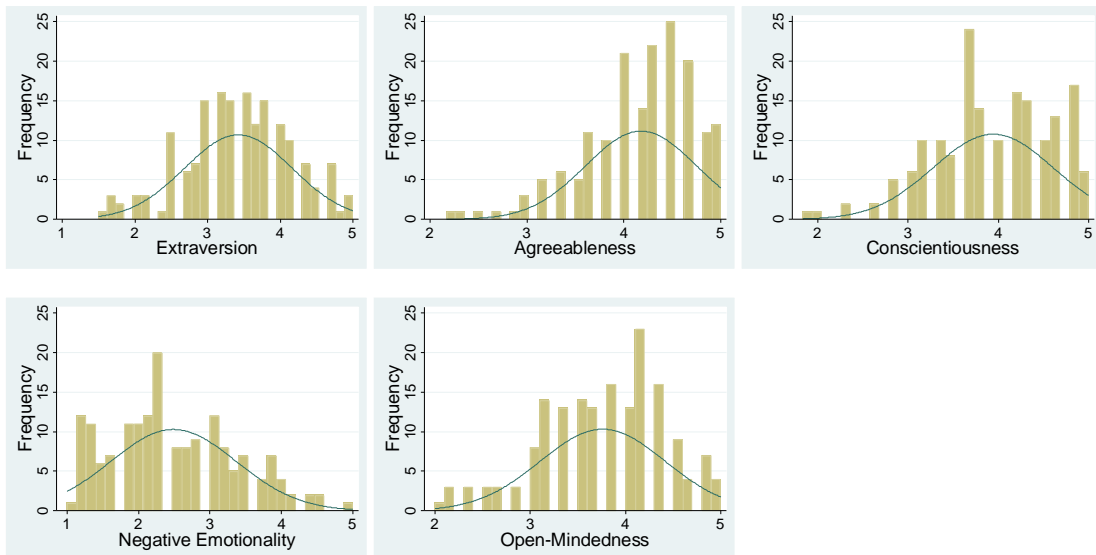
Figure 6.7 Distribution of self-reported scores on the BFI-2-S for each of the five personality traits

Table 6.13 displays the scores obtained for each of the five personality traits according to peer-ratings on the BFI-2-S. The distribution of these scores are demonstrated in Figure 6.8.

Table 6.13 Peer-ratings on the BFI-2-S

Scale (BFI-2-S)	N	Minimum	Maximum	Mean	SD
Extraversion	108	1.33	5	3.65	0.85
Agreeableness	108	1.17	5	4.29	0.72
Conscientiousness	108	1.33	5	4.23	0.68
Negative Emotionality	108	1	5	2.38	0.90
Open-Mindedness	108	2.17	5	3.70	0.57
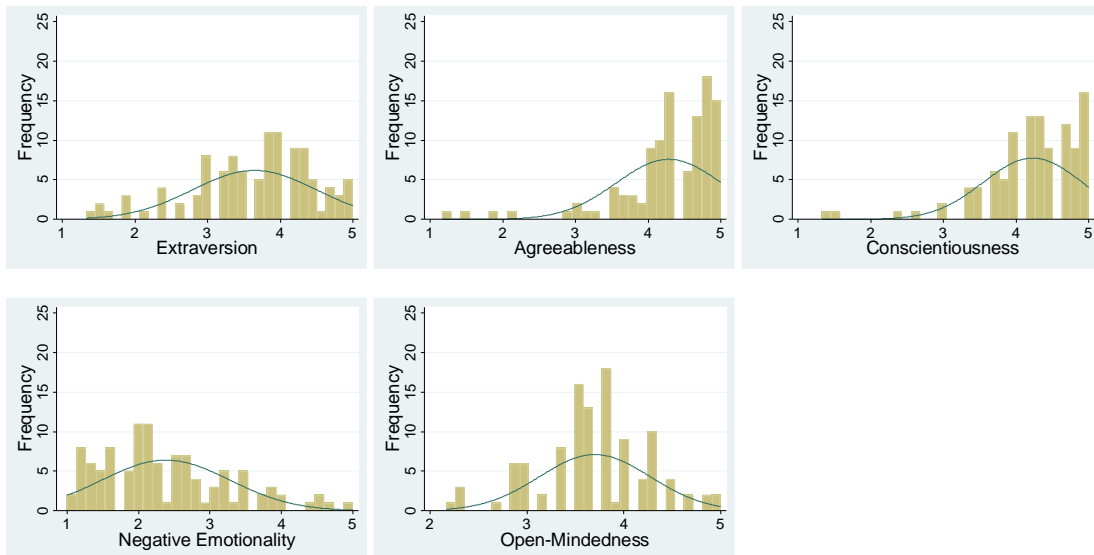
Figure 6.8 Distribution of peer-reported scores on the BFI-2-S for each of the five personality traits

6.4.5 SJT scoring approaches

Spearman's Rho correlation coefficient was used to explore the relationship between SJT scores and job performance ratings provided by colleagues and supervisors. No relationships were observed when all ratings were included in the analyses. The correlation coefficients were as follows for *effectiveness* and *professionalism*, respectively (N=152): SME polytomous scoring (r_s=0.06, p=0.44; r_s=0.04, p=0.62); SME binary scoring (r_s=0.07, p=0.36; r_s=0.08, p=0.30); dichotomous modal consensus scoring (r_s=0.10, p=0.22; r_s=0.01, p=0.87); and proportion modal consensus scoring (r_s=0.13, p=0.12, N=152; r_s=0.05, p=0.58). As noted earlier, colleagues appeared to be more generous than supervisors when providing feedback regarding job performance. A decision was thus made to include only supervisor ratings in the analyses; as mentioned previously, the 'supervisor' in this context includes managers and professional leads.

This section reports on the relationships observed between SJT scores and personality and job performance ratings using four alternative scoring approaches: raw SME polytomous scoring (0,1,2,3), raw SME binary scoring (0,1), dichotomous modal consensus scoring (0,1) and proportion modal consensus scoring (a continuous

proportion; as described in Table 6.6). The range of scores obtained across all four

scoring approaches are displayed in Table 6.14. The distribution of SJT scores across

both forms of the SJT using all four scoring approaches can be viewed as Appendix O.

Using the mean and standard deviation obtained by the respondents for each SJT

form, the SJT scores were standardised (mean 0, sd 1). This crudely equated the SJT

scores from the two forms to permit pooled analyses. The distribution of z scores for

each of the scoring approaches are depicted in Figure 6.9.

Table 6.14 SJT scores for each of the four scoring approaches

Variable	*N*	Minimum	Maximum	Mean	SD
Raw SME polytomous scoring (0,1,2,3)					
- Overall sample	170	98	134	120	7.18
- Form 1	92	98	132	117.4	7.33
- Form 2	78	108	134	123.1	5.61
Raw SME binary scoring (0,1)					
- Overall sample	170	33	49	43.5	2.90
- Form 1	92	33	49	42.7	3.11
- Form 2	78	38	48	44.4	2.35
Dichotomous modal consensus scoring (0,1)					
- Overall sample	170	20	42	31.61	4.83
- Form 1	92	20	41	30.84	4.89
- Form 2	78	21	42	32.51	4.62
Proportion modal consensus scoring					
- Overall sample	170	18.8	31.4	26.0	2.41
- Form 1	92	18.8	28.9	25.1	2.09
- Form 2	78	19.4	31.4	27.1	2.34

The maximum possible score that could be obtained on the SJT was 150 for raw SME polytomous scoring, 50 for raw SME binary scoring and 50 for dichotomous modal consensus scoring.

6.4.5.1 *Relationship between SJT scores and personality ratings*

Prior to the analyses being conducted, it was hypothesised that SJT scores may

correlate with scores obtained on the BFI-2-S, particularly scores relating to an

individual's *conscientiousness*, *agreeableness* and *negative emotionality* as these traits

are particularly important in mental health services. As demonstrated in Table 6.15,

self-ratings of *agreeableness* were statistically significantly related to all SJT scores,

except raw SME binary scores. Peer-ratings on the BFI-2-S were not associated with SJT

scores for any of the scoring approaches.
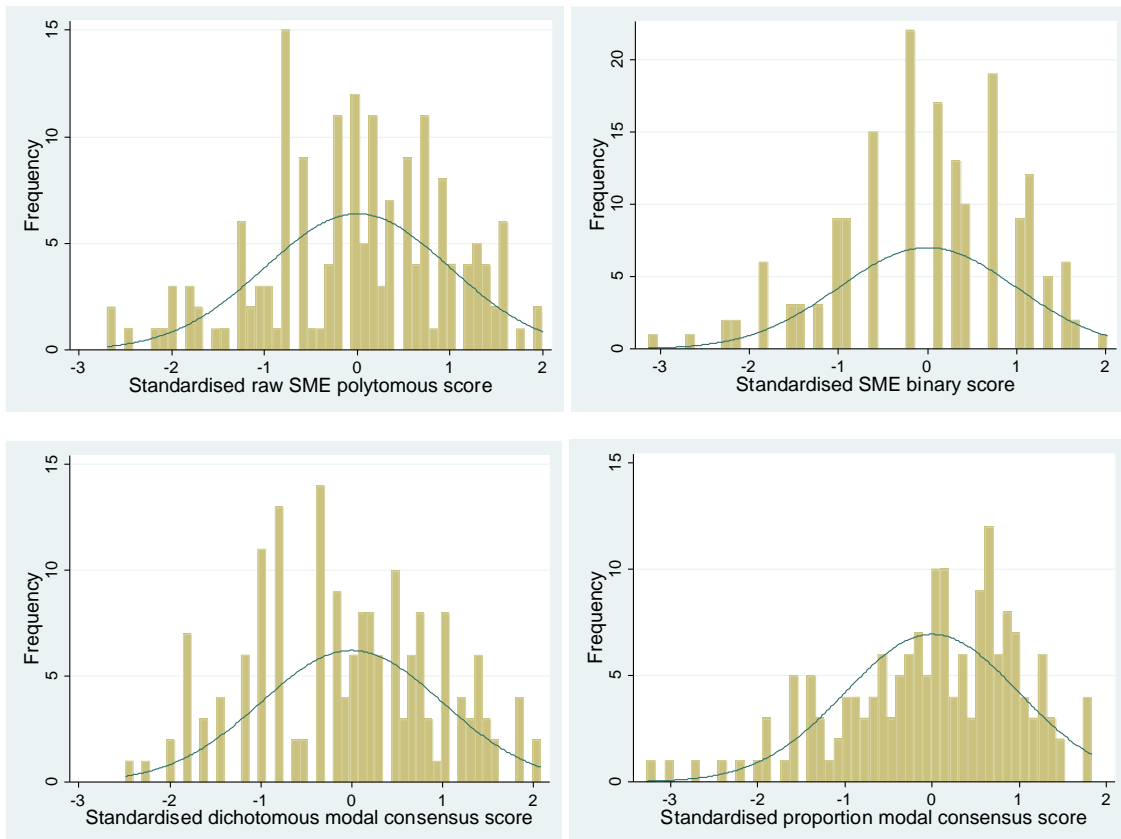
Figure 6.9 Histograms displaying the distribution of standardised SJT scores

6.4.5.2 *Relationship between SJT scores and job performance*

The main aim of the current study was to assess the criterion-related validity of the SJT

scores. That is, do the scores obtained on the SJT predict job performance. Spearman's

Rho correlation coefficient was used to explore the relationship between SJT scores

and supervisor ratings of *professionalism* or *effectiveness*. No significant relationships

were observed between SJT scores and supervisor ratings of *professionalism* or

effectiveness using raw polytomous scoring, raw SME binary scoring and dichotomous

modal consensus scoring (see Table 6.15). However, when proportion modal

consensus scoring was utilised, a significant correlation was observed between SJT

scores and supervisor ratings of both *effectiveness* (r_s=0.21, p=0.03, N=108) and

professionalism (r_s=0.21, p=0.03, N=108). Furthermore, a simple regression with

bootstrapping (employing 2000 replications) demonstrated that proportion modal

consensus SJT scores statistically significantly predicted ratings of *professionalism* and

effectiveness ($p<0.05$), thus demonstrating the SJT scores criterion-related validity.

Table 6.15 Means, Standard Deviations, and Intercorrelations of Study Variables (supervisor ratings only)

Variable	1	2	3	4	5	6	7	8	9	10
1. Experience (N=168)										
Standardised SJT score (N=170)										
2. Raw SME polytomous score	.015									
3. Raw SME binary score	.017	**-850*****								
4. Dichotomous modal consensus score	.043	**.540*****	**.323*****							
5. Proportion modal consensus score	.019	**.593*****	**.395*****	**.923*****						
Self-ratings on the BFI-2-S (N=170)										
6. S Extraversion	.066	.085	.046	.117	.112					
7. S Agreeableness	.042	**.168***	.129	**.247****	**.267*****	.131				
8. S Conscientiousness	.037	.093	.052	.077	.097	**.177***	**.238****			
9. S Negative Emotionality	-.132	-.073	-.067	-.116	-.112	**-.314*****	**-.269*****	**-.291*****		
10. S Open-Mindedness	.088	.103	.032	.114	.063	**.201****	.057	.032	-.008	
Peer-ratings on the BFI-2-S (N=108)										
11. P Extraversion	.058	-.088	-.107	-.024	-.023	**.445*****	.001	.154	-.187	.056
12. P Agreeableness	**-.207***	.023	.065	-.019	.031	-.149	.138	.040	.091	-.030
13. P Conscientiousness	-.158	.128	.127	-.017	.018	-.064	.027	**.317*****	.027	-.105
14. P Negative Emotionality	.041	.109	.000	.020	.017	.005	.006	**-.198***	**.416*****	-.036
15. P Open-Mindedness	-.006	.022	.053	-.110	-.065	.050	-.007	.050	-.035	.186
Job performance (N=108)										
16. Effectiveness	.011	.069	.077	.112	**.207***	.088	.190	**.207***	-.134	.060
17. Professionalism	.045	.091	.132	.133	**.207***	-.121	.169	.110	-.038	-.023
M	13.58	0	0	0	0	3.42	4.18	3.94	2.49	3.76
SD	9.40	1	1	1	1	.74	.57	.66	.88	.66

Bold = significant result (*p<0.05; **p<0.01; ***p<0.001)

Calculated using Spearman's Rho

Table 6.15 *continued*

Variable	11.	12.	13	14.	15.	16.	17.
1. Experience (N=168)							
Standardised SJT score (N=170)							
2. Raw SME polytomous score							
3. Raw SME binary score							
4. Dichotomous modal consensus score							
5. Proportion modal consensus score							
Self-ratings on the BFI-2-S (N=170)							
6. S Extraversion							
7. S Agreeableness							
8. S Conscientiousness							
9. S Negative Emotionality							
10. S Open-Mindedness							
Peer-ratings on the BFI-2-S (N=108)							
11. P Extraversion							
12. P Agreeableness	.106						
13. P Conscientiousness	**.341***	**.273**					
14. P Negative Emotionality	**-.351***	**-.326***	**-.380***				
15. P Open-Mindedness	**.380***	**.317***	**.190***	**-.298**			
Job performance (N=108)							
16. Effectiveness	**.379***	.175	**.441***	**-.426***	**.190***		
17. Professionalism	.173	**.366***	**.346***	**-.343***	.148	**.791***	
M	*3.65*	*4.29*	*4.23*	*2.38*	*3.70*	*79.55*	*81.00*
SD	*.85*	*.72*	*.68*	*.90*	*.57*	*14.20*	*14.35*

Bold = significant result (*p<0.05; **p<0.01; ***p<0.001)

Calculated using Spearman's Rho

6.4.5.3 *Group differences*

As SJTs are often used in high stakes selection settings, it is important that they do not unfairly discriminate against certain groups (for example, by BAME status). Of participants that received supervisor ratings, all but seven described themselves as white. Due to the limited participants of non-white ethnicity, it was not possible to evaluate group differences based on ethnicity. An independent-samples t-test showed that there was no statistically significant difference between females' and males' SJT scores using raw SME polytomous scoring, raw SME binary scoring, and dichotomous modal consensus scoring ($p > 0.05$ in all cases). A linear regression analysis with bootstrapping (employing 2000 replications) observed no statistically significant difference between females and males scores on the SJT using proportion modal consensus scoring ($p > 0.05$).

6.4.6 Scoring the SJT by discipline – nurses and allied health professionals (AHPs)

Prior research reports that SJTs should be developed based on a thorough job analysis and be specific to the attributes required of a role (Patterson et al., 2016d; 2018). Nurses and AHPs typically have more face-to-face contact with patients than the other professional groups (e.g., psychologists) and exploratory analyses found that for dichotomous modal consensus scoring, and proportion modal consensus scoring the validity coefficients for the SJT scores were larger for this group of participants. Nurses and AHPs were also the two largest staff groups in the study sample. These two factors meant from this point forward the results focus on the relationship between consensus scored SJTs and the outcomes of interest in these two professional groups. In total, 69 of the 109 nurses and AHPs that participated obtained feedback from their supervisors, which included managers and professional leads. A linear regression and logistic regression observed no difference between participants that did and did not receive supervisor ratings with regards to their age ($p=.942$, 95% CI -4.12-3.83) and gender ($p=.492$, 95% CI .47-4.54), respectively. Of the 109 nurses and AHPs that participated in the study, 59 completed form 1 and 50 completed form 2. The nurses'

and AHPs' demographics and professional characteristics are provided in Table 6.16

and Table 6.17, respectively. Data regarding the setting and the specialty nurses and

AHPs worked is provided in Table 6.18.

Table 6.16 Demographics of nurses and allied health professionals

Demographics	Overall (n=109) N (%)	Form 1 (n=59) N (%)	Form 2 (n=50) N (%)
Age, years (mean, SD)	41.5 (SD – 10.04)	42.1 (SD – 9.75)	40.9 (SD – 10.44)
Gender			
- Female	92 (84.4)	51 (86.4)	41 (82.0)
- Male	17 (15.6)	8 (13.6)	9 (18.0)
Ethnicity			
- BAME	1 (0.9)	1 (1.7)	0 (0.0)
- White	108 (99.1)	58 (98.3)	50 (100.0)
Sexuality			
- Bisexual	5 (4.6)	1 (1.7)	4 (8.0)
- Gay or Lesbian	3 (2.8)	1 (1.7)	2 (4.0)
- Heterosexual or Straight	100 (91.7)	57 (96.6)	43 (86.0)
- Prefer not to say	1 (0.9)	-	1 (2.0)
Religion			
- Christian	51 (46.8)	30 (50.9)	21 (42.0)
- No religion	56 (51.4)	29 (49.2)	27 (54.0)
- Other	1 (0.9)	-	1 (2.0)
- Prefer not to say	1 (0.9)	-	1 (2.0)
Disability			
- No	99 (90.8)	53 (89.8)	46 (92.0)
- Yes	9 (8.3)	6 (10.2)	3 (6.0)
- Prefer not to say	1 (0.9)	-	1 (2.0)
Primary language is English			
- No	1 (0.9)	1 (1.7)	0 (0.0)
- Yes	108 (99.1)	58 (98.3)	50 (100.0)

Table 6.17 Professional characteristics of nurses and allied health professionals

	Overall (n=109) N (%)	Form 1 (n=59) N (%)	Form 2 (n=50) N (%)
Professional discipline			
- Allied health professional	36 (33.0)	22 (37.3)	14 (28.0)
- Nurse	73 (67.0)	37 (62.7)	36 (72.0)
Years registered			
- 1 – 5 years	23 (21.1)	12 (20.3)	11 (22.0)
- 6 – 10 years	26 (23.9)	14 (23.7)	12 (24.0)
- 11 – 15 years	22 (20.2))	16 (27.1)	6 (12.0)
- 16 – 20 years	12 (11.0)	4 (6.8)	8 (16.0)
- 21 – 25 years	8 (7.3)	2 (3.4)	6 (12.0)
- 26 – 30 years	8 (7.3)	6 (10.2)	2 (4.0)
- 31 – 35 years	8 (7.3)	4 (6.8)	4 (8.0)
- 36 – 40 years	1 (0.9)	1 (1.7)	-
- 41 – 45 years	1 (0.9)	-	1 (2.0)
AFC Band			
- 5	15 (13.8)	8 (13.6)	7 (14.0)
- 6	51 (46.8)	27 (45.8)	24 (48.0)
- 7	28 (25.7)	16 (27.1)	12 (24.0)
- 8	13 (11.9)	7 (11.9)	6 (12.0)
- Other	2 (1.8)	1 (1.7)	1 (2.0)

Table 6.18 Setting and specialty worked by nurses and allied health professionals

	Overall (n=109) N (%)	Form 1 (n=59) N (%)	Form 2 (n=50) N (%)
Setting			
- Community services	87 (79.8)	45 (76.3)	42 (84.0)
- Inpatient services	27 (24.8)	18 (30.5)	9 (18.0)
- Corporate services	7 (6.4)	4 (6.8)	3 (6.0)
Specialty			
- All services, Trustwide	8 (7.3)	4 (6.8)	4 (8.0)
- Adult Mental Health services	44 (40.4)	23 (39.0)	21 (42.0)
- Child and Adolescent services	21 (19.3)	10 (17.0)	11 (22.0)
- Forensic Learning Disability services	7 (6.4)	5 (8.5)	2 (4.0)
- Forensic Mental Health services	14 (12.8)	7 (11.9)	7 (14.0)
- Learning Disability services (adults)	5 (4.6)	4 (6.8)	1 (2.0)
- Learning Disability services (children)	-	-	-
- Mental Health Services for Older People	27 (24.8)	15 (25.4)	12 (24.0)

NB: some individuals worked in more than one specialty and/or setting, thus, the figures noted above tally to more than the number of participants

Consensus scoring keys were generated for the SJT taking only the responses provided by nurses and AHPs into account. The range of SJT scores obtained by nurses and AHPs using dichotomous modal consensus scoring and proportion modal consensus scoring are displayed in Table 6.19. The distribution of SJT scores obtained by nurses and AHPs, using discipline specific consensus scoring approaches are displayed in Appendix P.

Table 6.19 Nurses and allied health professionals SJT scores using consensus scoring

Variable	*N*	Minimum	Maximum	Mean	SD
Dichotomous modal consensus scoring (0,1)					
- Overall sample	109	23	42	33.36	4.28
- Form 1	59	26	42	32.68	3.80
- Form 2	50	23	42	34.16	4.70
Proportion modal consensus scoring					
- Overall sample	109	20.98	32.13	27.18	2.44
- Form 1	59	21.24	29.92	26.30	1.99
- Form 2	50	20.98	32.13	28.23	2.53

Similar to that reported previously, SJT scores were standardised (mean 0, sd 1) as a means of (crudely) equating the scores from the two forms of the test allowing pooled analyses to be conducted. The mean and standard deviation used for this calculation related to nurses' and AHPs' scores only. The distribution of these scores is depicted in Figure 6.10.

6.4.6.1 ***Dichotomous modal consensus scoring (nurses and allied health professionals)***
No statistically significant relationships were observed between dichotomous modal consensus SJT scores (0,1) and self or peer BFI-2-S ratings (see Table 6.20). However, a linear regression analysis observed that nurses' and AHPs' SJT scores significantly predicted ratings of both *professionalism* (β = 0.31, p=0.009) and *effectiveness* (β = 0.32, p=0.007) using discipline specific dichotomous modal consensus scoring (N=69).
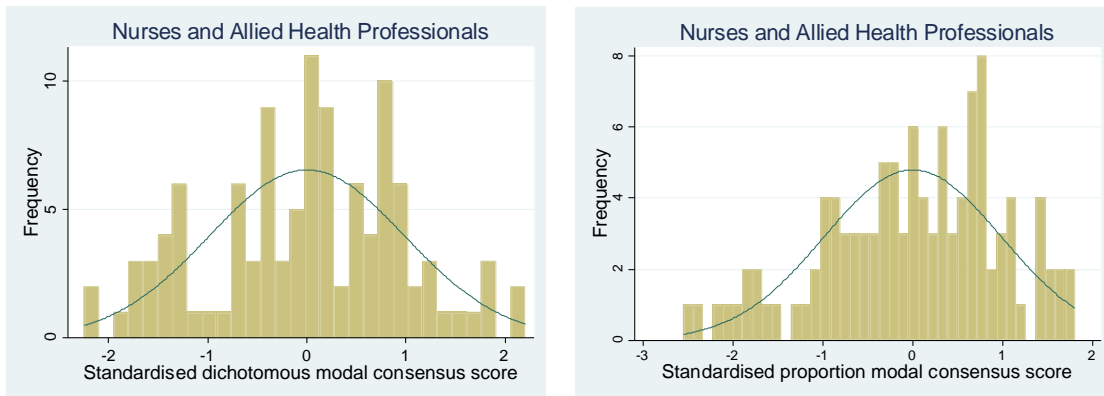
Figure 6.10 Histogram displaying the distribution of standardised SJT scores across nurses and allied health professionals

6.4.6.2 ***Proportion modal consensus scoring (nurses and allied health professionals)***
Similar to that reported for the wider sample, a statistically significant relationship was observed between nurses and AHPs discipline specific proportion modal consensus SJT scores and self ratings of *agreeableness* (see Table 6.20). No further relationships were observed between proportion modal consensus SJT scores and other ratings on the BFI-2-S. A linear regression demonstrated that the SJT scores statistically significantly predicted supervisor ratings of *professionalism* and *effectiveness* using discipline specific proportion modal consensus scoring (β = 0.31, p=0.010; β = 0.34, p=0.005), respectively.

Table 6.20 Means, Standard Deviations, and Intercorrelations of Study Variables (across nurses and allied health professionals; supervisor ratings only)

Variable	1	2	3	4	5	6	7	8	9	10
1. Experience (N=109)										
Standardised SJT score (N=109)										
2. Dichotomous modal consensus score	.026									
3. Proportion modal consensus score	.013	**.901*****								
Self-ratings on the BFI-2-S (N=109)										
4. S Extraversion	.063	-.010	-.024							
5. S Agreeableness	.084	.188	**.255****	.118						
6. S Conscientiousness	-.052	-.039	.021	.126	**.242***					
7. S Negative Emotionality	-.111	-.016	-.001	**-.262****	**-.248****	**-.265****				
8. S Open-Mindedness	.077	.090	.050	**.239***	-.010	.001	.080			
Peer-ratings on the BFI-2-S (N=69)										
9. P Extraversion	-.050	.060	.095	**.431*****	.117	**.240***	-.154	.039		
10. P Agreeableness	-.138	-.040	.034	-.091	.202	.023	.045	-.039	.169	
11. P Conscientiousness	**-.246***	.071	.091	.006	.212	**.448*****	-.052	-.185	**.360****	**.279***
12. P Negative Emotionality	.013	-.054	-.092	.038	-.030	**-.310****	**.482*****	.022	**-.344****	**-.319****
13. P Open-Mindedness	-.066	.018	.034	.007	.145	.042	.001	.141	**.386****	**.427*****
Job performance (N=69)										
14. Effectiveness	-.057	**.302***	**.328****	.131	**.357****	**.295***	-.198	.096	**.309****	.147
15. Professionalism	.073	**.264***	**.265***	-.133	**.260***	.148	-.024	-.025	.128	**.335****
M	14.36	0	0	3.52	4.19	3.92	2.52	3.72	3.73	4.38
SD	9.62	1	1	.72	.54	.66	.87	.69	.84	.67

Bold = significant result (*p<0.05; **p<0.01; ***p<0.001)
Calculated using Spearman's Rho

Table 6.20 *continued*

Variable	11	12	13	14	15
1. Experience (N=109)					
Standardised SJT score (N=109)					
2. Dichotomous modal consensus score					
3. Proportion modal consensus score					
Self-ratings on the BFI-2-S (N=109)					
4. S Extraversion					
5. S Agreeableness					
6. S Conscientiousness					
7. S Negative Emotionality					
8. S Open-Mindedness					
Peer-ratings on the BFI-2-S (N=69)					
9. P Extraversion					
10. P Agreeableness					
11. P Conscientiousness					
12. P Negative Emotionality	**-.368****				
13. P Open-Mindedness	.203	**-.310****			
Job performance (N=69)					
14. Effectiveness	.393***	**-.411*****	**.289***		
15. Professionalism	**.337****	**-.392*****	**.331****	**-80*****	
M	4.26	2.39	3.66	80.57	82.92
SD	.65	.96	.59	14.50	12.64

Bold = significant result (*p<0.05; **p<0.01; ***p<0.001)
Calculated using Spearman's Rho

6.4.6.3 *Group differences across the nurse and allied health professional sample*

All but one individual in the nurse and AHP sample were of white ethnicity; thus, group differences according to ethnicity were not explored for either consensus scoring approach. An independent-samples t-test and a Mann-Whitney U test found no difference between females and males SJT scores using either dichotomous modal consensus scoring (t(107)=0.604, p=0.547) or proportional modal consensus scoring (p=0.260), respectively. Prior studies have observed a relationship between individuals' experience on the job and their SJT scores. A Spearman's Rho analysis observed no statistically significant relationship between the number of years nurses and AHPs had been registered to the profession and their SJT scores using either dichotomous modal consensus scoring or proportion modal consensus scoring (p>0.05). Spearman's Rho correlation coefficients are presented in Table 6.20.

6.4.7 Most valid scoring approach

Analyses observed that effect sizes were larger for nurses and AHPs, as opposed to the wider sample, when using discipline specific dichotomous modal consensus scores and proportion modal consensus scores to predict supervisor ratings of job performance. It is thus recommended that the current SJT be used with nurses and AHPs for personnel selection. SJT scores statistically significantly predicted ratings of *professionalism* and *effectiveness* for nurses and AHPs using both consensus scoring approaches. It may be more pragmatic to use dichotomous modal consensus scoring for the final assessment however, given that both scoring approaches are similarly effective. A T-test observed no difference between nurses and AHPs that sat form 1 as opposed to form 2 regarding the scores they obtained on the shared items of the SJT assessment (t(107)=0.211, p=0.833); this finding indicates that these two groups of participants were unlikely to differ in their overall ability to answer the SJT.

6.4.8 Face validity and acceptability

The face validity of an SJT is important as this may influence prospective employees' perceptions of fairness, as well as their perception of the appointing organisation. The face validity and acceptability of the current SJT was evaluated by asking participants whether they thought the test was relevant to their role, appropriate to their grade, suitable for the selection and recruitment of staff into mental health services, and fair to all applicants, regardless of their profession and demographic characteristics. The responses provided by nurses and AHPs are reported in Table 6.21 and Table 6.22, respectively. Results that incorporate the views of all participants, regardless of profession, are presented in Appendix Q.

Table 6.21 Face validity of the SJT according to nurses

	Overall (n=73) N (%)	Form 1 (n=37) N (%)	Form 2 (n=36) N (%)
Relevant to role			
- Yes	70 (95.9)	36 (97.3)	34 (94.4)
- No	3 (4.1)	1 (2.7)	2 (5.6)
Appropriate difficulty for grade			
- Yes	66 (90.4)	35 (94.6)	31 (86.1)
- No	7 (9.6)	2 (5.4)	5 (13.9)
Suitable for recruitment			
- Yes	68 (93.2)	33 (89.2)	35 (97.2)
- No	5 (6.9)	4 (10.9)	1 (2.8)
Fair to applicants			
- Yes	66 (90.4)	33 (89.2)	33 (91.7)
- No	7 (9.6)	4 (10.8)	3 (8.3)

Table 6.22 Face validity of the SJT according to allied health professionals

	Overall (n=36) N (%)	Form 1 (n=22) N (%)	Form 2 (n=14) N (%)
Relevant to role			
- Yes	35 (97.2)	22 (100)	13 (92.3)
- No	1 (2.8)	0 (0)	1 (7.1)
Appropriate difficulty for grade			
- Yes	35 (97.2)	22 (100)	13 (92.3)
- No	1 (2.8)	0 (0)	1 (7.1)
Suitable for recruitment			
- Yes	33 (91.7)	20 (90.9)	13 (92.9)
- No	3 (8.3)	2 (9.1)	1 (7.1)
Fair to applicants			
- Yes	29 (80.6)	18 (81.8)	11 (78.6)
- No	7 (19.4)	4 (18.2)	3 (21.4)

6.4.8.1 *Feedback*

Most nurses stated that the SJT was relevant to their role, an appropriate difficulty for their grade, suitable for recruitment and fair to all applicants. Like nurses, the majority of AHPs noted that the SJT was relevant to their role, an appropriate difficulty for their grade and suitable for recruitment. Whilst the majority of AHPs felt the SJT was fair to all applicants, some AHPs did not. There was a concern that differing expectations would be placed on individuals at different points in their career, and that multiple choice items may be quite limiting. A couple of comments provided by AHPs with regards to the suitability of the SJT for personnel selection are presented below:

> "*Different grades would have different expectations i.e., someone in a senior role would be addressing issues themselves where as support staff maybe more expected to report behaviours.*" (AHP)

> "*I'm not sure a multiple choice would work well for getting a full picture of what the person is like. Perhaps if this, combined with a chance for the person to speak openly (not multiple choice) about situations they have experienced may work well.*" (AHP)

6.4.9 Selecting the final SJT items

Whilst a decision was made to assess the internal consistency of the SJT scores (see heading 6.3.13), no items were removed based on these findings. SJTs are known to be heterogeneous in nature and tap into multiple separate domains; one would, therefore, not want to increase the reliability of the instrument at the expense of its predictive validity. Instead of removing items that had lower internal consistency, items were removed if they had a negative relationship with professionalism (i.e., odds ratio < 0.98), or if the participant responses did not vary using dichotomous modal consensus scoring (see Appendix R for the related odds ratios; items that are not scored in the final SJT assessment are denoted by an Asterix).

6.4.9.1 *Simulation to determine the criterion validity of trimmed version*

Once items had been shortlisted for the trimmed version of the test, a 'simulation' was performed to determine the predictive validity of the resulting SJT items using discipline specific dichotomous modal consensus scoring. That is, a 'final test score' was created and the ability of these scores to predict the professionalism and effectiveness rating was evaluated. In this regard, firstly, First, a total mean score was calculated for each nurse and AHP based on their responses to the final items only. Following this, z scores (mean 0, sd 1) were calculated for each of these individuals using the mean and standard deviation obtained for each form separately. Z scores were subsequently regressed against *professionalism* and *effectiveness*. Taking the final items only into consideration, nurses and AHPs modal consensus SJT z scores statistically significantly predicted ratings of both *professionalism* (N=69; β = 0.41, p<0.001) and *effectiveness* (N=69; β = 0.38, p=0.001).

As noted previously, standardised scores (mean 0, sd 1) were calculated so that the SJT forms could be analysed together. The assumption that this equating was, at least, crudely effective at resulting in scores from the two test forms with a similar interpretation was subsequently evaluated. This was by observing the relationship between the two separate test forms standardised dichotomous modal consensus

scores and the primary outcomes of interest (supervisor ratings of professionalism and effectiveness). Incorporating only the final items in the analysis across the nurse and AHP sample, the regression coefficients were similar for ratings of both *professionalism* (β = 0.53, p=0.003; β = 0.36, p=0.025) and *effectiveness* (β = 0.45, p=0.013; β = 0.37, p=0.022) for forms 1 and 2, respectively.

6.4.10 Incremental validity of the SJT for nurses and allied health professionals

Some previous studies have observed SJT scores to provide incremental validity over and above personality traits in predicting job performance (e.g. Clevenger et al., 2001). Noting that it would be more pragmatic to use dichotomous modal consensus scoring for the final assessment, the current study sought to evaluate whether dichotomous modal consensus SJT scores possessed incremental validity over and above personality ratings when predicting job performance. Although this study collected peer-ratings on the BFI-2-S as well as self-ratings, it is typically not possible to obtain personality ratings from prospective employers. The current author thus decided to explore only whether SJT scores possessed incremental validity over and above 'self-ratings' only, for nurses and AHPs.

Due to the limited number of participants, only variables that significantly predicted ratings of job performance were put into the incremental validity analyses in order to increase study power. A Spearman's Rho analysis observed a statistically significant relationship between supervisor ratings of *effectiveness* and self-ratings of both *conscientiousness* and *agreeableness* among the nurse and AHP sample (see Table 6.20). Self-ratings of *agreeableness* were statistically significantly associated with supervisor ratings of *professionalism*. Controlling for self-ratings of *agreeableness* and *conscientiousness*, a regression observed that dichotomous modal consensus SJT scores provided no statistically significant additional predictive validity with regards to ratings of *effectiveness* (β = 0.20, p=0.098). The SJT scores did, however, provide incremental validity over and above self-rated *agreeableness* in predicting supervisor ratings of *professionalism* (β = 0.26, p<0.05).

226

6.4.11 Group differences with the trimmed version

It is important that any selection procedure does not unfairly discriminate against certain groups, such as disability. Whilst the results relating to an exploration of gender differences in SJT scores have previously been reported, further analyses sought to evaluate whether there were any other group differences, among the sample, using the shortlisted SJT items only. It is recognised that experience could be considered a proxy for a staff member's grade, however as suggested by an AHP during the study '*Different grades would have different expectations.*' The grade of a staff member may therefore influence the response a staff member provides on the SJT, regardless of their experience on the job. Whilst a relationship was not observed between SJT scores and the number of years experience staff members possessed, a Kruskal-Wallis H test was also conducted to determine if dichotomous modal consensus z scores for the final items differed between AFC Band 5 staff (n=15), AFC Band 6 staff (n=51), AFC Band 7 staff (n=28) and AFC Band 8 staff (n=14). The Kruskal-Wallis H test showed that there was no difference in SJT scores between the four groups, $\chi2 = 0.575$, p = 0.902. A linear regression also observed no relationship between dichotomous modal consensus z scores and age (p=0.891). Finally, a Wilcoxon signed rank test found no significant difference in dichotomous modal consensus z scores obtained between those that did or did not report having a disability (Z = 0.195, p=0.846), those that did and did not identify as being heterosexual (Z = 1.571, p=0.116), and those that identified themselves as being a Christian or having no religion (Z = 0.219, p=0.827).

6.4.12 Factor analysis / dimensionality

Noting that factor structure is linked to reliability, exploratory factor analysis was performed on the final versions of each SJT form. No clear factor structure was observed. Rather, the structure of the response data appeared 'essential unidimensional' (Nandakumar, 1991), also previously referred to as 'fuzzy unidimensionality' (Tiffin et al., 2020). This is a very common picture in the factor analysis of SJT response data, due to both the effects of dependency of responses

within the same scenarios (stems) but also that even individual response options may tap into multiple traits or abilities (Guenole et al., 2017). The main findings from the factor analyses can be viewed as Appendix S.

6.4.13 Evaluation of method effects and reliability

Analyses next explored whether question type influenced the predictive validity coefficients of the trimmed version; the lead researcher was interested to know whether *appropriateness* or *importance* items tapped into one's *professionalism* or *effectiveness* more than the other. Regression coefficients were therefore compared for each set of items (see Table 6.23).

Table 6.23 Relationship between question type and supervisor ratings of professionalism and effectiveness using the trimmed version for the nurse and allied health professional sample

Question type - domain	Coefficient	SE (95% CI)	p
Importance - professionalism	0.0538	0.0265 (0.0019 to 0.1058)	0.04*
Appropriateness - professionalism	0.0840	0.0258 (0.0334 to 0.1345	0.00***
Importance - effectiveness	0.0563	0.0214 (0.0144 to 0.0982)	0.01**
Appropriateness - effectiveness	0.0643	0.0204 (0.0242 to 0.1043)	0.00***

*p<0.05; **p<0.01; ***p<0.001

Whilst both question formats predicted ratings of job performance, interestingly, a Spearman's Rho correlation coefficient observed no relationship between the total scores obtained for the final appropriateness items and the total scores obtained for the final importance items (r_s=0.11, p=0.261, N=109). This finding indicates that the different format of questions are both orthogonal, but relevant traits for predicting job performance.

A Kuder Richardson KR20 reliability analysis was carried out on the final items for each SJT form using dichotomous modal consensus scores. Form 1, which has 44 items, obtained an alpha coefficient of 0.45 and Form 2, which has 41 items, obtained an alpha coefficient of 0.38 indicating low reliability. In addition to analysing the reliability

of item scores using the trimmed version, item scores were summed for each scenario that they corresponded to; ordinal summed scores were subsequently fed into a reliability analysis and resulted in McDonald's omega values of 0.47 for form 1 and 0.64 for form 2, thus also indicating relatively low reliability. Whilst both forms obtained poor internal consistency values, this is not atypical for traditional SJTs.

6.5   Discussion

This study was undertaken to provide evidence for the validity of an SJT that was developed for use in mental health services. This involved evaluating the feasibility of using the SJT in this setting. In addition, the criterion-related validity of four alternative scoring approaches for the SJT were evaluated. Evidence relating to convergent and divergent validity was explored in relation to the scores obtained on a personality assessment, namely the BFI-2-S. The incremental validity of the SJT scores, in relation to personality self-report scores was also assessed.

6.5.1   Summary of key findings

The study found that when consensus scoring approaches were utilised, which included dichotomous modal consensus scoring (0,1) and proportion modal consensus scoring (a continuous proportion between 0 and 1), the SJT scores statistically significantly predicted supervisor ratings of perceived *professionalism* and *effectiveness* in a sample of nurses and AHPs working in mental health services (n=69). It was found that whilst supervisor ratings had a strong relationship with SJT scores, colleague ratings did not. In this regard, on average, colleagues rated participants more favourably and were possibly more hesitant in providing negative feedback. SJT scores did not predict supervisor ratings of job performance when SME-based scoring approaches were utilised. Across nurses and AHPs, scores obtained on the SJT were statistically significantly associated with self-ratings of *agreeableness* when using proportion modal consensus scoring. In addition, self-ratings of *agreeableness* were associated with supervisor ratings of perceived *professionalism* and *effectiveness*. Self-

ratings of *conscientious* were also associated with supervisor ratings of *effectiveness*, but not perceived *professionalism*. Furthermore, the modal consensus SJT scores were found to possess incremental validity over and above self-rated *agreeableness* in predicting supervisor ratings of *professionalism*.

The above findings are important because they support the SJT (consensus) scores as valid measures that have some ability to predict mental health nurses' and AHPs' performance on the job, at least as perceived by managers and supervisors. Furthermore, the SJT scores were observed to possess incremental validity over self-rated personality traits in predicting supervisor ratings of perceived *professionalism* at work. The SJT could therefore be used as a cost-effective, scalable approach to facilitate the selection of suitable staff for working in mental health services.

6.5.2 Comparisons with the existing literature

The findings of this study support Motowidlo and Beier's theory that SJTs assess individuals' implicit trait policies (ITPs, 2010). Self-rated *agreeableness* for instance, as identified via the BFI-2-S was associated with SJT scores across three of the four scoring approaches. It has been recommended previously that a job analysis be undertaken prior to developing an SJT in order to identify the key attributes required for a role (Patterson and Driver, 2018). However, Motowidlo and Beier (2010) suggest that where ITPs have a level of stability across jobs and organisational settings, which could include mental health roles, it may be beneficial to develop an SJT that would assess those ITPs even if the specific knowledge required for these roles differ. It is not so surprising that the SJT scores were related to self-reported agreeableness because ITPs related to agreeableness are beneficial for staff that work in mental health services and provide face to face care.

Prior research has demonstrated that a text-based SJT can predict supervisor ratings of performance among undergraduate medical and dental school admissions (Patterson et al., 2017). SJT scores have been found to predict end of training assessment scores

in postgraduate settings also (Patterson et al., 2016c; Webster et al., 2020). The current study builds on these findings by demonstrating that an SJT, used in an electronic format, and delivered online, can predict ratings related to actual clinical practice among mental health nurses and AHPs. The online format of the SJT increases its accessibility, which is particularly beneficial in circumstances where it is difficult to deliver assessments in person, such as during the recent Covid-19 pandemic. Whilst there have previously been concerns about delivering SJT content online for reasons such as test fraud, procedures can be undertaken to mitigate against these concerns. For example, one firm currently requests applicants to provide government issued ID and use a webcam whilst undertaking other security measures too (Altus assessments, 2022).

The current study observed that SJT scores alone predicted supervisor ratings of job performance, however, it is unusual for SJTs to be the only means of selecting candidates for a role. Instead, SJTs are typically used to complement other selection assessments. Such tests are often used as a selection step before other more resource intensive selection procedures, such as multiple-mini interviews (Patterson et al., 2016a).

6.5.3 Interpretation of findings

It was clear at the beginning of the analyses that colleagues rated participants more favourably than supervisors and that this negatively impinged on the predictive validity of the SJT scores. It is possible that participants chose their preferred colleagues to provide feedback, particularly due to the additional financial incentive. In turn, colleagues may not have provided as honest feedback as supervisors and subsequent analyses therefore explored the relationship between SJT scores and supervisor ratings only. The results demonstrated that when the proportion modal consensus scoring was used, SJT scores predicted supervisor ratings of perceived *professionalism* and *effectiveness* for the overall sample, as well as for nurses and AHPs. Dichotomous modal consensus scores also predicted ratings of *professionalism* and *effectiveness*

when used across mental health nurses and AHPs only. The SJT therefore possessed greater predictive validity when used with nurses and AHPs as opposed to other professional groups working in mental health services.

There are a number of reasons that may have resulted in the above finding. Nurses and AHPs typically provide more face-to-face care than psychiatrists or psychologists, with the latter tending to fill more strategic, supervisory, consultancy and leadership roles. As noted previously, SJT developers have recommended that a thorough job analysis be conducted prior to developing an SJT in order to determine the key attributes required for a role. As mental health roles differ, it is understandable that there is no 'one size fits all' when it comes to an SJT. With that in mind however, it is important to recognise that the proportion modal consensus scores still predicted supervisor ratings among the entire sample. Again, this may be for various reasons. First, nurses and AHPs made up the majority of the sample and their results will have, therefore, dominated the findings. Second, proportion modal consensus scores were 'continuous,' real positive numbers, and therefore contained more information than dichotomous scores (0,1). A drawback of using proportion modal consensus scores however would be that they are harder to model and implement in practice. An alternative theory as to why proportion modal consensus scores had predictive validity among the overall sample is that whilst mental health roles differ, there are core attributes required of all professionals working in mental health services. The latter concept aligns to the findings of the earlier systematic review (see chapter 4) and qualitative study (see chapter 5). That is, whilst there are discipline-specific skills and competencies required of mental health practitioners, there are also generic attributes required of all professionals working in mental health services, such as the ability to remain secure, stable, calm and confident in times of distress (Ljungberg et al., 2015).

An additional finding of the study was that self-rated *agreeableness* was associated with SJT scores across all scoring approaches, except raw SME binary scoring. The

positive association observed between proportion modal consensus SJT scores and supervisor ratings of perceived *professionalism* and *effectiveness* across all participants may have resulted from the SJT tapping into core, generic, non-academic attributes that are desired of all professionals working in mental health services (e.g., agreeableness). Nevertheless, when looking at a subset of the sample that had a similar job role (i.e. nurses and AHPs), the predictive validity of the SJT scores increased substantially and resulted in validity coefficients similar to face-to-face selection processes (McDaniel et al., 1994).

6.5.4 Strengths and limitations

As noted above, this study resulted in the production of validity evidence for an SJT that was developed to be used in mental health services. Patients, carers, and staff members from a range of disciplines had involvement in the development of the SJT through their participation in focus groups and interviews, as well as their review and initial scoring of the pilot SJT items; the latter facilitated the development of the SME-based scoring key. The involvement of patients and carers in the development and scoring of items could be considered a strength of this study; it is possible, however, that the involvement of patients and carers lessened the relationship between SME-based SJT scores and ratings of job performance due to patients and carers contrasting beliefs, regarding professionalism, in comparison to members of staff. This matter is discussed further in the next chapter. Another potential reason that no relationships were observed between SME SJT scores and ratings of job performance could be that the SME group included staff with varying levels of experience. It is common to include more senior staff in SME panels as, for example, they would be more likely to understand the most appropriate course of action in a particular situation. Should a further panel be developed by the researcher, efforts would be made to recruit staff with more experience and/or expertise in the field.

The SJT was initially developed with a view to using it across all mental health professional disciplines and this provided the rationale for including a variety of

professional groups within the development of the SJT item content. As the results demonstrate, the SJT is more suited to mental health nurses and AHPs with the SJT possessing greater predictive validity for this sample. Had this finding been known in advance of the study, efforts may have been made to focus on these professional groups when developing the SJT. That being said, it is common for staff to be involved in multidisciplinary practice in mental health services. Thus, a relative 'outsider' perspective provided by other professions being involved may have resulted in more comprehensive SJT content that takes account of the bigger picture. This would include the need for multi-disciplinary teamwork in the delivery of mental health care. Indeed, a finding of the aforementioned qualitative study was that staff must understand problems from a wider perspective, taking account of the whole picture (see chapter 5, section 5.5.2). A further strength of this study was that the SJT was piloted in an NHSFT that covers a large geographical area in the Northeast of England. There were limited participants with a BAME background however, which meant that it was not possible to assess for differences in SJT scores according to ethnicity.

A limitation of the study was that participants were already employed by the NHS site where the research took place, and the findings may have differed had an applicant sample been utilised. First, for example, staff working in the role will have gained additional wisdom and experience from doing the job and may therefore have performed better, on average, on the SJT than applicants as a whole. Second, incumbents may be less inclined to distort their responses in comparison to applicants, with the latter potentially revising their response in order to be viewed more favourably (Whetzel and McDaniel, 2009). As the SJT has been developed for the selection of staff, a further pilot on an applicant sample is recommended. Nevertheless, the research findings observed a *strong* positive association between consensus SJT scores and ratings of job performance; given the large validity coefficients observed, it is anticipated that the association would remain in an applicant sample also. It is important to note that the SJT in question was designed in a

knowledge-based format, which is more difficult to fake (McDaniel and Nguyen, 2001; Nguyen et al., 2005). That is, the test-taker either knew what the best responses were, or they didn't. It is therefore likely, given the current SJT's format, that the scores would differ less between incumbents and applicants than had had the questions in the SJT been of a behavioural-tendency format instead (i.e., 'what <u>would</u> you do?').

The use of an incumbent sample also resulted in a restriction of range amongst the data. Had individuals not suitable for working in mental health services participated in the research, there may have been more variation in SJT scores, which in turn could influence their reliability and validity (McManus et al., 2013). This range restriction is an almost universal challenge with validation studies for personnel selection assessments (Davison et al., 2016). Moreover, in this case both the selection assessment and the outcomes could only be observed. This precluded the usual adjustments that can be made in these contexts for direct and indirect restriction of range (Alexander, 1990; Schmit and Ryan, 1992; Thorndike, 1947). Nevertheless, despite this, as statistically significant and substantively meaningful correlations were observed in the current study it can be assumed that the adjusted correlations would have been even larger. Thus, evidence of the validity of the SJT scores could still be demonstrated despite this limitation.

Whilst ratings were typically requested one week following completion of the SJT, a temporary hold was placed on recruitment for the study because of the Covid-19 pandemic. Also, prior to closing the study, colleagues were sent one final reminder to submit feedback. Thus, in some instances, there was a much longer time lapse between individuals completing the SJT and individuals receiving feedback from their colleagues and supervisors. The longest duration from completing the SJT to receiving feedback was 210 days (i.e., 30 weeks). In theory this may have attenuated the relationship between the SJT scores and supervisor ratings; however, if knowledge of

professionalism is fairly stable in experienced professionals, it is less likely that this will have attenuated the relationship.

6.5.5   Implications for policy and practice

This thesis commenced by introducing the reader to the failings of Winterbourne View care home. Sadly, there has many further instances since where patients have failed to receive adequate care from staff that are paid to look after them. An inquiry is currently underway to investigate the deaths of 1,500 people that were being cared for by an NHS mental health trust; these individuals had '*unexpected, unexplained or self-inflicted'* deaths between 2000 and 2020 whilst they were either an inpatient or within three months of being discharged from the mental health trust (Campbell, 2022). Leading on the inquiry, Dr Geraldine Strathdee identified three recurring failings at the trust, including: "*serious concerns about patients' physical, mental and sexual safety while on a ward …; big differences in the quality of care patients received, both in staff attitudes and in the use of effective treatments; and patients and their families being given too little information about their treatment, likely length of stay and chances of recovery*" (Campbell, 2022). These findings highlight further, the necessity for the mental health workforce to improve in their delivery of safe, effective, and person-centred care. The current SJT would facilitate this endeavour by helping select staff with the appropriate knowledge to deliver mental health care. In this respect it is anticipated that the SJT would at least flag applicants that provide 'unusual' responses to the items in the assessment. Such candidates could either be rejected at a fairly early stage of the selection process, especially if there were other negative indicators relating to the individual, or alternatively 'atypical' responses to SJT items could be flagged and then explored with the candidate at a face-to-face interview. Either way, it is hoped that this would lead to less suitable candidates not being employed in mental health services.

A model developed by the Centre for Mental Health (an independent UK mental health charity), alongside NHS colleagues, forecast that in England up to 10 million people will

need either new or additional mental health care as a result of the Covid-19 pandemic, equating to an increase of almost 20% of the population (O'Shea, 2020; 2021); the model estimates that a minimum of 10 million additional people will be required to deliver this support. The current SJT could provide individuals with a realistic preview of the job, which may in turn increase their desire to work in the field. Specifically, if prospective applicants saw challenging situations portrayed, but were also shown skilled ways of responding and resolving these, it could improve the attractiveness of these roles. The workforce shortages in mental health care have very real consequences. Previous research conducted by the University of Manchester reported that deaths of patients under observation tended to occur for four reasons, which included when there were staff shortages (National Confidential Inquiry into Suicide and Homicide by People with Mental Illness, 2015). More recently, following an inspection of Norfolk and Suffolk NHSFT, the Care Quality Commission commented that '*The trust did not consistently maintain safe staffing levels or ensure there were enough suitably qualified staff to meet the needs of people using services. We found this was impacting on the level of safety for staff and patients. It also impacted on governance within teams, multidisciplinary team effectiveness and patient safety*.' (Care Quality Commission, 2022). Given the current staffing crisis, there is an even greater need for efficient, scalable, and effective values-based recruitment in the NHS and the mental health sector, for which is something this thesis provides evidence.

Student nurses' person-centred values have been observed to diminish as they gain experience in an adult nursing programme (Groothuizen, 2020). It is possible that the relationship between work experience and the worsening of professional values may also apply to staff employed in mental health services. SJTs have been used for training and development purposes previously within the fields of medicine (Goss et al., 2017) and pharmacy (Patterson et al., 2019); using the SJT in this manner could help mitigate against a decline in mental health professionals values and subsequent clinical practice. That is, the content could serve as a reminder to staff about the desirable

behaviours and related values expected to be demonstrated in that clinical setting. Additional content from the SJT item pool, which was developed during this project, could therefore be utilised for the training and development of staff working in mental health services, which would ensure that efforts are made not just to select staff with the appropriate knowledge and behaviours for mental health practice, but that staff continue to deliver good practice throughout the course of their mental health careers.

6.5.5.1 *Implementation of the final SJT*

Based on the findings of the pilot study, the following approach is recommended when using the SJT for personnel selection in mental health services. First, until the scoring key has been adapted for other professional disciplines, it is recommended that the SJT be used for the selection of nurses and AHPs only. Second, whilst both dichotomous and proportion modal consensus scoring approaches were similarly effective in predicting job performance, the former (i.e., dichotomous modal consensus scoring) is advised as this is a more pragmatic approach and would be easier to model using Item Response Theory. Third, it is recommended that all items be used in the final assessment, however only the validated items should be scored. The use of non-scored items will provide camouflage and hopefully minimise the risk of coaching effects on test-takers scores as well as reduce the risk of test-takers learning the content of the SJT from others. Fourth, where individuals are required to sit the SJT on more than one occasion, it is recommended they are given alternate forms to minimise any test-retest effects.

Individuals' scores on the SJT can be compared with the normative reference scores generated during the validation study. A mechanism has already been developed where test-takers will receive a Band from 1 to 6, dependent on what percentile their standardised SJT score falls within; Band 1 equates to the top 5% of test takers, whereas those that score lowest on the assessment would obtain a Band 6. Until further validity evidence is obtained in an applicant sample, it is advised that the SJT be used to support the interview process as opposed to being used to screen out

applicants. Feedback, like that presented in Figure 6.11, can then be used to help

guide, but not dictate decisions made by the interview panel.

Candidate feedback:

<u>**SJT score:**</u>

Band 5 – This individual obtained a very low score compared to the other clinicians in the original study sample.

This score would place them in the lowest 25%, but above the bottom 5% of test-takers.
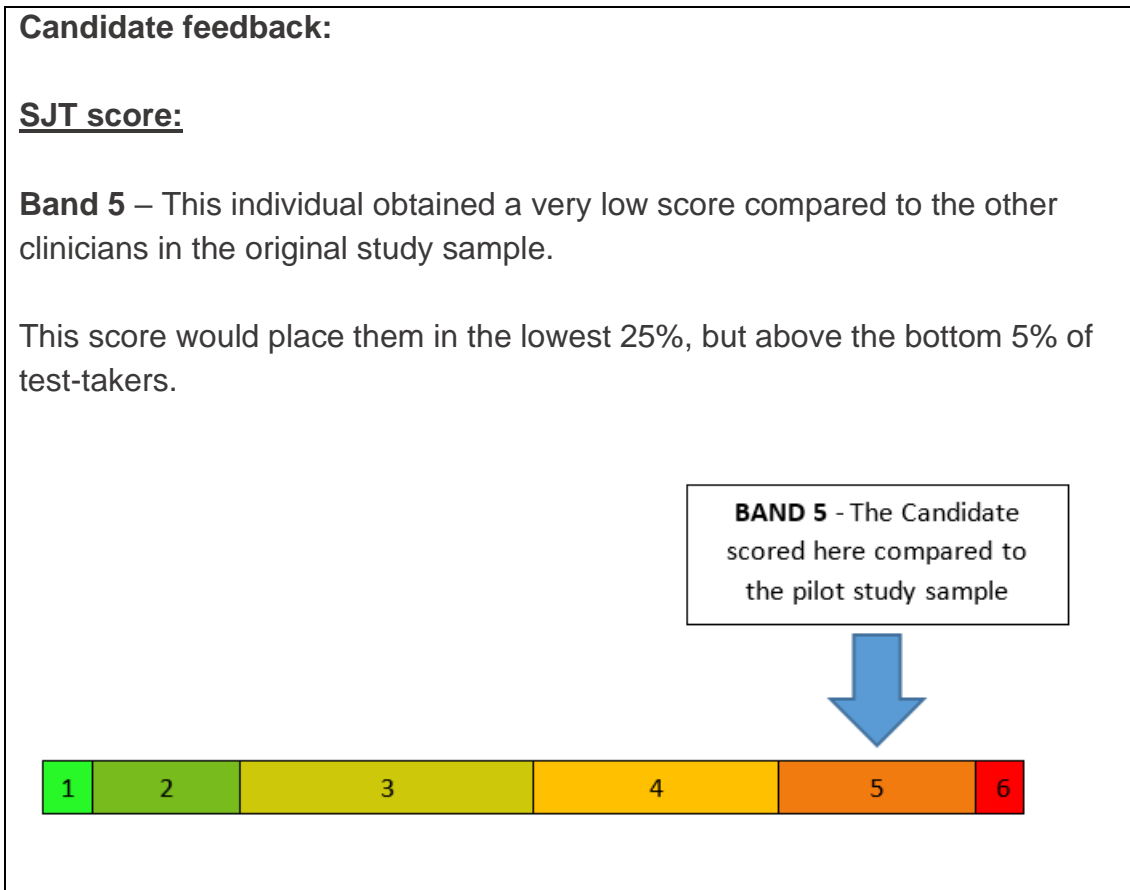
Figure 6.11 Example feedback for the interview panel

To support the personnel selection process in mental health services, a system has

been developed to flag SJT items that the test-taker provides unusual responses too.

The latter is based on two conditions being met: first, the test-taker must provide a

response that differs to that provided by the SME panel, which included patients and

carers; second, the test-taker's response must have been chosen by less than 30% of

their peers (i.e., other nurses and AHPs). Flagged items can then be presented to the

interview panel so that the panel can discuss these with the applicant during the

interview itself. The opportunity to provide a rationale for choosing particular

responses was a theme that cropped up within the acceptability feedback provided by participants in the research study. That is, participants remarked that the multiple-choice format of the SJT was somewhat limiting, and they would have preferred the opportunity to justify and contextualise their responses. For example, one participant noted the following:

"*I think if you were using them as value based questions even if the person answered what might be deemed wrong/incorrect I think then the answers should be explored. I don't think all the questions had a right or wrong answer and would depend on the variances in the situation*."
(nurse)

The ability to discuss items at interview should hopefully enhance people's experience of sitting the assessment, whilst also providing the interview panel further insight regarding an applicant's reasoning and professional judgement.

6.5.6   Recommendations for further research

A finding of the current study was that SJT scores obtained for appropriateness items were not related to the SJT scores obtained for importance items. This provides evidence to support that proposed by Tiffin and Carter (2015); that is, that the different question formats enable the evaluation of different traits. Future studies could seek to replicate this finding in different healthcare disciplines and samples.

The results of the current study provide evidence that the SJT scores predict supervisor ratings of perceived *professionalism* and *effectiveness* in a sample of nurses and AHPs that currently work in a mental health and learning disability NHSFT. As the sample incorporated current employees, as opposed to applicants for the role, it would be helpful to undertake 'post-marketing surveillance' of the SJT once it is implemented in personnel selection. This would allow the collection of further validity evidence when used with an applicant sample. For example, researchers may choose to evaluate the predictive validity of the SJT scores in relation to scores obtained during interviews. A

longitudinal study would also provide further evidence regarding the SJT scores predictive validity. For future SJT validation studies, criterion could include future ratings of job performance, whether the SJT scores predicted how long individuals retained their roles, or whether applicants received complaints or faced later disciplinary action. In this respect it is worthy to note that SJT scores have previously been found to predict future disciplinary action among a national cohort of UK-based medical students (Tiffin et al., 2022).

The analyses observed that the SJT possessed low internal consistency-reliability. However, this is common for some SJTs used in personnel selection due to their 'essential unidimensional' nature. In these situations, there may be one main factor underlying response patterns but numerous additional 'mini-factors' with few items loading on each. The current SJT was not designed to assess one specific construct and was instead created using a traditional approach to SJT development, aiming to assess procedural knowledge on a range of skills and abilities. This undoubtedly resulted in construct heterogeneity at the item level and internal consistency estimates are therefore not the most appropriate means of assessing the SJT's reliability (Motowidlo et al., 1990; Whetzel and McDaniel, 2009). A more suitable approach to assessing the reliability of SJTs that are multidimensional is to evaluate (1) their test-retest reliability (i.e. the consistency of the test over time) or alternatively, (2) their parallel forms reliability (i.e. the equivalence across tests that are believed to assess the same construct; Lievens et al., 2008; Mcdaniel et al., 2007; O'Connell et al., 2007; Whetzel et al., 2020). It would therefore be useful to assess the reliability of the SJT using either of these methods as opposed to putting too much emphasis on the internal consistency reliability estimates, which are known to underestimate the reliability of SJT scores (Catano et al., 2012; Whetzel and McDaniel, 2009). A protocol for a test-retest reliability study has already been drafted by the author. Future studies could also be conducted within a different UK or global geography to better assess for potential item

or test bias for different, underrepresented groups, particularly those identifying as of minority ethnicity.

The results demonstrated that the current SJT scores had predictive validity across nurses and AHPs working in mental health services. Nurses and AHPs formed the largest number of participants in the study however and the consensus scoring key was therefore generated using a larger sample of participants. A further pilot could be conducted with a larger sample of psychiatrists and/or psychologists to determine whether the SJT consensus scores subsequently predicted ratings of perceived *professionalism* and *effectiveness* for these disciplines. Thus, it may be possible, without altering content dramatically, to develop bespoke consensus scoring systems for other professional groups working in mental health services.

6.5.7 Conclusion

This study, at the time of submission of this thesis, was the first to develop and pilot an SJT for the purpose of selection in mental health services. The results show that the SJT scores obtained by nurses and AHPs working in mental health services predicted ratings of job performance, namely perceived *professionalism* and *effectiveness*. This predictive validity is important as it supports the SJT as offering recruiting organisations a cost-effective, scalable, objective measure of professional judgement that is not yet routinely available. In the absence of further validity evidence, it is anticipated that the SJT would be useful to enhance and complement existing recruitment processes, which are likely to also involve face-to-face interviews.

Chapter 7: Overarching Discussion

The overall aim of this doctoral work was to develop and validate a Situational Judgement Test (SJT) that would assess staff members' knowledge of professionalism for a mental health services context. To fulfil this aim, four objectives were set. These objectives consisted of 1) to develop an operational definition of 'professionalism' suitable for a mental health services context, 2) to develop a pool of SJT items that could be used to evaluate individuals' knowledge of professionalism for a mental health services context, 3) to develop and validate an SJT for personnel selection in mental health services, and 4) to develop a bespoke workplace behaviours rating tool.

7.1   Methods used to meet the thesis objectives

Adopting a mixed methods approach, a programme of work was undertaken to meet the project objectives (see chapter 3). The thesis commenced by presenting the context for this doctoral work (chapter 1). Subsequently, a narrative review was presented regarding the application of SJTs for personnel selection (chapter 2). Following the background and methods section, the author reported on a rapid, systematic review that was completed to derive an operational definition of professionalism for a mental health services context (chapter 4). The review was followed up by a qualitative study, which built on the findings of the systematic review and facilitated the generation of item content for an SJT (chapter 5). Finally, chapter 6 reported on the development and validation of the SJT in a mental health services setting.

It's not the intention of the author to reiterate every thesis section here, however the key findings of the project are drawn upon within the remainder of this chapter. The strengths and limitations of the thesis are subsequently highlighted prior to recommendations being made for future research. Implications for policy and practice have been integrated throughout the discussion as opposed to being presented in a separate section at the end of the chapter.

7.2 Key findings

Chapter 1 sets out the context for this doctoral work and in turn, provided a rationale for developing an SJT that would help select staff to work in mental health services.

7.2.1 Rationale for developing an SJT

As noted previously, SJT-format tests, in this context, can assess knowledge of interpersonal attributes that are not typically evaluated by more traditional assessment methods (e.g., cognitive tests; Lievens, 2013; Lievens and Sackett, 2012). This means that they may have incremental validity when predicting performance in jobs that involve emotional labour (Clevenger et al., 2001; Mcdaniel et al., 2007). Due to the nature of their illness, patients using mental health services are potentially more vulnerable to being victims of unprofessional practice or misconduct; it is therefore paramount that job applicants are subjected to valid assessments during the selection process that would evaluate their suitability for the role. Therefore, the aim of this doctoral work was to develop and validate an SJT that would assess individuals' knowledge of professionalism for a mental health services context.

Chapter 1 highlights various selection approaches that can be used for values-based recruitment, which includes structured interviews. It is acknowledged by the author that scores obtained during interviews predict job performance, yet they are resource intensive and can be costly to develop and implement (Schmidt and Hunter, 1998). SJTs on the other hand can be computerised, delivered online, and machine marked, thus providing a valid, yet cost-effective approach to personnel selection.

7.2.1.1 *Utilising a traditional, as opposed to a construct-driven approach to SJT development*

Having chosen to develop an SJT that would assess ones' knowledge of professionalism for mental health services, the next decision was whether to develop an SJT using a construct-driven or traditional approach. Construct-driven SJTs have been developed to assess various unidimensional constructs (see Guenole et al., 2017). However, given the complexity and ambiguity of mental health practice, one particular trait wouldn't

244

always, necessarily, result in the best behavioural response to every situation. In addition, traditional format SJTs may have high levels of predictive validity where there has been greater exposure to 'real-life work situations' (Tiffin et al., 2020), as is the case for mental health clinicians that have already worked in healthcare placements during their professional training. The current thesis therefore reports on the development of a traditional format SJT as it was hoped that this would assess individuals' procedural knowledge of professionalism in a range of work-relevant settings and contexts. Construct-driven SJT scores are more unidimensional that traditional SJT scores because they focus on different levels of the same targeted trait (Lievens, 2017a). Unlike construct-driven SJTs, traditional SJTs have no well defined 'measurement model' and it is not always clear what is actually being assessed (Tiffin et al., 2020); for this reason, factor analysis was performed, but the output was not used to help reduce items on the SJT. While construct-driven SJTs have their advantages, they may be more prone to faking and social desirability bias, having more in common with personality measures than traditional SJTs (Peeters and Lievens, 2005). This was deemed of critical importance given that the current SJT was being developed for use in personnel selection, where test-takers are more likely to alter their responses in order to provide a more favourable impression.

Recent evidence suggests that traditional SJTs may inadvertently assess prosocial personality traits, such as agreeableness (Tiffin et al., 2020). For example, a meta-analysis reported a mean population correlation coefficient of .25 between agreeableness and SJT scores; the correlation between agreeableness and SJT score was higher for behavioural-tendency items, as opposed to a knowledge-based format (.37 vs .19; Mcdaniel et al., 2007). SJT scores obtained during the pilot study were observed to correlate with self-ratings of agreeableness for most of the scoring approaches. This observation is likely due to the SJT assessing test-takers implicit trait policies (ITPs); that is, the beliefs they hold regarding how effective various levels of trait expression are (i.e. agreeableness in this instance; Motowidlo et al., 2006a;

2006b). Motowidlo et al. (2018) proposed that the procedural knowledge scores obtained on interpersonally-orientated SJTs are correlated with agreeableness because the procedural knowledge measured by these SJTs includes knowledge about the effectiveness of prosocial actions in interpersonal situations; the authors suggest that agreeable individuals obtain higher SJT scores, because agreeable people are more likely to possess this knowledge. The fact that individuals' ITPs regarding agreeableness appear to be assessed by the SJT partially explain the relationship observed between SJT scores and supervisor ratings of job performance. Afterall, agreeableness is a prosocial trait that understandably would facilitate a mental health practitioners' effectiveness at work. It is worth adding at this point that the SJT scores provided incremental validity over self-rated agreeableness in predicting supervisor ratings of professionalism. Whilst the SJT may partially assess ITPs regarding agreeableness, the latter finding indicates that the SJT is assessing something else also, which is possibly their procedural knowledge or other ITPs that are associated with high levels of professionalism.

When developing SJTs, equating is important to ensure fairness across alternate test forms. However, the lack of a well-defined 'measurement model' with traditional SJTs can make the equating of SJT forms particularly difficult (Tiffin et al., 2020). Therefore, as part of this doctoral work, a blueprinting process was utilised. This process included obtaining the views of Subject Matter Experts (SMEs) regarding professionalism in mental health services and then subsequently mapping these professional domains as equally as possible across the alternate SJT forms.

7.2.2 Conceptualising and defining professionalism for a mental health services context – the basis of a dynamic social contract

Prior to the development of any assessment, an operational definition of the criterion to be assessed is crucial. Whilst many efforts have been paid to explore professionalism in medicine and medical education (Finn et al., 2010; Hodges et al., 2011), there are nuances in mental health services that make them different to other

healthcare fields. The review reported on in chapter 4 was therefore undertaken to obtain an enhanced conceptualisation of professionalism, specifically for a mental health services context. Based on the findings of the review, professionalism was conceptualised on two levels: on a societal level, professionalism was described as the basis of a 'dynamic social contract;' and on an individual level, professionalism was described as 'being representatives of the profession.' The operational definitions derived during the review are provided, in full, in Table 7.1. Findings from the review have begun to have impact with early citations being logged (Cheesmond, 2020; Fitzgerald, 2019; Hale and Wright, 2021; Oates et al., 2020; Rioli et al., 2020; Song et al., 2021; Yahyavi et al., 2021).

Table 7.1 Operational definitions of professionalism for mental health services

At a societal level –	*Professionalism forms the basis of a dynamic social contract between professions and society. This contract (which can have both tacit and explicit elements) specifies that society will remunerate the members and permit the profession to self-regulate on the understanding that the profession use their skills for patient and public good.*
At an individual level –	*On an individual level, professionalism can be conceptualised as a latent trait, composed of elements of intrapersonal, interpersonal, and working professionalism. This trait may only be observed through manifest behaviours in certain situations. Such behaviours will be in keeping with society's expectations and demonstrate a commitment to ethical practice, cultural-sensitivity, self-awareness and reflection and self-discipline.*

Noting that professionalism forms the basis of a dynamic social contract between professions and society, the review findings highlighted the need for professions to regularly renegotiate this contract, incorporating both patients' and carers' views. It is suggested that patients also have a part to play in the social contract, by having a 'shared responsibility' for their own health and wellbeing (Bhugra, 2008b; Cruess, 2006). Potential examples of this that have been provided previously in this thesis

include patients turning up to their appointments on time and being polite to professionals. However, when a patient is having an acute episode of mental illness, especially if they're enduring psychosis, is it fair to expect a patient to uphold these behaviours at all times?

7.2.3   The added value of listening to patients and carers

Whilst it is suggested that patients need to be involved in renegotiating the social contract, and in turn establishing the expectations held of a profession, the systematic review observed a lack of patient presence in the literature on professionalism in mental health services. It has been argued by some professional groups that it is professionals that have the knowledge to make informed decisions regarding what services should be delivered, and to what standard they should be met (Randall and Kindiak, 2008). However, prior research has demonstrated that patients and practitioners do not always place the same regard on certain professional behaviours (Green et al., 2009).

Acknowledging the importance of patient and carer perspectives, a subsequent qualitative study, reported on in chapter 5, sought to explore the views of patients, carers, and professionals on the topic; specifically, the qualitative study aimed to determine how the lived experience of participants aligned with medically defined, generic professionalism standards. Chapter 5 reported that in comparison to a conventional medical definition of professionalism, additional themes and differing emphases were observed for mental health and learning disability services. One specific theme identified during the study was the importance of professionals working with carers. Carers reported being ignored by professionals who wouldn't even talk to them because '*the patient had not given consent for the carer to be involved*.' The limited involvement of carers in healthcare delivery remains present to this day. Indeed, as reported earlier in this thesis, a current inquiry investigating deaths at a mental health trust has reported that patients and their families were given too little information regarding their treatment, likely length of stay in hospital and chances of

248

recovery (Campbell, 2022). This thesis explicitly adds 'Working with Carers' as a separate entity to the Professional Attributes Framework, noting that this is a key characteristic required of all professionals working in mental health services. Used alongside other guidance such as the Triangle of Care (Worthington et al., 2013), this professional attribute being made explicit in the professional attributes framework could enhance the development and training of professionals and therefore improve the delivery of clinical practice.

7.2.4   Involving patients and carers in SJT development and scoring

The SJT content developed during this project had input from patients and carers as well as professional staff. The involvement of patients and carers aligns with the idea that all parties have a part to play in deciding what services professions should deliver and what expectations should be placed on professionals. As mentioned earlier however, there is a discrepancy amongst what professionals and patients value when it comes to healthcare delivery. This is a key observation when one considers the findings reported on in chapter 6. Chapter 6 observed a relationship between consensus SJT scores, and ratings of perceived *professionalism* and *effectiveness* that were provided by managers and supervisors. There is a concern however that whilst supervisors value the professionalism of individuals that score highly on the SJT using consensus scoring approaches, patients may not. The consensus scoring approaches essentially provide test-takers a score based on how well their SJT responses align with their peers' responses (i.e., the responses of other professionals in their discipline). During the pilot study, an additional scoring key was also developed with an SME group, which incorporated the views of patients and carers alongside professionals of varying disciplines, thus providing a multi-disciplinary perspective. The scoring key was based on *agreement* among the SME panel; however, it is possible that the views of professionals, patients and carers are incompatible. Indeed, this could have been the underlying reason for the lack of a relationship observed between ratings of job performance and SJT scores whilst using the SME scoring key. The idea that patients,

carers, and staff members beliefs regarding professional behaviours differ is further supported by the findings of this thesis; for example, the qualitative study observed a patient and staff member to have strong contrasting beliefs regarding documentation.

7.2.5 Role-modelling and organisational culture in mental health services

Another observation made during the qualitative study was the presence of negative role modelling, especially with regards to professionals mentoring students. At this point, the notion of the Hidden Curriculum warrants further consideration. The Hidden Curriculum (HC) is defined as, "*the tacit, implied, unwritten, unofficial, and often unintended behaviours, lessons, values, and perspectives that students learn during their education*" (as cited in Finn et al., 2022). While originally defined within the context of education, the hidden curriculum is equally as relevant within workplace settings. Think of it as 'the way things are done around here.' Using hidden curriculum as a term is perhaps a misnomer, it is very much about organisational culture in this context, but in the seminal work it is about the educational culture (Hafferty, 1998). Examples of positive and negative things that could be learned or observed via the hidden curriculum may include the way one communicates with colleagues, students, or patients. The same applies in clinical practice. For example, using terms like the 'revolving door patient' during multi-disciplinary meetings to describe patients that are frequently admitted to mental health services, which includes individuals with severe mental illness (Botha et al., 2010; Fakhoury and Priebe, 2002; Weiden and Glazer, 1997), may reduce a student's empathy and increase stigma towards this patient group. On the other hand, positive role-models not only have a positive impact on students' development, but are also believed to influence students' future career choices (Ambrozy et al., 1997; Murakami et al., 2009). Role models have even been suggested to be more influential than the specialty itself when students are choosing future career options (Archdall et al., 2013). The ways that healthcare professionals tacitly learn about how to 'behave' is especially relevant to interpreting and using SJTs for both personnel selection, and possibly development. For example, one study noted

that student nurses scores on an SJT assessing person-centred values actually declined with increased exposure to the healthcare workplace (Groothuizen, 2020).

7.2.5.1 *Humour*

The qualitative study findings highlighted instances where gallows humour was used, and staff would ridicule patients' behaviour. Whilst it is inappropriate for this to occur in any setting, doing so in front of students suggests that it is okay to make fun of patients. When students have previously been asked *how* they know their participation in humour was acceptable, one student responded with the comment "*If the ice has been broken by someone of a higher level, then it is okay to say anything*", to which many other students nodded their heads (Wear et al., 2006, p.458). Wear et al. reported that students were expected to laugh, and certainly not object, if someone of a higher rank used derogatory humour, even if the student did not appreciate or find the humour funny.

The use of humour has been contested previously in both healthcare settings (Dueñas et al., 2020) and undergraduate curricula (Finn et al., 2010). It has been reported that there are particular groups of patients that comments and jokes would usually be targeted at, including patients whose health problems are believed to be brought on by themselves, such as those that use alcohol, drugs, or food to excess (Wear et al., 2006). This may be influenced by students' and practitioners' difficulty to empathise with this patient group (i.e., individuals with illnesses that are perceived to be self-inflicted). Patients with mental health illnesses are also perceived as more vulnerable to being the target of humour, such as name calling, imitating unusual behaviours, and making fun of their stories and beliefs (Wear et al., 2006).

7.2.5.2 *'Good' and 'bad' patients*

Various illnesses and symptoms are suggested to make patients more or less liked by practitioners; for instance, many authors report that individuals with psychiatric conditions produce negative evaluations from staff (see Kelly and May, 1982 for a

review). Kelly and May observed various characteristics in the literature that are attributed to 'good' and 'bad' patients. For example, 'good' patients make drastic recoveries, or are seen as understanding, amusing, optimistic, cheerful, and grateful, whereas 'bad' patients are unappreciative, will not help themselves, or refuse to accept that there is anything wrong with them.

When patients do not always adhere to treatment regimes, even if it is just because they forget to take their medication on a morning, they may be labelled non-compliant. 'Difficult patients' (AKA bad patients) are often targets of derogatory humour; medical students have described these patients as being 'non compliant', 'demanding', 'aggressive', 'talkative', 'disrespectful', 'persistent' and/or 'periodic' (Wear et al., 2006). Koekkoek et al. (2006) report on four dimensions of difficult behaviours among patients receiving mental health care, which include 'being withdrawn and hard to reach', 'demanding and claiming', 'attention seeking and manipulating', and 'aggressive and dangerous'. The authors also hypothesise that there are three subgroups of 'difficult patients', these being: care avoiders, which includes severely psychotic patients that do not consider themselves unwell and view mental health care as an interference; care seekers, which includes patients with chronic mental illness that struggle to maintain a steady relationship with caregivers; and care claimers, which includes patients that do not need long-term care, but need some short term benefit from mental health care, such as housing.

It is suggested that patients are sometimes labelled as problematic and difficult because they experience problems that are difficult to resolve by the psychiatric system (Koekkoek et al., 2006). Weiner and colleagues (1988) have also observed that patients with behavioural/mental problems are perceived to be more responsible for their condition, and are liked less, evoke more anger, and less pity, than people with other conditions that are perceived to be uncontrollable. Professionals may also

experience annoyance or anger if their competence is questioned (Breeze and Repper, 1998), or when their authority is challenged (May and Kelly, 1982).

7.2.5.3 *Ethical erosion*

The concept of ethical erosion has been described as '*where empathy and sympathy declines with increasing clinical experience*' (Stratta et al., 2016, p.286). This phenomena has mostly been discussed in relation to medical education (Bellini and Shea, 2005; Bellini et al., 2002; Diseker and Michielutte, 1981; Hojat et al., 2004; Neumann et al., 2011), however there are similarities in the nursing literature where a decline in altruism (Miers et al., 2007) and values following professional registration (Maben, 2003) are reported. Values have been observed to be compromised and re-prioritised during nurse training also, as students are introduced to the clinical practice environment (Groothuizen, 2020). These findings indicate the importance of role-modelling during both training and clinical practice. Positive role modelling has been observed to improve medical students' empathy skills (Tavakol et al., 2012), however, a lack of empathy can also be learned from senior role models (Feudtner et al., 1994).

We must pay attention to the organisational culture of mental health services as this impacts the quality and safety of care, as well as staff wellbeing (Looi et al., 2022). One study found also that the organisational culture and climate of community mental health teams significantly predicted improvements in perceived mental and physical health status among patients (Morris et al., 2007).

Recognising the impact that organisational culture can have on one's healthcare practice, it could be questioned whether assessing one's professionalism, or knowledge of professionalism, is actually going to help at all. As mentioned in chapter 2 (section 2.3.1.4), a decision was made to develop a knowledge-based format SJT, as opposed to using behavioural tendency instructions; the former is correlated more highly with cognitive ability, but is less prone to faking and was thus deemed the most appropriate format for personnel selection. Afterall, if an individual does not know

what is the most appropriate behaviour in a given situation, it is unlikely that they'll behave accordingly; thus, knowing what to do is a pre-requisite to behaving appropriately on the job. The mental health system is highly complex and it should be considered whether assessing an individual's knowledge is going to influence healthcare practice at all, especially when there are various organisational pressures and 'values attrition' at play. Whilst it is worthy to consider the sustainability of professionalism, prior studies have found SJT scores, generally, validly predict aspects of interpersonal performance relevant to clinical practice (Webster et al., 2020). Communication skills can serve as a protective factor against burnout (Pérez-Fuentes, 2019). Therefore, assessing one's professionalism, which incorporates a professional's ability to communicate effectively, may in turn lead to a reduction in fitness to practice situations (Coffey, 2020). Furthermore, by employing staff with an increased knowledge of professionalism, it is more likely that you will have an 'emergent' culture of professionalism at a system level. Professionalism requires a staff member to challenge the unprofessionalism of others; therefore, as staff with higher levels of professionalism begin to fill roles, it is likely that unprofessional behaviours will diminish at a system level too.

7.2.6 Additional uses for the SJT content

Whilst the current SJT was developed with personnel selection in mind, SJT content developed during the project could also be utilised for scenario-based learning, as has been done previously (Cox et al., 2017). Such training could counter the effect of negative-role-modelling and help maintain ethical standards in mental health services. Monrouxe et al. (2011) advocate that more sense-making opportunities are made available to students under the supervision of clinical educators. The SJT content could assist this process as it would provide material that would hopefully facilitate open discussion among students and staff. This is especially important given that, as noted previously, students have found the school environment inconducive to discussing ethical concerns (Satterwhite et al., 2000). Application of the SJT could also set the

tone for current and potential employees regarding the culture of the organisation. That is, use of the SJT in the personnel selection process could translate to staff that the organisation places deep regard on the importance of professionalism within the healthcare environment and that the organisation pursues staff that will maintain the highest levels of professional behaviour.

7.2.7   Professionalism in the nursing and allied health professions

Had the author been able to anticipate the findings of this doctoral work in advance of the project, the methods utilised may have differed to that reported on in this thesis. First, it is likely that greater attention would have been paid to the specific professional guidelines adopted by the nursing and allied health professions. For instance, there is already documentation on professionalism published by the Nursing and Midwifery Council (2017), and the Health and Care Professions Council (2014). Guidelines and codes of conduct used by these professional groups are tangible expressions of professionalism (Sox, 2007). Nonetheless, the primary intention of the author was to look at professional behaviours specific to mental health services, which may differ to the behaviours expected of a practitioner in other healthcare specialties. Furthermore, the author initially set out to develop an SJT that could be employed for all mental health practitioners, regardless of the professional discipline they aligned to. However, chapter 6 found that the SJT scores possessed greater predictive validity for nurses and allied health professionals (AHPs), emphasising the different expectations of these professional groups.

7.2.8   Is it time to review the mental health workforce?

Given the differences among the professional disciplines, which includes the training that is delivered, the fact that the SJT scores possessed more validity in a nurse and AHP sample, as opposed to other professional groups, is somewhat underwhelming. As mentioned previously, nurses and AHPs typically deliver more face-to-face care than other professional groups working in this setting. However, whilst the expectations placed on each of the professions vary, the overarching objective of mental health

services is to provide the best possible care and improve patients' mental wellbeing regardless of ones' professional discipline. It is questionable whether one professional group fulfil this aim better than others. Given the expression 'too many chefs spoil the broth,' is it worth considering whether all the professions are needed? Many of the professions now fulfil similar roles; for example, some nurses prescribe medications, which would previously have been performed by psychiatrists only. Nurses also may deliver psychological therapies and staff from various professional groups may fulfil the role of being a patient's care co-ordinator. An observation of the literature review was that mental health nurses would develop skills from outside of their profession to '*make themselves more professional as nurses*' (Crawford et al., 2008, p.1060). However, it is important to consider whether we are further complicating an already complex system, and if all these 'overlapping' roles are confusing for patients? Essentially, is now a suitable time to review the mental health workforce?

George Bernard-Shaw famously said, "*all professions are conspiracies against the laity*." It is therefore important that regulators and professional bodies have transparent processes to uphold the standards and reputation of their professions (Finn et al., in press). Given that patients are the recipients of care delivered in mental health services, their views on what is helpful and what is unhelpful are extremely important. Whilst professions may argue that patients are unaware of the challenges professionals face and therefore cannot make informed decisions regarding what, and how care should be delivered, if the professions do not listen to patient's and carer's, and change their clinical practices based on patients' and carers' wishes and expectations, then it is possible that they will end up working in a 'professional hall of mirrors' (Tiffin, 2021; personal communication).

7.2.9 A potential route forward? - utilising the expertise of professional staff with lived experience of mental health services

Even if one agrees that it is only professionals that have the knowledge to make informed decisions regarding what services can and should be delivered, it cannot be

denied that patients' and carers' views are a means of establishing whether these services are delivered as intended. A means of obtaining patients' and carers' views whilst maintaining a professionals' insight, could be to explore the experience of 'professional staff' that also have lived experience of receiving mental healthcare, whether this be in a patient or carer capacity. Patient and public involvement is highly advocated in the design and conduct of research (Brett et al., 2014) and education (Towle et al., 2010), yet it is less common in the creation and delivery of healthcare services. More recently however, providers of mental health care are increasingly creating lived experience roles, which includes roles in a leadership capacity. For example, a Trust in the North-East of England has recently appointed two individuals with Lived Experience of mental healthcare into leadership roles at the core of the organisation, with the aim to '*co-create a great patient, carer and colleague experience and being a great partner'* (Lightfoot, 2022). Efforts such as these set a precedent regarding the culture of the organisation. Similarly, the use of an SJT that was co-created with patients and carers can help set the tone of the organisation, thus creating a cultural shift for the better. For example, from day 1, onlookers will perceive the organisation to take professional conduct and patients' and carers' views very seriously.

7.3   Strengths and limitations

The specific strengths and limitations for each section of the research project can be found within their respective chapters. Here however the author will evaluate the overall methodology underlying the research project.

7.3.1 Quantitative, qualitative and mixed-methods research

Research designs differ in terms of the risk of error and bias introduced into their results and some research methods provide better evidence than others when seeking answers to specific research questions; the randomized controlled trial for instance has commonly been viewed as providing the highest level of evidence over other research

methods, especially when evaluating health interventions (Evans, 2003). Quantitative

methods are not as well suited to measuring other aspects of healthcare provision

however, such as patient perceptions regarding the quality of care (Curry et al., 2009).

Both qualitative and quantitative methods were utilised for this project, as is the case

for many studies in health care research that have been published recently (Sale et al.,

2002). However, as Sale et al. (2002) note, whilst qualitative and quantitative

techniques are integrated often, it does not follow that it is always appropriate to do

so. Indeed, qualitative and quantitative purists would advocate the '*incompatibility*

thesis' (Howe, 1988); that is, that the qualitative and quantitative paradigms should

not be mixed (Johnson and Onwuegbuzie, 2004). As noted by Sale et al. (2002), when

undertaking mixed-methods it is helpful to label the phenomena that is being

examined with each approach (both qualitative and quantitative); during this project, a

qualitative study was undertaken to explore peoples' 'perceptions' of professionalism,

which in turn informed a quantitative 'measure' of professional judgement. Such an

approach is commonly adopted for psychological assessment (i.e., psychometrics).

7.3.1.1 *Measurement in the human sciences*

There have been many debates as to whether psychometric traits can be measured at

all. For example, with regards to the concept of professionalism, we cannot directly

see the amount of professionalism an individual has. This contrasts with physical

measures such as length or weight. Instead, to assess professionalism, we must first

clarify this latent trait and establish what behaviours we would expect from a person

that has a high or low level of professionalism. Indeed, this was the rationale for

conducting the qualitative study reported on in chapter 5. By clarifying the latent trait,

we can then attempt to assess this using suitable assessment tools. Of course,

individuals may disagree about what professionalism is; therefore, multiple

perspectives were sought during the current doctoral project, in addition to

undertaking a systematic review on the topic.

Norman Campbell previously claimed that measurement requires an ordering system whereby additivity can be illustrated by physical concatenation (as cited by Bond et al., 2013). That is, arbitrary units can be added together on a physical linear measurement scale, thus representing fundamental measurement. Fundamental measurement applies to entities such as weight and time. However, for some physical entities, the additive nature must be discovered and can not be measured directly. Bond et al. (2013) highlight this by using the example of density; density can be calculated using a product's mass and volume and is therefore a derived measure. Bond et al. (2013) propose that the measurement of temperature (i.e. thermometry) may be a more useful analogy for the measurement of human attributes. Among other descriptors, Bond et al. note that temperature must be estimated by observing its effects on other substances, such as alcohol or bimetallic strips. This approach is similar when one tries to measure 'latent traits.' As mentioned above, once a latent trait has been conceptualised, we can then attempt to measure this trait through the use of assessment tools. We cannot measure latent traits directly, like we can with height and weight; instead, we can only infer how much of a trait someone possesses by measuring its effect on other objects, as is the case for temperature. The current doctoral project sought to assess individuals' professionalism through the 'test scores' individuals obtained on the SJT.

7.3.1.2 *Reflexivity resumed*

As can be observed in the methodology chapter of this thesis, reflexivity was utilised increasing the validity of the research (Pillow, 2003). Having previously been an inpatient in mental health services, the author acknowledges their own biases in undertaking this work; how the data were generated and how the results were interpreted was influenced by the author's background and beliefs (Berger, 2015).

It is worthy to note that over a decade ago the author of this thesis spent a couple of years at university training alongside student nurses. One could therefore question whether this earlier experience influenced the development of the SJT item content.

As highlighted in chapter 6, the author generated item stems (i.e., scenarios) and response options using an inductive and deductive approach. It is therefore possible that the author's experience and prior learning helped generate scenarios that were more aligned to the nursing role, as opposed to the other professions. However, having other individuals review the draft SJT items during development, as was the case with the current project, helped mitigate against creating items solely suited to nurses. It must be added that the author did not complete a nurse training programme and is therefore not a registered nurse. Also, the lead item reviewer (PT) is/was an honorary consultant psychiatrist and therefore trained and worked for an alternative healthcare discipline. In addition, the author worked as an assistant psychologist immediately prior to starting the PhD.

The author did not declare that they were a former patient to participants during the research process as they did not want their position to influence the data generated. That being said, some staff knew the author as a former patient, staff member, or healthcare student, which may have influenced some of the discussions held. To minimise bias during data analysis, a second researcher helped analyse the data that were generated during the qualitative study. Despite knowing certain individuals in the organisation, the author wanted to ensure that all professionals had an equal opportunity to participate. For this reason, the study was advertised widely across the NHS site.

The author acknowledges that the design of both qualitative and quantitative research is typically influenced by a researcher's research paradigm, and that individuals may possess alternative research paradigms (Brown and Dueñas, 2020). Different paradigms may result in different methods, and as one would expect, different methods may result in a different outcome. At the start of the doctoral project, the current author identified as a pragmatist and thus relied most on methods that best fulfil the research aims and objectives. As noted throughout this thesis, the overall aim

of the doctoral project was to develop an SJT that can be used for personnel selection in mental health services. To meet the objectives of this project, both qualitative and quantitative methods were needed. As the contents of this thesis demonstrate, rigorous methods have been utilised and the resulting SJT assessment has demonstrated predictive validity in a sample of mental health professionals, namely nurses and AHPs.

The current author initially identified as a pragmatist, however on reflection the author resonates more so with a transformative worldview. As discussed in chapter 3, section 3.3.4, like pragmatism, transformative researchers' methods are not dictated by methodological assumptions (Mertens, 2017). According to the transformative paradigm however, there is a potential strength of combining both qualitative and quantitative methods; qualitative research helps obtain community perspectives during the research process, while quantitative research helps demonstrate outcomes that have credibility amongst community members and scholars (Mertens, 2007). The SJT created during this project was developed to create change in mental health services, thus aligning well to the change-orientated nature of a transformative worldview.

Recognising that certain individuals occupy a position of greater power than others, the central tenet of the transformative paradigm is that power must be addressed at each stage of the research process (Mertens, 2007). The current author is grateful to the patients and carers, as well as the staff members that participated in this research. Given their input, it feels somewhat unfair that the author is the sole beneficiary of the title obtained for these works (i.e., Dr). Recognising the limited patient presence in the literature, the author feels some regret that they did not share the patients views more widely; instead, the author could have used their own position (i.e. of privilege) to advocate for patients more so. If the author were to develop a further SJT, similar

methods would be adopted; however, more attention would be paid to obtaining the views of patients during the SJT pilot study.

7.3.2 Patient and carer involvement

The involvement of patients and carers is a strength of this thesis. A lack of patient presence was observed in the literature early in the project. However, the patient and carer voice, which is presented throughout this thesis can help inform future research and practice in mental health services. That being said, patients and carers could have had more involvement throughout the doctoral project too. For example, whilst patients and carers were involved in the development of item content for the SJT, it would've been useful to also obtain feedback from patients and carers, as well as supervisors and managers, to help validate the SJT. Afterall, professionals must be able to meet the expectations of patients, whilst at the same time working within the boundaries of the service, and adhering to legal and ethical expectations (Rankin, 2013).

Group differences according to ethnicity could not be explored in the current sample, due to the limited number of BAME participants in the pilot study. Whilst patients and carers were involved in generating scenarios for the SJT, they could have been specifically asked to provide their views regarding the fairness of the SJT items. That is, it is important that the SJT does not discriminate amongst test-takers on the basis of ethnicity, sexuality, gender or disability, among other protected equality characteristics. Whilst patient and carer views on the topic could facilitate the revision of items, where necessary, piloting the SJT in a larger sample would also help us assess group differences.

The systematic review reported on in chapter 4 also found no papers that specifically referred to learning disability services; the content of this thesis hopefully advocates for this staff group as both the qualitative and quantitative study involved staff from a range of specialties, which included learning disability services.

7.3.3 Difficulty in assessing the reliability of SJT scores

A limitation of the quantitative study was the lack of more appropriate reliability estimates being calculated. Whilst the internal consistency reliability estimates were low, this is typical for SJTs given their often-heterogeneous nature. It was not possible to assess the test-retest reliability or parallel-forms reliability of the SJT in the study reported on above, however this should be a future endeavour. The author has since drafted a research proposal for this work.

7.4   Recommendations for further research

The following recommendations build on the suggestions for future research previously proposed in chapters 4, 5 and 6. The current project explored the view of patients, carers, and staff members regarding the concept of professionalism in general. However, future research may want to explore views regarding each of the separate professions. For example, do patients have different expectations and desires when being cared for by a nurse as opposed to a psychologist? What if the patient is then receiving therapies; would the patient expect a nurse and psychologist to deliver therapies in the same manner, and if not, how would these expectations differ?

It is hoped that the SJT developed during this doctoral work would help identify failing students, or staff, even when supervisors are unprepared to do so. The phenomenon of failure to fail is common across the healthcare professions, a barrier of which includes unsatisfactory development and evaluation tools, among other factors (Yepes-Rios et al., 2016). With regards to SJTs, Tiffin et al. (2020) advised that these are generally experienced as relatively easy tests that are therefore best used as a 'screen out' assessment. It is worth noting however that harder items are less likely to obtain consensus using rational scoring because an SME panel are less likely to agree on these items. In turn, harder items will likely be removed from the pool of items during the prioritisation process. In future studies, it may be worth incorporating items with low

consensus in an assessment and using empirical scoring to see if these items differentiate more clearly between individuals of higher ability.

A decline in empathy over the course of training and clinical practice was highlighted earlier in this chapter. Using the SJT content for training and development was also proposed. It is possible that video-format SJTs would be more influential in teaching empathy than a paper and pencil format given their increased fidelity; this could be a topic of future research. Each SME panel in the current project incorporated three patients and/or carers. It would be interesting to have a distinct panel of patients and carers to help develop a separate scoring key. Of course, an appropriate criterion would need to be determined in order to evaluate how effective the selection assessment was (Patterson, 2018). For example, as alluded to above, a scoring key developed by patients and carers may result in SJT scores that do not align with the *professionalism* and *effectiveness* ratings provided by professionals' managers and supervisors; a more appropriate criterion may therefore be *effectiveness* ratings provided by patients and carers, as opposed to obtaining the views of individuals already working in the service.

Whilst the current SJT appeared to somewhat assess agreeableness, it did not meet the criteria for a construct-driven SJT because it was developed using a traditional approach (Tiffin et al., 2020). Construct-driven SJTs could be developed however, to assess traits that are perceived to be beneficial for practitioners working in mental health services, such as having 'resilience' or 'emotional intelligence.' Alternatively, researchers could explore what behaviours exhibit these traits across a multitude of scenarios. By disguising the trait being assessed by an SJT, it may be possible to develop construct-driven SJTs that are less prone to faking.

7.5　Conclusion

This work contributes to the existing literature by developing a greater understanding and conceptualisation of professionalism within a mental health services context. The

doctoral project resulted in the development of an SJT, which possesses validity evidence for predicting nurses' and AHPs' performance on the job. This is a valuable advancement in the field that can importantly facilitate the values-based recruitment of staff in a discipline where patients are vulnerable and reliant on staff to have the procedural knowledge, interpersonal skills and behaviours to deliver safe, effective and person-centred mental health care.

Chapter 8: Thesis summary

Patients with mental health problems are vulnerable and open to abuse and exploitation. The needs of this patient group can differ to that of patients in other healthcare settings; for example, patients may suffer from psychoses and be unable to distinguish what is real from what is unreal. In turn, the skills required of professionals working in mental health services may differ to that required of staff working in other healthcare disciplines. Furthermore, professionals working in mental health services typically endure mental health difficulties more than practitioners in other healthcare specialties and of concern, high stress and poor wellbeing in staff has been found to result in more fitness to practice hearings.

Values Based Recruitment (VBR) incorporates selecting staff with the right values to work in healthcare services. Since the introduction of VBR, many efforts have been made to implement these processes in the NHS, yet there is no single 'gold standard' assessment used for the selection of staff into mental health services. Various selection procedures are used for personnel selection, however, Situational Judgement Tests (SJTs) show more promise than other approaches given their cost-effectiveness and criterion-related validity, especially with regards to predicting job performance. Therefore, the overall aim of this project was to develop and validate an SJT that would assess staff members' knowledge of professionalism for a mental health services context.

Chapter 1 discusses the context leading up to the doctoral project. The importance of selecting staff with the appropriate skills and behaviours to deliver mental healthcare was highlighted. Subsequently, various selection methods were introduced, and a rationale was provided regarding the decision to develop an SJT that could assess individuals' non-academic attributes. In this instance, a decision was made to assess individuals' knowledge of professionalism for a mental health services context.

Chapter 2 builds on chapter 1 by extending the discussion regarding the use of SJTs for personnel selection. A brief history on the use of SJTs was provided and the fundamental beliefs regarding how and why SJTs are useful were discussed. Reference was also made regarding how best to develop and score SJTs. Various psychometric properties of SJTs were addressed and some potential pitfalls regarding the use of SJTs for personnel selection were examined.

Chapter 3 discusses various theoretical orientations and approaches that guide researchers in their work. The author's paradigmatic stance was highlighted, and the methods used to deliver on the overarching aims and objectives of the project were noted.

Chapter 4 addresses objective 1 of the research project. In this chapter, a systematic review, which sought to derive an operational definition of professionalism for a mental health services context was reported alongside its findings.

Chapter 5 builds on the findings of chapter 4 by exploring professionalism from the views of key stakeholders in the field, namely patients, carers and staff members working in mental health services. The qualitative study reported on in this chapter also sought to generate content for the SJT, therefore addressing objectives 1 and 2 of the research project.

Chapter 6 addresses objective 3 of the research project and reported on the development and validation of an SJT for mental health services.

Chapter 7 provides a general discussion regarding the findings of this doctoral work. One such finding was that performance on the SJT validly predicts supervisor perceptions of actual work performance for nurses and allied health professionals. Use of the SJT is therefore a potentially valid and cost-effective approach to supporting VBR in this setting.

List of references

AGUILAR, A., STUPANS, L. & SCUTTER, S. 2011. Assessing students' professionalism: considering professionalism's diverging definitions. *Education for Health,* 24**,** 599.

ALDERSON, M., PARENT-ROCHELEAU, X. & MISHARA, B. 2015. Critical Review on Suicide Among Nurses. *Crisis: The Journal of Crisis Intervention and Suicide Prevention,* 36**,** 91-101.

ALEXANDER, R. A. 1990. Correction formulas for correlations restricted by selection on an unmeasured variable. *Journal of Educational Measurement,* 27**,** 187-189.

ALTUS ASSESSMENTS. 2022. *Altus assessments: terms of use* [Online]. Available: https://takealtus.com/terms-of-use [Accessed 13 May 2022].

AMBROZY, D. M., IRBY, D. M., BOWEN, J. L., BURACK, J. H., CARLINE, J. D. & STRITTER, F. T. 1997. Role models' perceptions of themselves and their influence on students' specialty choices. *Academic Medicine,* 72**,** 1119-1121.

AMERICAN PSYCHOLOGICAL ASSOCIATION. 2020. *APA Dictionary of Psychology* [Online]. Available: https://dictionary.apa.org/prosocial [Accessed 15 December 2021].

ANDERSON, J. R. & LEBIERE, C. J. 2014. *The Atomic Components of Thought*. New York: Psychology Press.

ARCHDALL, C., ATAPATTU, T. & ANDERSON, E. 2013. Qualitative study of medical students' experiences of a psychiatric attachment. *The Psychiatrist,* 37**,** 21-24.

ARKSEY, H. & KNIGHT, P. 1999. *Interviewing for Social Scientists: An introductory resource with examples,* London: SAGE Publications ltd.

ARNOLD, J. D. 2020. *What Does Behavioral Consistency Really Mean?: Predictive Validity of Content and Method Factors in Employee Selection.* Doctoral Dissertation, Florida State University.

ARTHUR JR, W., GLAZE, R. M., JARRETT, S. M., WHITE, C. D., SCHURIG, I. & TAYLOR, J. E. 2014. Comparative evaluation of three situational judgment test response formats in terms of construct-related validity, subgroup differences, and susceptibility to response distortion. *Journal of Applied Psychology,* 99**,** 535-545.

ASHER, J. J. & SCIARRINO, J. A. 1974. Realistic work sample tests: A review. *Personnel Psychology,* 27**,** 519-533.

AVERITT, E. 2005. *The Relationship Among Sense of Humor, Defense Style, Levels of Trait Anxiety, and Locus of Control.* Masters Theses & Specialist Projects. Paper 454. Western Kentucky University.

AYLOTT, L., TIFFIN, P. A., BROWN, S. & FINN, G. M. 2022. Great expectations: Views and perceptions of professionalism amongst mental health services staff, patients and carers. *Journal of Mental Health,* 31**,** 139-146.

AYLOTT, L., TIFFIN, P. A., SAAD, M., LLEWELLYN, A. R. & FINN, G. M. 2019. Defining professionalism for mental health services: a rapid systematic review. *Journal of Mental Health,* 28**,** 546-565.

BAER, W. & SCHWARTZ, A. C. 2011. Teaching professionalism in the digital age on the psychiatric consultation-liaison service. *Psychosomatics,* 52**,** 303-309.

BARBOUR, R. 2007. *Doing focus groups,* London, SAGE Publications ltd.

BARBUI, C., PURGATO, M., ABDULMALIK, J., CALDAS-DE-ALMEIDA, J.M., EATON, J., GUREJE, O., HANLON, C., NOSÈ, M., OSTUZZI, G., SARACENO, B. & SAXENA, S. 2021. Efficacy of interventions to reduce coercive treatment in mental health services: umbrella review of randomised evidence. *The British Journal of Psychiatry,* 218, 185-195*.*

BARRICK, M. R., MOUNT, M. K. & STRAUSS, J. P. 1993. Conscientiousness and Performance of Sales Representatives: Test of the Mediating Effects of Goal Setting. *Journal of Applied Psychology,* 78**,** 715-722.

BATCHELOR, J. H. & MIAO, C. 2016. Extreme response style: A meta-analysis. *Journal of Organizational Psychology,* 16**,** 51-62.

BBC. 2022. *Edenfield Centre: Staff sacked over abuse at hospital* [Online]. Available: https://www.bbc.co.uk/news/uk-england-manchester-63261476 [Accessed 21 November 2022].

BEGG, C. B. & MAZUMDAR, M. 1994. Operating characteristics of a rank correlation test for publication bias. *Biometrics***,** 50, 1088-1101.

BELLINI, L. M., BAIME, M. & SHEA, J. A. 2002. Variation of mood and empathy during internship. *Journal of the American Medical Association,* 287**,** 3143-3146.

BELLINI, L. M. & SHEA, J. A. 2005. Mood change and empathy decline persist during three years of internal medicine training. *Academic Medicine,* 80**,** 164-167.

BERGER, R. 2015. Now I see it, now I don't: researcher's position and reflexivity in qualitative research. *Qualitative Research,* 15**,** 219-234.

BERGMAN, M. E., DRASGOW, F., DONOVAN, M. A., HENNING, J. B. & JURASKA, S. E. 2006. Scoring situational judgment tests: Once you get the data, your troubles begin. *International Journal of Selection and Assessment,* 14**,** 223-235.

BHUGRA, D. 2008a. Professionalism and psychiatry: the profession speaks. *Acta Psychiatrica Scandinavica,* 118**,** 327-329.

BHUGRA, D. 2008b. Renewing psychiatry's contract with society. *Psychiatric Bulletin,* 32**,** 281-283.

BHUGRA, D. 2009. Professionalism and psychiatry: past, present, future. *Australasian Psychiatry,* 17**,** 357-359.

BHUGRA, D. 2010. Teaching professionalism in psychiatry. *International Journal of Social Psychiatry,* 56**,** 323-325.

BHUGRA, D. & BROWN, N. 2007. Psychiatry: de-professionalisation. *International Journal of Social Psychiatry,* 53**,** 291-292.

BHUGRA, D. & GUPTA, S. 2010. Medical professionalism in psychiatry. *Advances in Psychiatric Treatment,* 16**,** 10-13.

BHUGRA, D. & GUPTA, S. 2011. Alienist in the 21st century. *Asian Journal of Psychiatry,* 4**,** 92-95.

BIEDENBACH, T. & JACOBSSON, M. 2016. The open secret of values: the roles of values and axiology in project research. *Project Management Journal,* 47**,** 139-155.

BIRDEN, H. 2012. *Professionalism in medicine. What is it and how can it be taught?* Doctoral Dissertation, The University of Sydney.

BIRDEN, H., GLASS, N., WILSON, I., HARRISON, M., USHERWOOD, T. & NASS, D. 2014. Defining professionalism in medical education: a systematic review. *Medical Teacher,* 36**,** 47-61.

BIRKELAND, S. A., MANSON, T. M., KISAMORE, J. L., BRANNICK, M. T. & SMITH, M. A. 2006. A meta-analytic investigation of job applicant faking on personality measures. *International Journal of Selection and Assessment,* 14**,** 317-335.

BLEDOW, R. & FRESE, M. 2009. A situational judgment test of personal initiative: Towards understanding construct based situational judgment tests. *Personnel Psychology,* 62**,** 229–258.

BLEGEBERG, B., BLOMBERG, A. & HEDELIN, B. 2008. Nurses conceptions of the professional role of operation theatre and psychiatric nurses. *Nordic Journal of Nursing Research & Clinical Studies / Vård i Norden,* 28**,** 9-13.

BRITISH MEDICAL ASSOCIATION. 2022. *Mental health pressures in England* [Online]. Available: https://www.bma.org.uk/advice-and-support/nhs-delivery-and-workforce/pressures/mental-health-pressures-data-analysis [Accessed 20 December 2022].

BOND, T. G., YAN, Z. & HEENE, M. 2013. *Applying the Rasch model: Fundamental measurement in the human sciences. Fourth edition,* New York, Routledge.

BOOTH, A. 2006. Australian supermodel?—A practical example of evidence-based library and information practice (EBLIP). *Health Information & Libraries Journal,* 23**,** 69-72.

BORMAN, W. C., BUCK, D. E., HANSON, M. A., MOTOWIDLO, S. J., STARK, S. & DRASGOW, F. 2001. An examination of the comparative reliability, validity, and accuracy of performance ratings made using computerized adaptive rating scales. *Journal of Applied Psychology,* 86**,** 965-973.

BOTHA, U. A., KOEN, L., JOSKA, J. A., PARKER, J. S., HORN, N., HERING, L. M. & OOSTHUIZEN, P. P. 2010. The revolving door phenomenon in psychiatry: comparing low-frequency and high-frequency users of psychiatric inpatient services in a developing country. *Social Psychiatry and Psychiatric Epidemiology,* 45**,** 461-468.

BOURAS, N. & IKKOS, G. 2013. Ideology, psychiatric practice and professionalism. *Psychiatriki,* 24**,** 17-26.

BRADLEY, C., MARSHALL, M. & GATH, D., 1995. Why do so few patients appeal against detention under Section 2 of the Mental Health Act?. *British Medical Journal*, 310, 364-367.

BRAQUEHAIS, M., LUSILLA, P., BEL, M., NAVARRO, M., NASILLO, V., DÍAZ, A., VALERO, S., PADRÓS, J., BRUGUERA, E. & CASAS, M. 2014. Dual Diagnosis Among Physicians: A Clinical Perspective. *Journal of Dual Diagnosis,* 10**,** 148-155.

BRAUN, V. & CLARKE, V. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology,* 3**,** 77-101.

BREEZE, J. A. & REPPER, J. 1998. Struggling for control: the care experiences of 'difficult'patients in mental health services. *Journal of Advanced Nursing,* 28**,** 1301-1311.

BRENDEL, D. H., CHU, J., RADDEN, J., LEEPER, H., POPE, H. G., SAMSON, J., TSIMPREA, G. & BODKIN, J. A. 2007. The price of a gift: an approach to receiving gifts from patients in psychiatric practice. *Harvard Review of Psychiatry,* 15**,** 43-51.

BRESSI, C., PORCELLANA, M., GAMBINI, O., MADIA, L., MUFFATTI, R., PEIRONE, A., ZANINI, S., ERLICHER, A., SCARONE, S. & ALTAMURA, A. 2009. Burnout among psychiatrists in Milan: a multicenter survey. *Psychiatric Services,* 60**,** 985-988.

BRETT, J. O., STANISZEWSKA, S., MOCKFORD, C., HERRON-MARX, S., HUGHES, J., TYSALL, C. & SULEMAN, R. 2014. A systematic review of the impact of patient and public involvement on service users, researchers and communities. *The Patient-Patient-Centered Outcomes Research,* 7**,** 387-395.

BRIDGE, P. D., MUSIAL, J., FRANK, R., ROE, T. & SAWILOWSKY, S. 2003. Measurement practices: methods for developing content-valid student examinations. *Medical Teacher,* 25**,** 414-421.

BRITISH MEDICAL ASSOCIATION 2017. *Breaking down barriers: the challenge of improving mental health outcomes*. [Online]. Available: https://unitementalhealth.files.wordpress.com/2018/02/breaking-down-barriers-mental-health-briefing-apr2017.pdf [Accessed 19 August 2022].

BRODY, H. & DOUKAS, D. 2014. Professionalism: a framework to guide medical education. *Medical Education,* 48**,** 980-987.

BROOKS, M. E. & HIGHHOUSE, S. 2006. Can good judgment be measured? *In:* WEEKLEY, J. A. & PLOYHART, R. E. (eds.) *Situational judgment tests: Theory, measurement, and application.* New Jersey: Lawrence Erlbaum Associates, Inc, 39-55.

BROWN, G. T. & PRANGER, T. 1992. Predictors of burnout for psychiatric occupational therapy personnel. *Canadian Journal of Occupational Therapy,* 59**,** 258-267.

BROWN, M. E. & DUEÑAS, A. N. 2020. A medical science educator's guide to selecting a research paradigm: building a basis for better research. *Medical Science Educator,* 30**,** 545-553.

BROWN, N. & BHUGRA, D. 2007. 'New' professionalism or professionalism derailed. *Psychiatric Bulletin,* 31**,** 281-283.

BROWNELL, A. K. W. & CÔTÉ, L. 2001. Senior Residents' Views on the Meaning of Professionalism and How They Learn about It. *Academic Medicine,* 76**,** 734-737.

BRUNN, M., DIEFENBACHER, A. & VOLPE-GILLOT, L. 2020. Psychiatrists and neurologists in dementia care: Professionalism, practice, and perspectives. *General Hospital Psychiatry,* 64**,** 105-107.

BRYMAN, A. 2016. *Social Research Methods,* Oxford, Oxford University Press.

BUNNISS, S. & KELLY, D. R. 2010. Research paradigms in medical education research. *Medical Education,* 44**,** 358-366.

BURFORD, B., MORROW, G., ROTHWELL, C., CARTER, M. & ILLING, J. 2014. Professionalism education should reflect reality: findings from three health professions. *Medical Education,* 48**,** 361-374.

CAFE, R. 2012. *Winterbourne View: Abuse footage shocked nation* [Online]. Available: https://www.bbc.co.uk/news/uk-england-bristol-20084254 [Accessed 13 December 2021].

CAMPBELL, D. 2022. *Inquiry investigates deaths of 1,500 NHS mental health patients in Essex* [Online]. Available: https://www.theguardian.com/society/2022/mar/28/inquiry-investigates-deaths-of-1500-nhs-mental-health-patients-in-essex [Accessed 11 May 2022].

CAMPION, M. C., PLOYHART, R. E. & MACKENZIE JR, W. I. 2014. The state of research on situational judgment tests: A content analysis and directions for future research. *Human Performance,* 27**,** 283-310.

CARE QUALITY COMMISSION. 2022. *Norfolk and Suffolk NHS Foundation Trust; all inspection reports and timeline* [Online]. Available: https://www.cqc.org.uk/provider/RMY/reports [Accessed 13 May 2022].

CATANO, V. M., BROCHU, A. & LAMERSON, C. D. 2012. Assessing the Reliability of Situational Judgment Tests Used in High-Stakes Situations. *International Journal of Selection and Assessment,* 20**,** 333-346.

CENTRE FOR REVIEWS AND DISSEMINATION 2009. *Systematic Reviews: CRD's guidance for undertaking reviews in health care.* [Online]. Available: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf [Accessed 19 August 2022].

CHAN, D. & SCHMITT, N. 1997. Video-based versus paper-and-pencil method of assessment in situational judgment tests: subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology,* 82**,** 143-159.

CHAN, D. & SCHMITT, N. 2002. Situational judgment and job performance. *Human Performance,* 15**,** 233-254.

CHANNEL FOUR TELEVISION CORPORATION. 2022. *Hospital Undercover Are They Safe? Dispatches* [Online]. Available: https://www.channel4.com/programmes/hospital-undercover-are-they-safe-dispatches [Accessed 21 November 2022].

CHEESMOND, N. 2020. *Making Mental Health Help-Seeking Easier: A Rural Australian Perspective.* Doctoral dissertation, University of Newcastle.

CHRISTIAN, M. S., EDWARDS, B. D. & BRADLEY, J. C. 2010. Situational judgment tests: Constructs assessed and a meta-analysis of their criterion-related validities. *Personnel Psychology,* 63**,** 83-117.

CLARK, V. L. P. & CRESWELL, J. W. 2008. *The mixed methods reader,* London, Sage Publications, Ltd.

CLEARY, M., FREEMAN, A., HUNT, G. E. & WALTER, G. 2005. What patients and carers want to know: an exploration of information and resource needs in adult mental health services. *Australian & New Zealand Journal of Psychiatry,* 39**,** 507-513.

CLEARY, M., FREEMAN, A. & WALTER, G. 2006. Carer participation in mental health service delivery. *International Journal of Mental Health Nursing,* 15**,** 189-194.

CLEVENGER, J., PEREIRA, G. M., WIECHMANN, D., SCHMITT, N. & HARVEY, V. S. 2001. Incremental validity of situational judgment tests. *Journal of Applied Psychology,* 86**,** 410-417.

COBB, P. 2001. Situated Cognition: Origins. *In:* SMELSER, N. J. & BALTES, P. B. (eds.) *International Encyclopedia of the Social & Behavioral Sciences.* Oxford: Pergamon, 14126-14129.

CODERRE, S., WOLOSCHUK, W. & MCLAUGHLIN, K. 2009. Twelve tips for blueprinting. *Medical Teacher,* 31**,** 322-324.

COFFEY, L. 2020. *The importance of maintaining fitness to practise, wellbeing and seeking help where necessary* [Online]. Available: https://www.hcpc-uk.org/news-and-events/blog/2020/the-importance-of-maintaining-fitness-to-practise-wellbeing-and-seeking-help-where-necessary [Accessed 13 December 2021].

CONSILIUM RESEARCH AND CONSULTANCY LTD 2016. *Study into the impact of a values based approach to recruitment and retention*. [Online]. Available: https://www.scie-socialcareonline.org.uk/study-into-the-impact-of-a-values-based-approach-to-recruitment-and-retention/r/a11G000000G63SlIAJ [Accessed 19 August 2022].

COOKE, A., SMITH, D. & BOOTH, A. 2012. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. *Qualitative Health Research,* 22**,** 1435-1443.

CORRY, M., PORTER, S. & MCKENNA, H. 2019. The redundancy of positivism as a paradigm for nursing research. *Nursing Philosophy,* 20**,** e12230.

CORSTJENS, J., LIEVENS, F. & KRUMM, S. (eds.) 2017. *Situational Judgement Tests for Selection*. *In* H. W. Goldstein, E. D. Pulakos, J. Passmore, & C. Semedo (Eds.), *The Wiley Blackwell handbook of the psychology of recruitment, selection and employee retention*. Hoboken, NJ: Wiley-Blackwell, 226-246.

CORTINA, J. M., GOLDSTEIN, N. B., PAYNE, S. C., DAVISON, H. K. & GILLILAND, S. W. 2000. The incremental validity of interview scores over and above cognitive ability and conscientiousness scores. *Personnel Psychology,* 53**,** 325-351.

COUSANS, F., PATTERSON, F., EDWARDS, H., WALKER, K., MCLACHLAN, J. C. & GOOD, D. 2017. Evaluating the complementary roles of an SJT and academic assessment for entry into clinical practice. *Advances in Health Sciences Education,* 22**,** 401-413.

COVERDALE, J. H. 2007. Virtues-based advice for beginning medical students. *Academic Psychiatry,* 31**,** 354-357.

COX, C. B., BARRON, L. G., DAVIS, W. & DE LA GARZA, B. 2017. Using situational judgment tests (SJTs) in training: Development and evaluation of a structured, low-fidelity scenario-based training method. *Personnel Review,* 46**,** 36-45.

CRAWFORD, P., BROWN, B. & MAJOMI, P. 2008. Professional identity in community mental health nursing: a thematic analysis. *International Journal of Nursing Studies,* 45**,** 1055-1063.

CRESWELL, J. W. 2014. *Research design: qualitative, quantitative and mixed methods approaches,* London: SAGE Publications Ltd.

CRESWELL, J. W. & POTH, C. N. 2007. *Qualitative inquiry and research design: Choosing among five approaches. Second edition,* California, Sage publications.

CROOK, A. E., BEIER, M. E., COX, C. B., KELL, H. J., HANKS, A. R. & MOTOWIDLO, S. J. 2011. Measuring relationships between personality, knowledge, and performance using single-response situational judgment tests. *International Journal of Selection and Assessment,* 19**,** 363-373.

CRUESS, S. R. 2006. Professionalism and medicine's social contract with society. *Clinical Orthopaedics and Related Research,* 449**,** 170-176.

CUMMINGS, J. & BENNETT, V. 2012. Compassion in practice. *Nursing, Midwifery and Care Staff. Our Vision and Strategy.* NHS Commissioning Board.

CURRY, L. A., NEMBHARD, I. M. & BRADLEY, E. H. 2009. Qualitative and mixed methods provide unique contributions to outcomes research. *Circulation,* 119**,** 1442-1452.

DAVISON, I., MCMANUS, C. & TAYLOR, C. 2016. *Evaluation of GP specialty selection*. [Online]. Available: http://pure-oai.bham.ac.uk/ws/portalfiles/portal/47719222/GP_Specialty_Selection_Evaluation.pdf [Accessed 19 August 2022].

DAY-CALDER, M. 2021. When friendly is too friendly, and knowing where to draw the line. *Nursing Standard,* 36**,** 27-28.

DE LENG, W. 2019. *Measuring Integrity for Selection into Medical School: Development of a Situational Judgement Test.* Doctoral dissertation, Erasmus University.

DE LENG, W. E., STEGERS-JAGER, K. M., HUSBANDS, A., DOWELL, J. S., BORN, M. P. & THEMMEN, A. P. N. 2017. Scoring method of a Situational Judgment Test: influence on internal consistency reliability, adverse impact and correlation with personality? *Advances in Health Sciences Education,* 22**,** 243-265.

DE LENG, W. E., STEGERS-JAGER, K. M., BORN, M. P. & THEMMEN, A. P. 2018. Influence of response instructions and response format on applicant perceptions of a

situational judgement test for medical school selection. *BMC Medical Education,* 18**,** 1-10.

DE LENG, W. E., STEGERS-JAGER, K. M., BORN, M. P. & THEMMEN, A. P. 2018. Integrity situational judgement test for medical school selection: judging 'what to do'versus 'what not to do'. *Medical Education,* 52**,** 427-437.

DE LENG, W. E., STEGERS-JAGER, K. M., BORN, M. P. & THEMMEN, A. P. 2019. Faking on a situational judgment test in a medical school selection setting: Effect of different scoring methods?. *International Journal of Selection and Assessment,* 27, 235-248*.*

DE MEIJER, L. A., BORN, M. P., VAN ZIELST, J. & VAN DER MOLEN, H. T. 2010. Construct-driven development of a video-based situational judgment test for integrity. *European Psychologist,* 15**,** 229-236.

DE WAAL, H., MALIK, A. & BHUGRA, D. 2010. The psychiatric profession: an expertise under siege? *International Journal of Social Psychiatry,* 56**,** 647-656.

DEARY, I., AGIUS, R. & SADLER, A. 1996. Personality and Stress in Consultant Psychiatrists. *International Journal of Social Psychiatry,* 42**,** 112-123.

DEPARTMENT OF HEALTH 2012. *Transforming care: A national response to Winterbourne View Hospital: Department of Health Review Final Report*. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/ attachment_data/file/213215/final-report.pdf [Accessed 19 August 2022].

DEPARTMENT OF HEALTH 2013a. *The Cavendish Review: An Independent Review into Healthcare Assistants and Support Workers in the NHS and social care settings,* London, The Stationary Office. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/ attachment_data/file/236212/Cavendish_Review.pdf [Accessed 19 August 2022].

DEPARTMENT OF HEALTH 2013b. *Delivering high quality, effective, compassionate care: Developing the right people with the right skills and the right values. A mandate from the Government to Health Education England: April 2013 to March 2015.* [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/ attachment_data/file/203332/29257_2900971_Delivering_Accessible.pdf [Accessed 19 August 2022].

DEPARTMENT OF HEALTH & SOCIAL CARE. 2021. *Guidance: The NHS Constitution for England* [Online]. Available: https://www.gov.uk/government/publications/the-nhs-constitution-for-england/the-nhs-constitution-for-england [Accessed 2 November 2021].

DEPARTMENT OF HEALTH. 2014. *Delivering high quality, effective, compassionate care: Developing the right people with the right skills and the right values. A mandate from the Government to Health Education England: April 2014 to March 2015*. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/

attachment_data/file/310170/DH_HEE_Mandate.pdf [Accessed 19 August 2022].

DEPTULA, P. & CHUN, M. B. J. 2013. A literature review of professionalism in surgical education: suggested components for development of a curriculum. *Journal of Surgical Education,* 70**,** 408-422.

DICICCO-BLOOM, B. & CRABTREE, B. F. 2006. The qualitative research interview. *Medical Education,* 40**,** 314-321.

DINGLE, A. D. & STUBER, M. L. 2008. Ethics education. *Child & Adolescent Psychiatric Clinics of North America,* 17**,** 187-207.

DISEKER, R. A. & MICHIELUTTE, R. 1981. An analysis of empathy in medical students before and following clinical experience. *Medical Education,* 56**,** 1004–1010.

DOMENIGHETTI, G., TOMAMICHEL, M., GUTZWILLER, F., BERTHOUD, S. & CASABIANCA, A. 1991. Psychoactive drug use among medical doctors is higher than in the general population. *Social Science & Medicine,* 33**,** 269-274.

DORE, K. L., KREUGER, S., LADHANI, M., ROLFSON, D., KURTZ, D., KULASEGARAM, K., CULLIMORE, A. J., NORMAN, G. R., EVA, K. W., BATES, S. & REITER, H. I. 2010. The reliability and acceptability of the multiple mini-interview as a selection instrument for postgraduate admissions. *Academic Medicine,* 85**,** S60-S63.

DUEÑAS, A. N., KIRKNESS, K. & FINN, G. M. 2020. Uncovering hidden curricula: use of dark humor in anatomy labs and its implications for basic sciences education. *Medical Science Educator,* 30**,** 345-354.

DUVAL, S. & TWEEDIE, R. 2000. Trim and fill: a simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics,* 56**,** 455-463.

EASTERBROOK, P. J., GOPALAN, R., BERLIN, J. & MATTHEWS, D. R. 1991. Publication bias in clinical research. *The Lancet,* 337**,** 867-872.

ELLIOTT, D. M. & GUY, J. D. 1993. Mental health professionals versus non-mental-health professionals: Childhood trauma and adult functioning. *Professional Psychology: Research and Practice,* 24**,** 83-90.

ELMAN, N. S. & FORREST, L. 2007. From trainee impairment to professional competence problems: Seeking new terminology that facilitates effective action. *Professional Psychology: Research and Practice,* 38**,** 501-509.

ERON, L. D. 1955. Effect of medical education on medical students' attitudes. *Academic Medicine,* 30**,** 559-566.

EVA, K. W., REITER, H. I., TRINH, K., WASI, P., ROSENFELD, J. & NORMAN, G. R. 2009. Predictive validity of the multiple mini-interview for selecting medical trainees. *Medical Education,* 43**,** 767-775.

EVANS, D. 2003. Hierarchy of evidence: a framework for ranking evidence evaluating healthcare interventions. *Journal of Clinical Nursing,* 12**,** 77-84.

FAKHOURY, W. & PRIEBE, S. 2002. The process of deinstitutionalization: an international overview. *Current Opinion in Psychiatry,* 15**,** 187-192.

FAY, J. 2013. The Baby and the Bathwater: An Unreserved Appreciation of Nick Totton's Critique of the Professionalisation of Psychotherapy. *Psychotherapy and Politics International,* 11**,** 26-33.

FERGUSON, E., MCMANUS, I. C., JAMES, D., O'HEHIR, F. & SANDERS, A. 2003. Pilot study of the roles of personality, references, and personal statements in relation to performance over the five years of a medical degree. Commentary: How to derive causes from correlations in educational studies. *British Medical Journal,* 326**,** 429-432.

FETTERS, M. D., CURRY, L. A. & CRESWELL, J. W. 2013. Achieving integration in mixed methods designs—principles and practices. *Health Services Research,* 48**,** 2134-2156.

FEUDTNER, C., CHRISTAKIS, D. A. & CHRISTAKIS, N. A. 1994. Do clinical clerks suffer ethical erosion? Students' perceptions of their ethical environment and personal development. *Academic Medicine,* 69**,** 670-679.

FINN, G., CRAMPTON, P., KEHOE, A., BALOGUN-KATUNG, A., TIFFIN, P., PAGE, M. & BUCHANAN, J. in press. *Exploring, understanding and evaluating experiences of Fitness to Practise at the General Dental Council: A realist evaluation. Final Report.*

FINN, G., GARNER, J. & SAWDON, M. 2010. 'You're judged all the time!' Students' views on professionalism: a multicentre study. *Medical Education,* 44**,** 814-825.

FINN, G. M., HAFFERTY, F. W. & QUINTON, H. 2022. The Significance of the Body in Health Professions Education. *In:* BROWN, M. E., VEEN, M. & FINN, G. M. (eds.) *Applied Philosophy for Health Professions Education.* Singapore: Springer, 85-101.

FINN, G. M., MWANDIGHA, L., PATON, L. W. & TIFFIN, P. A. 2018. The ability of 'non-cognitive'traits to predict undergraduate performance in medical schools: a national linkage study. *BMC Medical Education,* 18**,** 93.

FIRTH-COZENS, J. 2001. Interventions to improve physicians' well-being and patient care. *Social Science and Medicine,* 52**,** 215-222.

FIRTH-COZENS, J. 2007. Improving the health of psychiatrists. *Advances in Psychiatric Treatment,* 13, 161–168

FIRTH-COZENS, J., LEMA, V. & FIRTH, R. 1999. Specialty choice, stress & personality: their relationships over time. *Hospital Medicine,* 60**,** 751-755.

FISH, J. L. 2022. Genericism and managerialism: The limits to AMHP professionalisation and expertise. *International Journal of Law and Psychiatry,* 83**,** 101818.

FITZGERALD, S. 2019. *Managing Professional Roles in Home-Based Family Therapy: A Study of Marriage and Family Therapist Practices.* Doctoral dissertation, Antioch University.

FLANAGAN, J. C. 1954. The critical incident technique. *Psychological Bulletin,* 51**,** 327.

FLYNN, M. & CITARELLA, V. 2012. *Winterbourne View Hospital: a serious case review.* South Gloucestershire Safeguarding Adults Board. [Online]. Available: https://hosted.southglos.gov.uk/wv/report.pdf [Accessed 19 August 2022].

FOND-HARMANT, L., KANE, H., GOURRET BAUMGART, J., RUSCH, E., BRETON, H., EL-HAGE, W., DELOYER, J., LEBAS, M.C., MARAZZITI, D., THOME, J. & DENIS, F., 2021. International professional practices in mental health, organization of psychiatric care, and COVID-19: A survey protocol. *PloS one*, 16, e0261818.

FRANCIS, R. 2013. *Report of the Mid Staffordshire NHS Foundation Trust Public Inquiry: Executive Summary*. [Online]. Available: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/279124/0947.pdf [Accessed 19 August 2022].

FREEMAN, T. 2006. 'Best practice'in focus group research: making sense of different views. *Journal of Advanced Nursing,* 56**,** 491-497.

FREUD, S. 1905. Jokes and the species of the comic. *In:* STRACHEY, J. (ed.) *Jokes and their relation to the unconscious.* New York: W. W. Norton & Company, 181-236.

FUNDER, D. C. 2006. Towards a resolution of the personality triad: Persons, situations, and behaviors. *Journal of Research in Personality*, 40, 21-34.

GABBARD, G. O. 2019. Digital professionalism. *Academic Psychiatry,* 43**,** 259-263.

GAUGLER, B. B., ROSENTHAL, D. B., THORNTON, G. C. & BENTSON, C. 1987. Meta-analysis of assessment center validity. *Journal of Applied Psychology,* 72**,** 493-511.

GENERAL MEDICAL COUNCIL. 2021. *Good Medical Practice* [Online]. Available: https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/good-medical-practice [Accessed 2 November 2021].

GENERAL MEDICAL COUNCIL. 2022. *Professional behaviour and fitness to practise* [Online]. Available: https://www.gmc-uk.org/education/standards-guidance-and-curricula/guidance/student-professionalism-and-ftp/professional-behaviour-and-fitness-to-practise [Accessed 19 November 2022].

GHADIRIAN, F., SALSALI, M. & CHERAGHI, M. A. 2014. Nursing professionalism: An evolutionary concept analysis. *Iranian Journal of Nursing and Midwifery Research,* 19**,** 1-10.

GLAS, G. 2017. On the existential core of professionalism in mental health care. *Mental Health, Religion & Culture,* 20**,** 536-543.

GLAUSER, W. 2020. Does medical professionalism have a dark side? *CMAJ,* 192**,** E1646-1647.

GOFFIN, R. D., GELLATLY, I. R., PAUNONEN, S. V., JACKSON, D. N. & MEYER, J. P. 1996. Criterion validation of two approaches to performance appraisal: The behavioral observation scale and the relative percentile method. *Journal of Business and Psychology,* 11**,** 23-33.

GOFFIN, R. D., JELLEY, R. B., POWELL, D. M. & JOHNSTON, N. G. 2009. Taking advantage of social comparisons in performance appraisal: The relative percentile method. *Human Resource Management,* 48**,** 251-268.

GOLD, B. & HOLODYNSKI, M. 2015. Development and Construct Validation of a Situational Judgment Test of Strategic Knowledge of Classroom Management in Elementary Schools. *Educational Assessment,* 20**,** 226-248.

GOODYEAR, R., LICHTENBERG, J., HUTMAN, H., OVERLAND, E., BEDI, R., CHRISTIANI, K., DI MATTIA, M., DU PREEZ, E., FARRELLI, B., FEATHER, J., GRANT, J., HAN, Y. J., JU, Y., LEE, D. G., LEE, H., NICHOLAS, H., NIELSEN, J. J., SINACORE, A., TU, S. F. & YOUNG, C. 2016. A global portrait of counselling psychologists' characteristics, perspectives, and professional behaviors. *Counselling Psychology Quarterly,* 29**,** 115-138.

GOSS, B. D., RYAN, A. T., WARING, J., JUDD, T., CHIAVAROLI, N. G., O'BRIEN, R. C., TRUMBLE, S. C. & MCCOLL, G. J. 2017. Beyond Selection: The Use of Situational Judgement Tests in the Teaching and Assessment of Professionalism. *Academic Medicine,* 92**,** 780-784.

GOTTLIEB, M. C., YOUNGGREN, J. N. & MURCH, K. B. 2009. Boundary management for Cognitive Behavioral Therapies. *Cognitive and Behavioral Practice,* 16**,** 164-171.

GOURRET BAUMGART, J., KANE, H., EL-HAGE, W., DELOYER, J., MAES, C., LEBAS, M.C., MARAZZITI, D., THOME, J., FOND-HARMANT, L. & DENIS, F., 2021. The early impacts of the COVID-19 pandemic on mental health facilities and psychiatric professionals. *International Journal of Environmental Research and Public Health,* 18, 8034.

GREEN, M., ZICK, A. & MAKOUL, G. 2009. Defining professionalism from the perspective of patients, physicians, and nurses. *Academic Medicine,* 84**,** 566-573.

GROOTHUIZEN, J. E. 2020. *Values, situational judgement, and the clinical practice environment: A study of adult nursing students.* Doctoral dissertation, University of Surrey.

GROOTHUIZEN, J.E., CALLWOOD, A. & GALLAGHER, A., 2018. What is the value of Values Based Recruitment for nurse education programmes?. *Journal of advanced nursing,* 74, 1068-1077.

GROVES, L. C. & KERSON, T. S. 2011. The influence of professional identity and the private practice environment: Attitudes of clinical social workers toward addressing the social support needs of clients. *Smith College Studies in Social Work,* 81**,** 218-233.

GUBA, E. G. & LINCOLN, Y. S. 1982. Epistemological and methodological bases of naturalistic inquiry. *ECTJ,* 30**,** 233-252.

GUENOLE, N., CHERNYSHENKO, O. S. & WEEKLY, J. 2017. On designing construct driven situational judgment tests: Some preliminary recommendations. *International Journal of Testing,* 17**,** 234-252.

GUPTA, S., PATERSON, M. L., LYSAGHT, R. M. & VON ZWECK, C. M. 2012. Experiences of burnout and coping strategies utilized by occupational therapists. *Canadian Journal of Occupational Therapy,* 79**,** 86-95.

GUYON, H., KOP, J. L., JUHEL, J. & FALISSARD, B. 2018. Measurement, ontology, and epistemology: Psychology needs pragmatism-realism. *Theory & Psychology,* 28**,** 149-171.

HAFFERTY, F. W. 1998. Beyond curriculum reform: confronting medicine's hidden curriculum. *Academic Medicine,* 73**,** 403-407.

HAFFERTY, F. W. & FRANKS, R. 1994. The hidden curriculum, ethics teaching, and the structure of medical education. *Academic medicine.,* 69**,** 861–871.

HALE, T. & WRIGHT, C. 2021. Unprofessional practice and student professionalism dilemmas: What can radiography learn from the other health professions? *Radiography,* 27**,** 1211-1218.

HAPPELL, B. 2006. Would the real mental health nurse please stand up? The relationship between identification and professional identity. *International Journal of Mental Health Nursing,* 15**,** 155-156.

HARRIS, S. E. & KURPIUS, S. E. R. 2014. Social networking and professional ethics: Client searches, informed consent, and disclosure. *Professional Psychology-Research and Practice,* 45**,** 11-19.

HASHMI, W. T. A. A. 2018. *Developing a Situational Judgment Test for Admission into Initial Teacher Education in Oman: An Exploratory Study.* Doctoral Dissertation, University of York.

HAVERKAMP, B. E., ROBERTSON, S. E., CAIRNS, S. L. & BEDI, R. P. 2011. Professional issues in canadian counselling psychology: Identity, education, and professional practice. *Canadian Psychology,* 52**,** 256-264.

HAWTON, K., CLEMENTS, A., SAKAROVITCH, C., SIMKIN, S. & DEEKS, J. J. 2001. Suicide in doctors: a study of risk according to gender, seniority and specialty in medical practitioners in England and Wales, 1979–1995. *Journal of Epidemiology & Community Health,* 55**,** 296-300.

HEALTH & CARE PROFESSIONS COUNCIL. 2022. *Fitness to practise* [Online]. Available: https://www.hcpc-uk.org/concerns/what-we-investigate/fitness-to-practise/ [Accessed 19 November 2022].

HEALTH EDUCATION ENGLAND. 2014a. *Evaluation of Values Based Recruitment (VBR) in the NHS: Analysis of VBR Activity within NHS Trusts.* London: NHS Health Education England.

HEALTH EDUCATION ENGLAND 2014b. *Evaluation of Values Based Recruitment (VBR) in the NHS: Literature Review and Evaluation Criteria.* London: NHS Health Education England.

HEPONIEMI, T., AALTO, A.-M., PUTTONEN, S., VÄNSKÄ, J. & ELOVAINIO, M. 2014. Work-Related Stress, Job Resources, and Well-Being Among Psychiatrists and Other Medical Specialists in Finland. *Psychiatric Services,* 65**,** 796-801.

HICKS, L. E. 1970. Some properties of ipsative, normative, and forced-choice normative measures. *Psychological Bulletin,* 74**,** 167-184.

HODGES, B. D., GINSBURG, S., CRUESS, R., CRUESS, S., DELPORT, R., HAFFERTY, F., HO, M.-J., HOLMBOE, E., HOLTMAN, M., OHBU, S., REES, C., TEN CATE, O.,

TSUGAWA, Y., VAN MOOK, W., WASS, V., WILKINSON, T. & WADE, W. 2011. Assessment of professionalism: recommendations from the Ottawa 2010 Conference. *Medical Teacher,* 33**,** 354-363.

HOGAN, A. J. 2019. Social and medical models of disability and mental health: evolution and renewal. *CMAJ,* 191**,** E16-E18.

HOJAT, M., MANGIONE, S., NASCA, T. J., RATTNER, S., ERDMANN, J. B., GONNELLA, J. S. & MAGEE, M. 2004. An empirical study of decline in empathy in medical school. *Medical Education,* 38**,** 934-941.

HOLLOWAY, I. & GALVIN, K. 2017. *Qualitative Research in Nursing and Healthcare*, Wiley Blackwell.

HORDICHUK, C. J., ROBINSON, A. J. & SULLIVAN, T. M. 2015. Conceptualising professionalism in occupational therapy through a Western lens. *Australian Occupational Therapy Journal,* 62**,** 150-159.

HOWE, K. R. 1988. Against the quantitative-qualitative incompatibility thesis or dogmas die hard. *Educational Researcher,* 17**,** 10-16.

IKKOS, G. & MACE, C. 2009. Professionalising psychotherapy: Lessons from the development of psychiatry. *European Journal of Psychotherapy, Counselling and Health,* 11**,** 161-172.

IKKOS, G., MCQUEEN, D. & ST. JOHN-SMITH, P. 2011. Psychiatry's contract with society: What is expected? *Acta Psychiatrica Scandinavica,* 124**,** 1-3.

IRVINE, D. 1997. The performance of doctors. I: Professionalism and self regulation in a changing world. *British Medical Journal,* 314**,** 1540-1542.

IVANKOVA, N. V. & CRESWELL, J. W. 2009. Mixed Methods. *In:* HEIGHAM, J. (ed.) *Qualitative research in applied linguistics: a practical introduction.* New York: Palgrave Macmillan.

JAIN, S., DUNN, L. B., WARNER, C. H. & ROBERTS, L. W. 2011a. Results of a multisite survey of U.S. psychiatry residents on education in professionalism and ethics. *Academic Psychiatry,* 35**,** 175-183.

JAIN, S., HOOP, J. G., DUNN, L. B. & ROBERTS, L. W. 2010. Psychiatry residents' attitudes on ethics and professionalism: Multisite survey results. *Ethics & Behavior,* 20**,** 10-20.

JAIN, S., LAPID, M. I., DUNN, L. B. & ROBERTS, L. W. 2011b. Psychiatric residents' needs for education about informed consent, principles of ethics and professionalism, and caring for vulnerable populations: results of a multisite survey. *Academic Psychiatry,* 35**,** 184-90.

JAKOVLJEVIĆ, M. 2012. Professionalism in psychiatry and medicine: A hot topic. *Psychiatria Danubina,* 24**,** 342-345.

JHA, V., BEKKER, H., DUFFY, S. & ROBERTS, T. 2006. Perceptions of professionalism in medicine: a qualitative study. *Medical Education,* 40**,** 1027-1036.

JOHN, N. J., SHELTON, P. G., LANG, M. C. & INGERSOLL, J. 2017. Training psychiatry residents in professionalism in the digital world. *Psychiatric Quarterly***,** 88, 263-270.

JOHNSON, R. B. & ONWUEGBUZIE, A. J. 2004. Mixed methods research: A research paradigm whose time has come. *Educational Researcher,* 33**,** 14-26.

JOINER, A. B., HUSAIN, N., DUDDU, V. & CHAUDHRY, I. B. 2015. Medical professionalism education for psychiatry trainees: does it meet standards? *Australasian Psychiatry,* 23**,** 72-75.

JONES, J. W., MCCULLOUGH, L. B. & RICHMAN, B. W. 2006. Ethics and professionalism: Do we need yet another surgeons' charter? *Journal of Vascular Surgery,* 44**,** 903-906.

KANE, H., BAUMGART, J.G., RUSCH, E., DELOYER, J., FUENZALIDA, C., KELEMEN, G., KRZYSTANEK, M., MARAZZITI, D., MORAITOU, M., REUNANEN, M. & SHYHRETE, R., 2022. The impact of COVID-19 on psychiatric and mental health services in Europe: Suffering experienced by professionals. *BMC Health Services Research*, 22, 1-10.

KANNING, U. P., GREWE, K., HOLLENBERG, S. & HADOUCH, M. 2006. From the Subjects' point of view: Reactions to different types of situational judgement items. *European Journal of Psychological Assessment,* 22**,** 168-176.

KELLY, M. P. & MAY, D. 1982. Good and bad patients: a review of the literature and a theoretical critique. *Journal of Advanced Nursing,* 7**,** 147-156.

KHANGURA, S., KONNYU, K., CUSHMAN, R., GRIMSHAW, J. & MOHER, D. 2012. Evidence summaries: the evolution of a rapid review approach. *Systematic Reviews,* 1**,** 10.

KIGER, M. E. & VARPIO, L. 2020. Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher***,** 42, 846-854.

KITZINGER, J. 1995. Qualitative research. Introducing focus groups. *British Medical Journal,* 311**,** 299-302.

KOEKKOEK, B., VAN MEIJEL, B. & HUTSCHEMAEKERS, G. 2006. "Difficult patients" in mental health care: a review. *Psychiatric Services,* 57**,** 795-802.

KOMIĆ, D., MARUŠIĆ, S. L. & MARUŠIĆ, A. 2015. Research integrity and research ethics in professional codes of ethics: Survey of terminology used by professional organizations across research disciplines. *Plos One,* 10**,** e0133662.

KRIPPENDORFF, K. 1970. Estimating the reliability, systematic error and random error of interval data. *Educational and Psychological Measurement,* 30, 61-70.

KRUMM, S., LIEVENS, F., HÜFFMEIER, J., LIPNEVICH, A. A., BENDELS, H. & HERTEL, G. 2015. How "situational" is judgment in situational judgment tests? *Journal of Applied Psychology,* 100**,** 399-416.

KUDER, G. F. & RICHARDSON, M. W. 1937. The theory of the estimation of test reliability. *Psychometrika,* 2**,** 151-160.

LAMBERT, S. & LOISELLE, C. 2008. Combining individual interviews and focus groups to enhance data richness. *Journal of Advanced Nursing,* 62**,** 228–237.

LAMPSHIRE, D. 2012. Living the dream. *Psychosis,* 4**,** 172-178.

LAPID, M., MOUTIER, C., DUNN, L., HAMMOND, K. G. & ROBERTS, L. W. 2009. Professionalism and ethics education on relationships and boundaries: psychiatric residents' training preferences. *Academic Psychiatry,* 33**,** 461-469.

LATHAM, G. P., SAARI, L. M., PURSELL, E. D. & CAMPION, M. A. 1980. The situational interview. *Journal of Applied Psychology,* 65**,** 422-427.

LEPPMA, M., TAYLOR, J. M., SPERO, R. A., LEONARD, J. M., FOSTER, M. N. & DANIELS, J. A. 2016. Working with veterans and military families: An assessment of professional competencies. *Professional Psychology: Research and Practice,* 47**,** 84-92.

LEWIS-BECK, M. S., BRYMAN, A., & FUTING LIAO, T. 2004. *The SAGE encyclopedia of social science research methods.* California: Sage Publications, Inc.

LIEVENS, F. 2013. Adjusting medical school admission: assessing interpersonal skills using situational judgement tests. *Medical Education,* 47**,** 182-189.

LIEVENS, F. 2017a. Construct-Driven SJTs: Toward an Agenda for Future Research. *International Journal of Testing,* 17**,** 269-276.

LIEVENS, F. 2017b. Assessing personality–situation interplay in personnel selection: Toward more integration into personality research. *European Journal of Personality,* 31**,** 424-440.

LIEVENS, F., ONES, D. S. & DILCHERT, S. 2009. Personality scale validities increase throughout medical school. *Journal of Applied Psychology,* 94**,** 1514-1535.

LIEVENS, F. & 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**,** 927-940.

LIEVENS, F., PEETERS, H. & SCHOLLAERT, E. 2008. Situational judgment tests: A review of recent research. *Personnel Review,* 37**,** 426-441.

LIEVENS, F. & SACKETT, P. R. 2006. Video-based versus written situational judgment tests: A comparison in terms of predictive validity. *Journal of Applied Psychology,* 91**,** 1181 - 1188.

LIEVENS, F. & SACKETT, P. R. 2007. Situational judgment tests in high-stakes settings: Issues and strategies with generating alternate forms. *Journal of Applied Psychology,* 92**,** 1043-1055.

LIEVENS, F. & SACKETT, P. R. 2012. The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology,* 97**,** 460-468.

LIGHTFOOT, G. 2022. *Teesside NHS trust shakes up its management* [Online]. Available: https://www.thenorthernecho.co.uk/news/20198515.shake-up-nhs-trust-following-critical-report---were-getting-there/?ref=rss&IYA-mail=cb3adb00-7173-4883-a760-c40e1495e2d0 [Accessed 17 June 2022].

LINCOLN, Y. S. & GUBA, E. G. 1985. *Naturalistic inquiry,* California: Sage Publications, Inc.

LJUNGBERG, A., DENHOV, A. & TOPOR, A. 2015. The art of helpful relationships with professionals: A meta-ethnography of the perspective of persons with severe mental illness. *Psychiatric Quarterly,* 86**,** 471-95.

LLOYD, C. & KING, R. 2003. Consumer and carer participation in mental health services. *Australasian Psychiatry,* 11, 80-84.

LLOYD, C., KING, R. & CHENOWETH, L. 2002. Social work, stress and burnout: A review. *Journal of mental health,* 11**,** 255-265.

LOOI, J. C., ALLISON, S., BASTIAMPILLAI, T. & KISELY, S. R. 2022. Mental health services need action on organisational culture and justice. *Australasian Psychiatry***,** p.10398562221111945.

LORENZO-SEVA, U. & FERRANDO, P. J. 2006. FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods,* 38**,** 88-91.

LORENZO-SEVA, U. & FERRANDO, P. J. 2013. FACTOR 9.2 A Comprehensive Program for Fitting Exploratory and Semiconfirmatory Factor Analysis and IRT Models. *Applied Psychological Measurement,* 37**,** 497-498.

MABEN, J. 2003. *The fate of ideals and values: experiences of newly qualified nurses.* Doctoral Dissertation, University of Southampton.

MACCANN, C., ROBERTS, R. D., MATTHEWS, G. & ZEIDNER, M. 2004. Consensus scoring and empirical option weighting of performance-based emotional intelligence (EI) tests. *Personality and Individual Differences,* 36**,** 645-662.

MALHI, G. S. 2008. Professionalizing psychiatry: from 'amateur' psychiatry to 'a mature' profession. *Acta Psychiatrica Scandinavica,* 118**,** 255-258.

MALONE, J. L. 2012. Ethical professional practice: exploring the issues for health services to rural Aboriginal communities. *Rural Remote Health,* 12**,** 1891.

MANNION, R., DAVIES, H., POWELL, M., BLENKINSOPP, J., MILLAR, R., MCHALE, J. & SNOWDEN, N. 2019. Healthcare scandals and the failings of doctors: Do official inquiries hold the profession to account? *Journal of Health Organization and Management,* 33**,** 221–240.

MASA'DEH, R., ALHALAIQA, F., ABURUZ, M. E., AL-DWEIK, G. & AL-AKASH, H. Y. 2017. Perceived stress in nurses: A comparative study. *Global Journal of Health Science,* 9**,** 195-203.

MAY, D. & KELLY, M. P. 1982. Chancers, pests and poor wee souls: problems of legitimation in psychiatric nursing. *Sociology of Health & Illness,* 4**,** 279-301.

MAYS, N. & POPE, C. 2000. Assessing quality in qualitative research. *British Medical Journal,* 320**,** 50-52.

MCDANIEL, M., MORGESON, F., FINNEGAN, E., CAMPION, M. & BRAVERMAN, E. 2001. Use of situational judgment tests to predict job performance: a clarification of the literature. *Journal of Applied Psychology,* 86**,** 730-740.

MCDANIEL, M. A., HARTMAN, N. S., WHETZEL, D. L. & GRUBB III, W. L. 2007. Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology,* 60**,** 63-91.

MCDANIEL, M. A. & NGUYEN, N. T. 2001. Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment,* 9**,** 103-113.

MCDANIEL, M. A., PSOTKA, J., LEGREE, P. J., YOST, A. P. & WEEKLEY, J. A. 2011. Toward an understanding of situational judgment item validity and group differences. *Journal of Applied Psychology,* 96**,** 327-336.

MCDANIEL, M. A. & WHETZEL, D. L. 2005. Situational judgment test research: Informing the debate on practical intelligence theory. *Intelligence,* 33**,** 515-525.

MCDANIEL, M. A., WHETZEL, D. L., HARTMAN, N. S., NGUYEN, N. T. & GRUBB, W. L. 2006. Situational judgment tests: Validity and an integrative model. *In:* WEEKLEY, J. A. & PLOYHART, R. (eds.) *Situational judgment tests: Theory, measurement, and application.* New Jersey: Lawrence Erlbaum Associates, Inc, 1-10.

MCDANIEL, M. A., WHETZEL, D. L., SCHMIDT, F. L. & MAURER, S. D. 1994. The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology,* 79**,** 599-616.

MCLACHLAN, J. C., FINN, G. & MACNAUGHTON, J. 2009. The Conscientiousness Index: A Novel Tool to Explore Students' Professionalism. *Academic Medicine,* 84**,** 559-565.

MCMANUS, I. C., DEWBERRY, C., NICHOLSON, S., DOWELL, J. S., WOOLF, K. & POTTS, H. W. 2013. Construct-level predictive validity of educational attainment and intellectual aptitude tests in medical student selection: meta-regression of six UK longitudinal studies. *BMC Medicine,* 11, 1-21.

MEDICAL SCHOOLS COUNCIL 2011. *Improving Selection to the Foundation Programme: Final Report.* [Online]. Available: https://isfporguk.files.wordpress.com/2017/04/isfp-final-report-appendices.pdf [Accessed 19 August 2022].

MENTAL CAPACITY ACT 2005, c.9. [Online]. Available: https://www.legislation.gov.uk/ukpga/2005/9/contents [Accessed 7 November 2022].

MENTAL HEALTH ACT 1983 , c.20. [Online]. Available: https://www.legislation.gov.uk/ukpga/1983/20/part/II [Accessed 7 November 2022].

MERRIAM, S. B. 2010. Adult Learning. *In:* PETERSON, P., BAKER, E. & MCGAW, B. (eds.) *International Encyclopedia of Education (Third Edition).* Oxford: Elsevier.

MERTENS, D.M., 2007. Transformative paradigm: Mixed methods and social justice. *Journal of mixed methods research*, 1, 212-225.

MERTENS, D. M. 2017. Transformative research: personal and societal. *International Journal for Transformative Research,* 4**,** 18-24.

METHLEY, A. M., CAMPBELL, S., CHEW-GRAHAM, C., MCNALLY, R. & CHERAGHI-SOHI, S. 2014. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research,* 14**,** 1-10.

MIERS, M. E., RICKABY, C. E. & POLLARD, K. C. 2007. Career choices in health care: Is nursing a special case? A content analysis of survey data. *International Journal of Nursing Studies,* 44**,** 1196-1209.

MILLER, R. L. & BREWER, J. 2003. *The a-z of social research*. London: SAGE Publications Ltd.

MILTON, J., WATKINS, K. E., STUDDARD, S. S. & BURCH, M. 2003. The ever widening gyre: Factors affecting change in adult education graduate programs in the United States. *Adult Education Quarterly,* 54**,** 23-41.

MOHAJAN, H. K. 2017. Two criteria for good measurements in research: Validity and reliability. *Annals of Spiru Haret University. Economic Series,* 17**,** 58-82.

MOHER, D., LIBERATI, A., TETZLAFF, J., ALTMAN, D. & PRISMA GROUP*. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine,* 151**,** 264-269.

MCLACHLAN, J. C 2017. Afterword. *In:* MONROUXE, L. V. & REES, C. E. 2017. *Healthcare professionalism: improving practice through reflections on workplace dilemmas.* Oxford: John Wiley & Sons Ltd, 237-238.

MONROUXE, L. V., REES, C. E. & HU, W. 2011. Differences in medical students' explicit discourses of professionalism: acting, representing, becoming. *Medical Education,* 45**,** 585-602.

MORAN, C. C. & HUGHES, L. P. 2006. Coping with stress: Social work students and humour. *Social Work Education,* 25**,** 501-517.

MORGAN, D. L. 2007. Paradigms lost and pragmatism regained: Methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research,* 1**,** 48-76.

MORRIS, A., BLOOM, J. R. & KANG, S. 2007. Organizational and individual factors affecting consumer outcomes of care in mental health services. *Administration and Policy in Mental Health and Mental Health Services Research,* 34**,** 243-253.

MORRISON, J. & MORRISON, T. 2001. Psychiatrists Disciplined by a State Medical Board. *American Journal Psychiatry,* 158**,** 474–478.

MORROW, G., BURFORD, B., ROTHWELL, C., CARTER, M., MCLACHLAN, J. & ILLING, J. 2014. *Professionalism in healthcare professionals*. [Online]. Available: https://www.hcpc-uk.org/globalassets/resources/reports/professionalism-in-healthcare-professionals.pdf [Accessed 19 August 2022].

MOSS, F. 1926. Do you know how to get along with people? Why some people get head in the world while others do not. *Scientific American,* 135**,** 26-27.

MOTOWIDLO, S. J. & BEIER, M. E. 2010. Differentiating specific job knowledge from implicit trait policies in procedural knowledge measured by a situational judgment test. *Journal of Applied Psychology,* 95**,** 321-333.

MOTOWIDLO, S. J., CROOK, A. E., KELL, H. J. & NAEMI, B. 2009. Measuring procedural knowledge more simply with a single-response situational judgment test. *Journal of Business and Psychology,* 24**,** 281-288.

MOTOWIDLO, S. J., DUNNETTE, M. D. & CARTER, G. W. 1990. An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology,* 75**,** 640-647.

MOTOWIDLO, S. J., HOOPER, A. C. & JACKSON, H. L. 2006a. Implicit policies about relations between personality traits and behavioral effectiveness in situational judgment items. *Journal of Applied Psychology,* 91**,** 749-761.

MOTOWIDLO, S. J., HOOPER, A. C. & JACKSON, H. L. 2006b. A theoretical basis for situational judgment tests. *In:* WEEKLEY, J. A. & PLOYHART, R. E. (eds.) *Situational judgment tests: Theory, measurement, and application.* New Jersey: Lawrence Erlbaum Associates Publishers.

MOTOWIDLO, S. J., LIEVENS, F. & GHOSH, K. 2018. Prosocial implicit trait policies underlie performance on different situational judgment tests with interpersonal content. *Human Performance,* 31**,** 238-254.

MURAKAMI, M., KAWABATA, H. & MAEZAWA, M. 2009. The perception of the hidden curriculum on medical education: an exploratory study. *Asia Pacific Family Medicine,* 8**,** 1-7.

NANDAKUMAR, R. 1991. Traditional dimensionality versus essential dimensionality. *Journal of Educational Measurement,* 28**,** 99-117.

NATIONAL CONFIDENTIAL INQUIRY INTO SUICIDE AND HOMICIDE BY PEOPLE WITH MENTAL ILLNESS (NCISH) 2015. *In-patient Suicide Under Observation. Manchester.* [Online] Available: https://documents.manchester.ac.uk/display.aspx?DocID=37570 [Accessed 19 August 2022].

NEUMANN, M., EDELHÄUSER, F., TAUSCHEL, D., FISCHER, M. R., WIRTZ, M., WOOPEN, C., HARAMATI, A. & SCHEFFER, C. 2011. Empathy decline and its reasons: a systematic review of studies with medical students and residents. *Academic Medicine,* 86**,** 996-1009.

NGUYEN, N. T., BIDERMAN, M. D. & MCDANIEL, M. A. 2005. Effects of response instructions on faking a situational judgment test. *International Journal of Selection and Assessment,* 13**,** 250-260.

NHS ENGLAND, PUBLIC HEALTH ENGLAND, HEALTH EDUCATION ENGLAND, MONITOR, CARE QUALITY COMMISSION & NHS TRUST DEVELOPMENT AUTHORITY 2014. *Five Year Forward View.* [Online] Available: https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf [Accessed 19 August 2022].

NHS JOBS 2018. *NHS Jobs.*

NOBLE, L. M. & DOUGLAS, B. C. 2004. What users and relatives want from mental health services. *Current Opinion in Psychiatry,* 17**,** 289-96.

NURSING & MIDWIFERY COUNCIL. 2017. *Enabling professionalism in nursing and midwifery practice.* [Online]. Available: https://www.nmc.org.uk/globalassets/sitedocuments/other-publications/enabling-professionalism.pdf [Accessed 19 August 2022].

NURSING & MIDWIFERY COUNCIL. 2021. *Read The Code online* [Online]. Available: https://www.nmc.org.uk/standards/code/read-the-code-online [Accessed 2 November 2021].

NURSING & MIDWIFERY COUNCIL. 2022. *Aims and principles for fitness to practise* [online]. Available: https://www.nmc.org.uk/ftp-library/understanding-fitness-to-practise/using-fitness-to-practise/ [Accessed 19 November 2022].

O'CONNELL, M. S., HARTMAN, N. S., MCDANIEL, M. A., GRUBB III, W. L. & LAWRENCE, A. 2007. Incremental validity of situational judgment tests for task and contextual job performance. *International Journal of Selection and Assessment,* 15**,** 19-29.

O'REILLY, K. 2009. Positivism. *In:* O'REILLY, K. (ed.) *Key concepts in ethnography.* London: SAGE Publications Ltd, pp.163-167.

O'SHEA, N. 2020. *Covid-19 and the nation's mental health: Forecasting needs and risks in the UK: October 2020.* [Online] Available: https://www.centreformentalhealth.org.uk/sites/default/files/publication/download/CentreforMentalHealth_COVID_MH_Forecasting3_Oct20_0.pdf [Accessed 19 August 2022].

O'SHEA, N. 2021. *Covid-19 and the nation's mental health: Forecasting needs and risks in the UK: May, 2021.* [Online] Available: https://www.centreformentalhealth.org.uk/sites/default/files/publication/download/CentreforMentalHealth_COVID_MH_Forecasting4_May21.pdf [Accessed 19 August 2022].

OATES, J., BURRELL, C., EBRAHIM, S., TAYLOR, J., VEITCH, P. & BRANDON, T. 2020. Responsibility as professional leadership and decision making: Interviews with non-medical Responsible Clinicians. *International Journal of Law and Psychiatry,* 71**,** 101575.

PALEY, J. & LILFORD, R. 2011. Qualitative methods: an alternative view. *British Medical Journal,* 342**,** 956-958.

PAPADAKIS, M. A., HODGSON, C. S., TEHERANI, A. & KOHATSU, N. D. 2004. Unprofessional behavior in medical school is associated with subsequent disciplinary action by a state medical board. *Academic Medicine,* 79**,** 244-249.

PAPROCKI, C. M. 2014. When personal and professional values conflict: Trainee perspectives on tensions between religious beliefs and affirming treatment of LGBT clients. *Ethics & Behavior,* 24**,** 279-292.

PARKS, L. & GUAY, R. P. 2009. Personality, values, and motivation. *Personality and Individual Differences,* 47**,** 675-684.

PATTERSON, F. 2018. Designing and Evaluating Selection and Recruitment in Healthcare. *In:* PATTERSON, F. & ZIBARRAS, L. (eds.) *Selection and Recruitment in the Healthcare Professions.* Cham, Switzerland: Palgrave Macmillan, pp.1-26.

PATTERSON, F., ASHWORTH, V., ZIBARRAS, L., COAN, P., KERRIN, M. & O'NEILL, P. 2012. Evaluations of situational judgement tests to assess non-academic attributes in selection. *Medical Education,* 46**,** 850-868.

PATTERSON, F., COUSANS, F., EDWARDS, H., ROSSELLI, A., NICHOLSON, S. & WRIGHT, B. 2017. The predictive validity of a text-based situational judgment test in undergraduate medical and dental school admissions. *Academic Medicine,* 92**,** 1250-1253.

PATTERSON, F. & DRIVER, R. 2018. Situational Judgement Tests (SJTs). *In:* PATTERSON, F. & ZIBARRAS, L. (eds.) *Selection and Recruitment in the Healthcare Professions.* Cham, Switzerland: Palgrave Macmillan, pp.79-112.

PATTERSON, F., GALBRAITH, K., FLAXMAN, C. & KIRKPATRICK, C. M. 2019. Evaluation of a situational judgement test to develop non-academic skills in pharmacy students. *American Journal of Pharmaceutical Education,* 83, 7074.

PATTERSON, F., KERRIN, M., BARON, H. & LOPES, S. 2015. *Exploring the relationship between general practice selection scores and MRCGP examination performance. Final Report to General Medical Council*. [Online] Available: https://www.gmc-uk.org/-/media/documents/Exploring_the_Relationship_between_Recruitment_Scores_and_MRCGP_Examination_Performance_v5.4.pdf_63533914.pdf [Accessed 19 August 2022].

PATTERSON, F., KNIGHT, A., DOWELL, J., NICHOLSON, S., COUSANS, F. & CLELAND, J. 2016a. How effective are selection methods in medical education? A systematic review. *Medical Education,* 50**,** 36-60.

PATTERSON, F., PRESCOTT-CLEMENTS, L., ZIBARRAS, L., EDWARDS, H., KERRIN, M. & COUSANS, F. 2016b. Recruiting for values in healthcare: a preliminary review of the evidence. *Advances in Health Sciences Education: Theory and Practice.,* 21**,** 859-881.

PATTERSON, F., ROBERTS, C., HANSON, M., HAMPE, W., EVA, K., PONNAMPERUMA, G., MAGZOUB, M., TEKIAN, A. & CLELAND, J. 2018. 2018 Ottawa consensus statement: Selection and recruitment to the healthcare professions. *Medical Teacher,* 40**,** 1091-1101.

PATTERSON, F., ROWETT, E., HALE, R., GRANT, M., ROBERTS, C., COUSANS, F. & MARTIN, S. 2016c. The predictive validity of a situational judgement test and multiple-mini interview for entry into postgraduate training in Australia. *BMC medical Education,* 16**,** 87.

PATTERSON, F., ZIBARRAS, L. & ASHWORTH, V. 2016d. Situational Judgement Tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. *Medical Teacher,* 38**,** 3-17

PEEK, H. 2014. Psychiatry and professionalism in the digital age. *Psychiatric Times,* 31**,** 40G.

PEEK, H. S., RICHARDS, M., MUIR, O., CHAN, S. R., CATON, M. & MACMILLAN, C. 2015. Blogging and social media for mental health education and advocacy: a review for psychiatrists. *Current Psychiatry Reports,* 17**,** 88.

PEETERS, H. & LIEVENS, F. 2005. Situational judgment tests and their predictiveness of college students' success: The influence of faking. *Educational and Psychological Measurement,* 65**,** 70-89.

PELTO-PIRI, V., ENGSTROM, K. & ENGSTROM, I. 2012. The ethical landscape of professional care in everyday practice as perceived by staff: A qualitative content analysis of ethical diaries written by staff in child and adolescent psychiatric in-patient care. *Child and Adolescent Psychiatry and Mental Health,* 6**,** 18.

PÉREZ-FUENTES, M.D.C., MOLERO JURADO, M.D.M., MARTOS MARTÍNEZ, Á. & GÁZQUEZ LINARES, J.J. 2019. Analysis of the risk and protective roles of work-related and individual variables in burnout syndrome in nurses. *Sustainability*, 11, 5745.

PILLOW, W. 2003. Confession, catharsis, or cure? Rethinking the uses of reflexivity as methodological power in qualitative research. *International Journal of Qualitative Studies in Education,* 16**,** 175-196.

PINTRICH, P., SMITH, D., GARCIA, T. & MCKEACHIE, W. 1991. *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ).* [Online] https://files.eric.ed.gov/fulltext/ED338122.pdf [Accessed 19 August 2022].

PLOYHART, R. E. & EHRHART, M. G. 2003. Be careful what you ask for: Effects of response instructions on the construct validity and reliability of situational judgment tests. *International Journal of Selection and Assessment,* 11**,** 1-16.

POOLE, P. J., MORIARTY, H. J., WEARN, A. M., WILKINSON, T. J. & WELLER, J. 2009. Medical student selection in New Zealand: looking to the future. *New Zealand Medical Journal,* 122**,** 88-100.

POOLE, R. & BHUGRA, D. 2008. Should psychiatry exist? *International Journal of Social Psychiatry,* 54**,** 195-196.

POPAY, J., ROBERTS, H., SOWDEN, A., PETTICREW, M., ARAI, L., RODGERS, M., BRITTEN, N., ROEN, K. & DUFFY, S. 2006. *Guidance on the conduct of narrative synthesis in systematic reviews: a product from the ESRC methods programme.* [Online]. Available: http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=6&ved=0CF4QFjAF&url=http%3A%2F%2Fciteseerx.ist.psu.edu%2Fviewdoc%2Fdownload%3Fdoi%3D10.1.1.178.3100%26rep%3Drep1%26type%3Dpdf&ei=3QMQUIqxK-GK0AXfgoGoBA&usg=AFQjCNGrW1dB7FBraR9t3ITS0FjiWqqg-A [Accessed 19 August 2022].

PROJECT OF THE ABIM FOUNDATION, A. A. F., AND EUROPEAN FEDERATION OF INTERNAL MEDICINE 2002. Medical professionalism in the new millennium: a physician charter. *Annals of Internal Medicine,* 136**,** 243-246.

QUINTERO, A. J., SEGAL, L. S., KING, T. S. & BLACK, K. P. 2009. The personal interview: assessing the potential for personality similarity to bias the selection of orthopaedic residents. *Academic Medicine,* 84**,** 1364-1372.

RANDALL, G. E. & KINDIAK, D. H. 2008. Deprofessionalization or postprofessionalization? Reflections on the state of social work as a profession. *Social Work in Health Care,* 47**,** 341-354.

RANKIN, B., 2013. Emotional intelligence: Enhancing values-based practice and compassionate care in nursing. *Journal of advanced nursing*, 69, 2717-2725.

REDLEY, M., CLARE, I.C., LUKE, L. & HOLLAND, A.J., 2010. Mental Capacity Act (England and Wales) 2005: The emergent independent mental capacity advocate (IMCA) service. *British Journal of Social Work*, 40, 1812-1828.

REES, C. E. & KNIGHT, L. V. 2007. Viewpoint: The Trouble with Assessing Students' Professionalism: Theoretical Insights from Sociocognitive Psychology. *Academic Medicine,* 82**,** 46-50.

REICH, J. & SCHATZBERG, A. 2014. An empirical data comparison of regulatory agency and malpractice legal problems for psychiatrists. *Annals of Clinical Psychiatry,* 26**,** 91-96.

RIOLI, G., FERRARI, S., HENDERSON, C., VANDELLI, R., GALLI, G., MINARINI, A. & GALEAZZI, G. M. 2020. Users' choice and change of allocated primary mental health professional in community-based mental health services: a scoping review. *International Journal of Social Psychiatry,* 66**,** 373-381.

REISER, S., SCHACHT, L., THOMM, E., FIGALIST, C., JANSSEN, L., SCHICK, K., DÖRFLER, E., BERBERAT, P.O., GARTMEIER, M. & BAUER, J. 2022. A video-based situational judgement test of medical students' communication competence in patient encounters: Development and first evaluation. *Patient Education and Counseling,* 105, 1283-1289.

ROBERTS, L. W. 2009. Professionalism in psychiatry: a very special collection. *Academic Psychiatry,* 33**,** 429-430.

ROBERTS, L. W., JOHNSON, M. E., BREMS, C. & WARNER, T. D. 2006. Preferences of Alaska and New Mexico psychiatrists regarding professionalism and ethics training. *Academic Psychiatry,* 30**,** 200-204.

ROBERTS, L. W. & TERMUEHLEN, G. 2013. (Honest) letters of recommendation. *Academic Psychiatry,* 37**,** 55-9.

ROBERTSON, M. & WALTER, G. 2007. Overview of psychiatric ethics I: Professional ethics and psychiatry. *Australasian Psychiatry,* 15**,** 201-206.

ROGERS, A. 2009. Dare we do away with professionalism? *Therapy Today,* 20**,** 26-29.

T. ROSMAN, T., MAYER, A-K. & KRAMPEN, G. 2016. Measuring Psychology Students' Information-Seeking Skills in a Situational Judgment Test Format, Construction and Validation of the PIKE-P Test. *European Journal of Psychological Assessment,* 32**,** 220-229.

RÖSSLER, W. 2012. Stress, burnout, and job dissatisfaction in mental health workers. *European Archives Psychiatry Clinical Neuroscience,* 262**,** S65–S69.

SABIN, J. E. & HARLAND, J. C. 2017. Professional ethics for digital age psychiatry: boundaries, privacy, and communication. *Current Psychiatry Reports,* 19**,** 55.

SALE, J. E., LOHFELD, L. H. & BRAZIL, K. 2002. Revisiting the quantitative-qualitative debate: Implications for mixed-methods research. *Quality and Quantity,* 36**,** 43-53.

SALKIND, N. J. 2010. Encyclopedia of research design. *Positivism.* CA: SAGE Publications, Inc., Thousand Oaks.

SANDERS, K., SERVIS, M. & BOLAND, R. 2014. The four general competencies. *Academic Psychiatry,* 38**,** 268-274.

SATTERWHITE, R. C., SATTERWHITE, W. M. & ENARSON, C. 2000. An ethical paradox: the effect of unethical conduct on medical students' values. *Journal of Medical Ethics,* 26**,** 462-465.

SCHMIDT, F. L. & HUNTER, J. E. 1998. The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin,* 124**,** 262-274.

SCHMIT, M. J. & RYAN, A. M. 1992. Test-taking dispositions: A missing link? *Journal of Applied Psychology,* 77**,** 629-637.

SCHREIBER, G., AVISSAR, S., JOTKOWITZ, A. & HALPERIN, D. 2016. The patient-psychiatrist relationship on the axis of the other and the same. *Psychiatric Quarterly***,** 88, 675-686.

SCHWARTZ, A. C., KOTWICKI, R. J. & MCDONALD, W. M. 2009. Developing a modern standard to define and assess professionalism in trainees. *Academic Psychiatry,* 33**,** 442-450.

SCOTT JOHNSON, R., CHIU, L. L. & CZELUSTA, K. L. 2015. For residents, technology can put professionalism and reputation at risk. *Current Psychiatry,* 14**,** e3-e4.

SHARMA, S., GANGOPADHYAY, M., AUSTIN, E. & MANDAL, M. K. 2013. Development and validation of a situational judgment test of emotional intelligence. *International Journal of Selection and Assessment,* 21**,** 57-73.

SIMS, D. 2011. Reconstructing professional identity for professional and interprofessional practice: a mixed methods study of joint training programmes in learning disability nursing and social work. *Journal of Interprofessional Care,* 25**,** 265-71.

SMITH, D. T. & TIFFIN, P. A. 2018. Evaluating the validity of the selection measures used for the UK's foundation medical training programme: a national cohort study. *BMJ Open,* 8**,** p.e021918.

SMITH, J. K. & HESHUSIUS, L. 1986. Closing down the conversation: The end of the quantitative-qualitative debate among educational inquirers. *Educational Researcher,* 15**,** 4-12.

SMITH, P. C. & KENDALL, L. M. 1963. Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology,* 47**,** 149-155.

Social Work England. 2022. *A guide to fitness to practise* [Online]. Available: https://www.socialworkengland.org.uk/concerns/fitness-to-practise-guide/ [Accessed 19 November 2022].

SONG, X., JIANG, N., LI, H., DING, N. & WEN, D. 2021. Medical professionalism research characteristics and hotspots: a 10-year bibliometric analysis of publications from 2010 to 2019. *Scientometrics,* 126**,** 8009-8027.

SOTO, C. J. & JOHN, O. P. 2017. Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality,* 68**,** 69-81.

SOX, H. C. 2007. The ethical foundations of professionalism: a sociologic history. *Chest,* 131**,** 1532-1540.

SPILSBURY, K., THOMPSON, C., BLOOR, K., DALE, V., DEVI, R., JACKSON, C., MCCAUGHAN, D., SIMPSON, A. & MANNION, R., 2022. *Values Based Recruitment: What works, for whom, why, and in what circumstances?.* Available: https://eprints.whiterose.ac.uk/190125/1/PR-R10-0514-14002_VBR%20FINAL%20REPORT_20%20May%202022.pdf [Accessed 9 December 2022].

STACEY, G. & PEARSON, M. 2020. An inductive content analysis of formative feedback given by lived experience assessors in pre-registration mental health nurse education. *The Journal of Mental Health Training, Education and Practice,* 15**,** 33-42.

STAFFORD-BROWN, J. & PAKENHAM, K. I. 2012. The effectiveness of an ACT informed intervention for managing stress and improving therapist qualities in clinical psychology trainees. *Journal of Clinical Psychology,* 68**,** 592-613.

STEDMAN, J., HATCH, J. & SCHOENFELD, L. 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**,** 339–345.

STEMIG, M. S., SACKETT, P. R. & LIEVENS, F. 2015. Effects of organizationally endorsed coaching on performance and validity of situational judgment tests. *International Journal of Selection and Assessment,* 23**,** 174-181.

STEFFEN, E.M., TIMOTIJEVIC, L. & COYLE, A. 2020. A qualitative analysis of psychosocial needs and support impacts in families affected by young sudden cardiac death: The role of community and peer support. *European Journal of Cardiovascular Nursing*, 19, 681-690.

STEPHENSON, J. 2014. *NHS England to rollout '6Cs' nursing values to all health service staff* [Online]. Available: https://www.nursingtimes.net/roles/nurse-managers/exclusive-6cs-nursing-values-to-be-rolled-out-to-all-nhs-staff/5070102.article [Accessed 15 September 2018].

STRATTA, E. C., RIDING, D. M. & BAKER, P. 2016. Ethical erosion in newly qualified doctors: perceptions of empathy decline. *International Journal of Medical Education,* 7**,** 286-292.

TALBOTT, J. A. & MALLOTT, D. B. 2006. Professionalism, medical humanism, and clinical bioethics: The new wave-does psychiatry have a role? *Journal of Psychiatric Practice,* 12**,** 384-390.

TAVAKOL, S., DENNICK, R. & TAVAKOL, M. 2012. Medical students' understanding of empathy: a phenomenological study. *Medical Education,* 46**,** 306-316.

TEES ESK AND WEAR VALLEYS NHS FOUNDATION TRUST. 2021. *Services* [Online]. Available: https://www.tewv.nhs.uk [Accessed 2 June 2021].

THE WORLD BANK GROUP. 2018. *Data: World Bank Country and Lending Groups* [Online]. Available:

https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [Accessed 1 February 2017].

THORNDIKE, R. L. 1947. Research problems and techniques (No. 3). US Government Printing Office.

TIFFIN, P. & CARTER, M. 2015. *Understanding the measurement model of the UKCAT situational judgement test: summary report.* [Online] Available: https://www.ucat.ac.uk/media/1183/understanding-the-measurement-model-of-the-ukcat-sjt.pdf [Accessed 19 August 2022].

TIFFIN, P. A., PATON, L. W., O'MARA, D., MACCANN, C., LANG, J. W. & LIEVENS, F. 2020. Situational judgement tests for selection: Traditional vs construct-driven approaches. *Medical Education,* 54**,** 105-115.

TIFFIN, P. A., SANGER, E., SMITH, D. T., TROUGHTON, A. & PATON, L. P. 2022. Situational judgement test performance and subsequent misconduct in medical students. *Medical Education***,** 56, 754-763.

TORRACO, R. J. 2005. Writing integrative literature reviews: Guidelines and examples. *Human Resource Development Review,* 4**,** 356-367.

TOWLE, A., BAINBRIDGE, L., GODOLPHIN, W., KATZ, A., KLINE, C., LOWN, B., MADULARU, I., SOLOMON, P. & THISTLETHWAITE, J. 2010. Active patient involvement in the education of health professionals. *Medical Education,* 44**,** 64-74.

TRIGGLE, N. 2018. *The history of the NHS in charts* [Online]. Available: https://www.bbc.co.uk/news/health-44560590 [Accessed 2 November 2021].

TRINKOFF, A. M. & STORR, C. L. 1998. Substance use among nurses: differences between specialties. *American Journal of Public Health,* 88**,** 581-585.

UKCAT CONSORTIUM. 2019. *Practice Tests and Question Banks* [Online]. Available: https://www.ucat.ac.uk/prepare/practice-tests [Accessed 19 July 2022].

VAN DE CAMP, K., VERNOOIJ-DASSEN, M., GROL, R. & BOTTEMA, B. 2004. How to conceptualize professionalism: a qualitative study. *Medical Teacher,* 26**,** 696-702.

VARPIO, L., AJJAWI, R., MONROUXE, L., O'BRIEN, B. & REES, C. 2017. Shedding the cobra effect: problematising thematic emergence, triangulation, saturation and member checking. *Medical Education,* 51**,** 40–50.

WAGNER, P., HENDRICH, J., MOSELEY, G. & HUDSON, V. 2007. Defining medical professionalism: a qualitative study. *Medical Education,* 41**,** 288-294.

WAGNER, R. K. & STERNBERG, R. J. 1985. Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of Personality and Social Psychology,* 49**,** 436-458.

WAGONER, N. 2006. Admission to medical school: selecting applicants with the potential for professionalism. *In:* STERN, D. (ed.) *Measuring Medical Professionalism.* Oxford: Oxford University Press, pp.235-264.

WATT, A., CAMERON, A., STURM, L., LATHLEAN, T., BABIDGE, W., BLAMEY, S., FACEY, K., HAILEY, D., NORDERHAUG, I. & MADDERN, G. 2008. Rapid versus full

systematic reviews: Validity in clinical practice? *ANZ Journal of Surgery,* 78**,** 1037-1040.

WEAR, D., AULTMAN, J. M., VARLEY, J. D. & ZARCONI, J. 2006. Making fun of patients: medical students' perceptions and use of derogatory and cynical humor in clinical settings. *Academic Medicine,* 81**,** 454-462.

WEAR, D. & KUCZEWSKI, M. G. 2004. The professionalism movement: can we pause? *American Journal of Bioethics,* 4**,** 1-10.

WEBSTER, E. S., PATON, L. W., CRAMPTON, P. E. & TIFFIN, P. A. 2020. Situational judgement test validity for selection: A systematic review and meta-analysis. *Medical Education,* 54**,** 888-902.

WEEKLEY, J. A. & JONES, C. 1999. Further studies of situational tests. *Personnel Psychology,* 52**,** 679-700.

WEEKLEY, J. A. & PLOYHART, R. E. 2005. Situational judgment: Antecedents and relationships with performance. *Human Performance,* 18**,** 81-104.

WEEKLEY, J. A. & PLOYHART, R. E. 2006. An Introduction to Situational Judgment Testing. *In:* WEEKLEY, J. A. & PLOYHART, R. (eds.) *Situational judgment tests: Theory, measurement, and application.* New Jersey: Lawrence Erlbaum Associates, Inc, 1-10.

WEEKLEY, J. A., PLOYHART, R. E. & HOLTZ, B. C. 2006. On the development of situational judgment tests: Issues in item development, scaling, and scoring. *In:* WEEKLEY, J. A. & PLOYHART, R. E. (eds.) *Situational judgment tests: Theory, measurement, and application.* New Jersey: Lawrence Erlbaum Associates, Inc, 157-182.

WEIDEN, P. & GLAZER, W. 1997. Assessment and treatment selection for "revolving door" inpatients with schizophrenia. *Psychiatric Quarterly,* 68**,** 377-392.

WEINER, B., PERRY, R. P. & MAGNUSSON, J. 1988. An attributional analysis of reactions to stigmas. *Journal of Personality and Social Psychology,* 55**,** 738-748.

WENG, Q. D., YANG, H., LIEVENS, F. & MCDANIEL, M. A. 2018. Optimizing the validity of situational judgment tests: The importance of scoring methods. *Journal of Vocational Behavior,* 104**,** 199-209.

WERNIMONT, P. F. & CAMPBELL, J. P. 1968. Signs, samples, and criteria. *Journal of Applied Psychology,* 52**,** 372-376.

WHETZEL, D. L. & MCDANIEL, M. A. 2009. Situational judgment tests: An overview of current research. *Human Resource Management Review,* 19**,** 188-202.

WHETZEL, D. L., MCDANIEL, M. A. & NGUYEN, N. T. 2008. Subgroup differences in situational judgment test performance: A meta-analysis. *Human Performance,* 21**,** 291-309.

WHETZEL, D. L., SULLIVAN, T. S. & MCCLOY, R. A. 2020. Situational Judgment Tests: An Overview of Development Practices and Psychometric Characteristics. *Personnel Assessment and Decisions,* 6**,** Article 1.

WHITE, J., BROWNELL, K., LEMAY, J. F. & LOCKYER, J. M. 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.

WHITTEMORE, R. & KNAFL, K. 2005. The integrative review: updated methodology. *Journal of Advanced Nursing,* 52**,** 546-553.

WISE, E. H. 2008. Competence and scope of practice: ethics and professional development. *Journal of Clinical Psychology,* 64**,** 626-637.

WOODRUFF, J., ANGELOS, P. & VALAITIS, S. 2008. Medical professionalism: one size fits all? *Perspectives in Biology and Medicine,* 51**,** 525-534.

WORLD MEDICAL ASSOCIATION 2013. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *The Journal of the American Medical Association,* 310**,** 2191-2194.

WORTHINGTON, A., ROONEY, P. & HANNAN, R. 2013. *The Triangle of Care. Carers Included: A Guide to Best Practice in Mental Health in England. Second Edition.* [Online] Available: https://carers.org/downloads/resources-pdfs/triangle-of-care-england/the-triangle-of-care-carers-included-second-edition.pdf [Accessed 19 August 2022].

WU, H. C., TANG, I. C., LIN, W. I. & CHANG, L. H. 2012. Professional values and attitude of psychiatric social workers toward involuntary hospitalization of psychiatric patients. *Journal of Social Work,* 13**,** 419-434.

YAHYAVI, S. T., HOOBEHFEKR, S. & TABATABAEE, M. 2021. Exploring the hidden curriculum of professionalism and medical ethics in a psychiatry emergency department. *Asian Journal of Psychiatry,* 66**,** 102885.

YEPES-RIOS, M., DUDEK, N., DUBOYCE, R., CURTIS, J., ALLARD, R. J. & VARPIO, L. 2016. The failure to fail underperforming trainees in health professions education: A BEME systematic review: BEME Guide No. 42. *Medical Teacher,* 38**,** 1092-1099.

ZIJLSTRA-SHAW, S., ROBINSON, P. & ROBERTS, T. 2012. Assessing professionalism within dental education; the need for a definition. *European Journal of Dental Education,* 16**,** e128-136.

Appendix A: Great expectations publication

Great expectations: views and perceptions of professionalism amongst mental health services staff, patients and carers

Lauren M. E. Aylott, Paul A. Tiffin, Sally Brown & Gabrielle M. Finn

Published online: 10 Sep 2020.

Submit your article to this journal

Article views: 553

View related articles

View Crossmark data

Taylor & Francis
Taylor & Francis Group

RESEARCH AND EVALUATION

a OPEN ACCESS | Check for updates

Great expectations: views and perceptions of professionalism amongst mental health services staff, patients and carers

Lauren M. E. Aylott[a] (ID), Paul A. Tiffin[a,b] (ID), Sally Brown[c*] (ID) and Gabrielle M. Finn[a‡] (ID)

[a]Health Professions Education Unit, Hull York Medical School, University of York, York, England; [b]Department of Health Sciences, University of York, York, England; [c]Edinburgh Napier University, Edinburgh, Scotland

ABSTRACT

Background: Numerous studies have explored the concept of 'professionalism' in medicine, yet little attention has been paid to the concept in a mental health services context.

Aims: This study sought to determine how the lived experience of patients, carers and healthcare professionals in mental health services align with medically defined, generic, professionalism standards.

Method: Interviews and focus groups were conducted with patients, carers, nurses, occupational therapists, psychiatrists and psychologists. A framework analysis approach was used to analyse the data, based on the 'Improving Selection to the Foundation Programmes' Professional Attributes Framework.

Results: Fifty-six individuals participated. Data aligned to all nine attributes of the Professional Attributes Framework, however the expectations within each attribute varied from that originally cited. A tenth attribute was devised during the process of analysis; Working with Carers. This attribute acknowledges the need to liaise with, and support carers in mental health services. Situational examples included both online and offline behaviours and the topic of 'black humour' emerged.

Conclusions: Compared to a conventional medical definition of professionalism, additional themes and differing emphases were observed for mental health and learning disability services. These findings should be used to inform the teaching and evaluation of professionalism, especially for staff pursuing mental health service careers.

Introduction

What is Professionalism?

Professionalism is a multidimensional concept (van de Camp et al., 2004). Definitions have tautologically included "upholding professional values, exhibiting professional behaviours or demonstrating professional attitudes" (Aguilar et al., 2011). The context-dependent nature of professionalism has previously been highlighted (Aylott et al., 2019; Rees & Knight, 2007). Yet, whilst many studies have sought to explore 'professionalism' in medicine, there is a dearth of studies focusing on the concept in mental health services (MHS; Aylott et al., 2019).

Mental health problems are, reportedly, the single largest source of burden of disease in the UK; nonetheless, services are in receipt of inadequate funding, an understaffed workforce, and insufficient training (British Medical Association, 2017). When services are not resourced adequately, the likelihood of unprofessional behaviour may increase.

Each healthcare area has their own regulatory frameworks, as do each of the professions; nonetheless, there are nuances to MHS that warrant further exploration regarding the concept of professionalism. For example, patients may experience impaired mental capacity, rendering them less autonomous and vulnerable to exploitation. This adds complexity to the ethical and professional issues that face mental health professionals.

Professionalism in mental health services

A recent systematic review provided an operational definition of professionalism for MHS (Aylott et al., 2019). Practitioners are described as embodiments of their profession, possessing intrapersonal, interpersonal and working professionalism. A scarcity of patient presence was observed in the literature. Therefore, this study explored the perceptions of professionals, patients and carers regarding the

CONTACT Lauren M. E. Aylott ✉ Lmea500@york.ac.uk 🏛 Health Professions Education Unit, Hull York Medical School, University of York, John Hughlings Jackson Building, Heslington, York YO10 5DD, England
*Sally Brown has moved affiliation during the peer review process. Sally now works in the School of Health in Social Science, College of Arts, Humanities and Social Sciences, University of Edinburgh.
‡Gabrielle M. Finn has since commenced work in the Faculty of Biology, Medicine and Health, The University of Manchester. Gabrielle holds an Honorary position with Hull York Medical School.

professional attitudes and behaviours experienced in MHS. Having reviewed the literature, the authors determined that the findings would be best analysed against the professional attributes framework (PAF), developed as part of the 'improving selection to the foundation programme' project, commissioned by the Medical Schools Council (2011). This was considered a suitable framework, as most of the professionalism literature regarding MHS focuses on psychiatry (a medical profession). The PAF also helped inform the development of the topic guide. The research questions were: (1) what are patients, carers and professionals experiences of, and perceptions of professionalism in a mental health setting? (2) how does this experience align, if at all, with medically defined, generic, professional standards and attributes?

Methods

Ethics

All procedures were approved by the Health Sciences Research Governance Committee at the University of York and a favourable ethical opinion was obtained from London–Camden & Kings Cross Research Ethics Committee (18/LO/0630).

Design

Focus groups were undertaken with carers and professionals. Individual semi-structured interviews were undertaken with patients, as the ethics committee expressed that this approach would raise fewer ethical concerns than the use of focus groups.

Recruitment

The study was advertised in an NHS Trust using their electronic news bulletin and intranet. Emails were sent via professional leads, and flyers were placed in community centres, and distributed via a patient and public newsletter. A carers' network was also approached.

Participation was voluntary; however, patients and carers were offered a £20 gift voucher to reimburse them for their time and travel. Purposive sampling was used. Too much heterogeneity within each focus group could inhibit the discussion (Freeman, 2006); thus, separate focus groups were conducted with psychiatrists, psychologists, nurses, occupational therapists (OTs) and carers. Patients that had accessed MHS, within the last two years, participated in one to one interviews. Having had the objective of the study explained, written consent was obtained from all participants.

Data Collection

Interview questions were established based on the critical incident technique (Flanagan, 1954) and prior research on professionalism (Burford et al., 2014; Medical Schools Council, 2011). The schedule was piloted, confirming that the questions and prompts were fit for purpose (see Table 1). The lead author facilitated all interviews and focus groups; four focus groups were co-facilitated by a co-author. Interviews and focus groups were digitally recorded and transcribed verbatim. Member checking was used during focus groups and interviews to confirm the authors' interpretation of the data.

Data analysis

Data were managed using NVivo. As the researchers had apriori themes in mind, framework analysis (Ritchie et al., 2003; Ritchie & Spencer, 1994) was employed (see Figure 1). During the familiarisation stage, open coding was utilised; it was determined that the data fit the PAF (Medical Schools Council, 2011). Data were subsequently indexed against all nine attributes of the PAF. Codes and subsequent themes were created to portray an accurate description of the data.

Analysis was predominantly performed by two authors (LA, SB). An iterative process was used by SB to review a transcript against the codes and themes created by LA. Data were discussed, and codes amended through negotiation, resulting in themes and subthemes; hereby referred

Table 1. Topic guide used to generate discussion on professionalism in mental health services.

Topic stem	Prompts
Professionalism	What does professionalism, in general, mean to you?
	What does good 'professionalism' look like, in general?
	What does 'unprofessional' behaviour look like, in general?
	What are the most important aspects of 'professionalism' for staff working in 'mental health services'?
	What skills are important when working in mental health services?
	What interpersonal attributes are important when working in mental health services?
	What values are expected of staff working in mental health services?
Professional Behaviours	Is anyone able to discuss an example of professional or unprofessional behaviour, that they have recently observed, especially good or bad practice?
	What were the circumstances leading up to the incident?
	Would you please describe the professional's reaction?
	What did the person(s) do that was (un)professional?
	• Has anyone else got any other views regarding this behaviour(s)?
	• How else could the professional have responded?
Professional Dilemmas	Has anyone observed any dilemmas that a professional has recently encountered?
	What were the circumstances?
	Did the professional(s) behave in a professional or unprofessional manner?
	What did the professional do?
	• What was it about the behaviour that was (un)professional?

Figure 1. Stages of Framework Analysis (adapted from Ritchie & Spencer, 1994).

to as the professional attributes and expectations. To improve the validity of the analysis, all authors reviewed a sample of coding.

Results

In total, 6 focus groups and 13 interviews were conducted between August 2018 and January 2019. Participants included 18 men, 36 women, and 2 individuals that referred to their gender as 'other'. The total sample was 56, including 7 psychiatrists, 7 psychologists, 10 nurses, 10 OTs, 7 carers, and 15 patients. Ages ranged from 21 to 86 years. Professionals worked in a range of specialties, across inpatient and community settings.

Data aligned to all nine attributes, however the expectations within each attribute varied to that originally cited. A tenth attribute was devised during analysis; 'Working with Carers'. The ten professional attributes, which became the coding framework are described in turn here, with illustrative quotes provided. Each attribute includes a brief description to reflect its overall nature. As each attribute has multiple expectations, it is not possible to describe each one; an overview of this structure is reported within Table 2.

Commitment to Professionalism

Individuals should be committed to honouring their profession by adhering to guidelines and challenging poor practice. Professionals must have integrity and be a responsible practitioner

Challenging poor practice can make professionals unpopular, but they have "a duty" to challenge. Social media presented issues, specifically with regards to online dating and platforms such as Facebook, but the work and personal life dichotomy was made clear.

> ...it's about having clear boundaries as well isn't it, so I've got a Twitter account for purely professional stuff ... but then Facebook, I have strict settings on that ... (OT)

> I am a single person, but I have got an issue for example with internet dating as well because of that status you have got [as a psychiatrist], I don't do it .... I think there are some unwritten rules you are supposed to follow. (Psychiatrist)

Coping with Pressure

Practitioners must utilise their clinical judgement, particularly in times of uncertainty and ambiguity. Professionals must have resilience and be able to de-escalate situations when others are experiencing distress

Participants were aware of the need for professionals to separate themselves from the work that they do, to not let this impinge upon their own wellbeing.

> Professionalism for me is walking out of here and onto the next thing with a fresh mind, a fresh view ...you have got to detach. (Patient)

Acknowledging the need for de-escalation, a carer highlighted that some staff de-escalate situations better than others.

> I have seen physical restraint used detrimentally to the point even when people were saying 'oh should we do your hair' while they are restraining this person on the floor, whereas other staff would be able to talk to that person and dissolve the situation (Carer)

Humour was noted as a means of making light, defusing stress and normalising situations. An example was provided with respect to a suicidal patient who frequently swallowed a specific form of vegetation. After numerous incidents within a shift a staff member made a blasé remark, '*I can't wait for spring, when there are no [vegetation] around.*' Humour was viewed negatively in certain circumstances, such as when one is critical of a patient.

Effective Communication

Practitioners must communicate effectively, using both verbal and non-verbal communication. Professionals should have the ability to build rapport with patients, and validate the thoughts and feelings of others

Banter can be used to build rapport; however, it was recognised that this can go wrong very quickly, particularly "*if someone feels targeted, especially where they've got trauma histories, where they've been humiliated, bullied, and all the rest.*" (Psychologist)

Table 2. Professional attributes and expectations of staff working in mental health services (adapted from Medical Schools Council, 2011).

Professional attribute	Associated expectations
1. Commitment to Professionalism	Adheres to guidelines; Behaves according to expectations; Challenges the system accordingly; Challenges poor practice; Displays a commitment to the role; Is trustworthy and has integrity; Maintains confidentiality; Is a responsible practitioner; Possesses confidence and courage; Demonstrates awareness of ethical issues; Upholds the profession's and organisation's reputation; Uses Social Media appropriately.
2. Coping with Pressure	Checks the facts of a case; Utilises clinical judgement accordingly; Utilises de-escalation techniques appropriately; Possesses resilience and manages own wellbeing; Remains calm and in control of situations.
3. Effective Communication, with Patients, Carers and Colleagues	Utilises an appropriate style of communication; Advocates for patients when needed; Builds rapport with patients and is personable; Communicates effectively with colleagues; Is open and honest, whilst communicating in a proactive manner; Communicates sensitively, taking context into consideration; Listens effectively; Observes accordingly; Understands non-verbal communication; Utilises appropriate non-verbal communication; Validates the thoughts and feelings of others.
4. Learning and Professional Development	Is accepting of feedback; Applies knowledge and learning to practice; Utilises supervision; Has the appropriate knowledge and skills for the role; Undertakes continuing professional development and learns from practice; Possesses the relevant qualifications for the role; Undertakes research.
5. Organisation and Planning	Is efficient in the role; Maintains accurate records; Reads patient's case notes; Is able to make appropriate use of limited resources; Manages time accordingly; Wears appropriate attire at work.
6. Patient Focus	Attends to patients' physical healthcare needs; Acts as a human interface for the organisation; Builds therapeutic relationships; Demonstrates compassion; Contains the emotions of others; Does not impose own values on patients; Is friendly, but maintains an appropriate professional distance; Is altruistic and possesses humility; Is approachable; Is empathic and understands the impact of mental illness; Is genuine, honest and fulfils promises; Is non-judgemental; Provides reassurance; Maintains Safety; Treats Patients with respect; Utilises both a person-centered and recovery-focused approach.
7. Problem Solving and Decision Making	Is able to reason in an abstract manner; Helps patients to problem solve; Makes appropriate decisions based on all the relevant information; Understands problems from a wider perspective; Uses their initiative and adapts practice, to meet a patient's needs.
8. Self Awareness and Insight	Acknowledges one limits; Discloses appropriate amounts of information about self; Has self-awareness and reflects on practice; Recognises and maintains appropriate boundaries.
9. Working Effectively as Part of a Team	Able to identify and utilise the most appropriate person for a task; Acts as a positive role model; Maintains appropriate relationships with colleagues; Maintains consistency with colleagues; Supports colleagues; Works with other teams; Works effectively with colleagues.
10. Working with carers	Assesses a carer's motives accordingly; Involves carers and makes use of their expertise; Supports and validates the carer's perspective.

Learning and Development

Practitioners must possess the appropriate knowledge and skills for their role, utilising professional development opportunities. Professionals should accept feedback and utilise supervision accordingly

Supervision is a safe forum for discussing the difficulties one may face, particularly when struggling with patients.

> You could say the same thing, but in different contexts. And some are professional and some are not professional. (Psychologist)

Organisation and Planning

Practitioners must maintain accurate records and read case notes attentively. Professionals must effectively manage limited resources, and their time accordingly

One professional suggested that documentation was a waste of clinician time, whereas a patient commented that sometimes this must take precedence.

> All the notes we keep really are litigation proof that's all it is you know (Nurse)

> The planning of my care could depend on that, ... If it isn't recorded properly I might lose out on the appropriate care. (Patient)

Patient Focus

Practitioners must possess qualities that enable them to build therapeutic relationships with patients, such as altruism and humility. Professionals, also, must maintain an appropriate professional distance and not impose their own values on patients, delivering person-centered care

A service user indicated that, in some cases, respect is not always evident. This finding was supported by an OT's comments.

> Just the way someone speaks to you honest they speak to you like you're on their shoe. (Patient)

> when people open doors without knocking, or open curtains without getting permission ... It's low-grade stuff, but it's poor practice. (OT)

One patient discussed their anxiety at making telephone calls; noting that their nurse challenged them to do this, they felt they should be challenged in this way.

> It's finding ... something I need to do or want to do and pushing me to do that, as opposed to pushing me to do whatever they think most people who are depressed or hear voices or whatever do. (Patient)

301

Problem Solving

Practitioners must be able to reason with abstract information. Professionals also must understand problems from a wider perspective, having the ability to adapt their practice

Patients experience mental illness differently; professionals must therefore conceptualise with abstract information.

> The word schizophrenia does not describe people with schizophrenia, you know everybody's different and I think you have to … think outside the box a bit in mental health that you maybe don't have to do so much in sort of medical care. (Nurse)

Self-Awareness and Insight

Practitioners must possess self-awareness and acknowledge the limits to their competence. It is appropriate to disclose some information about self, but professionals must recognise and maintain boundaries when doing so

Self-disclosure can break down some of the boundaries in a staff–patient relationship. However, the disclosure must be appropriate for each patient.

> It also seems a bit of a balance sometimes because you expect a patient to share so much of themselves… And then you kind of go, but I won't tell you anything about me. (Psychologist)

Professionals must reflect-in-action.

> I've seen a fair few emails that maybe come across as disrespectful and unprofessional, because they're emotive emails about a topic that people have strong opinions on … I think professionalism is about taking a step back and thinking right, I'm quite annoyed about this, maybe I should send this later or ask someone else to look over it. (Psychologist)

Working Effectively as Part of a Team

Practitioners must work alongside colleagues effectively, acknowledging people's strengths and capabilities. Professionals must also be a positive role model

Discussions arose regarding negative role modelling and the detrimental impact this can have on students and newly qualified professionals.

> When they're qualified, you can't go home and have a cup of tea and I think that sometimes with students, … we set the wrong examples. (OT)

Working with Carers

Professionals must involve carers where possible, whilst adhering to the bounds of confidentiality. Carers generally want to be involved and should feel supported in their role

The study found that professionals were perceived to hide behind policy. Carers want to be supported, yet that support is not always available.

> confidentiality, he is using that against me because all he is sees is a carer … he is not making that connection between if he supports me professionally as a carer, he is also supporting my son in the future (Carer)

Carers feel left to the wayside; professionals should work with carers, where possible, to improve patient care.

the majority of carers want to be involved and want to help and get very fed up of being left in a corner. (Carer)

Discussion

Professionalism is widely cited as the basis of a social contract between professions and society (see Aylott et al., 2019). With regards to medicine, physicians can expect their core activities to be protected and governed by licensing laws (Cruess, 2006); which permit members to practice. Professions are permitted and expected to self-regulate (Cruess, 2006). However, in return, the views of the patients they serve, who place their trust in the profession, must be incorporated (Irvine, 1997). Indeed, the social contract must be regularly renegotiated (Bhugra, 2008; Bhugra & Gupta, 2010).

The study sought to explore the professional attributes desired within mental health services. The authors hoped to determine how this experience aligned to medically defined, generic, professional standards and attributes.

Main Findings

The findings demonstrate that patients', carers' and professionals' views on professionalism in MHS mostly align to medically defined, generic, professional standards and attributes; however, there are differing emphases, and carers must be factored into the equation. When working with clients exhibiting challenging behaviours, it is important that one maintains self-awareness and insight to maintain a neutral, professional stance with patients. Effective communication is paramount across settings, yet this is more challenging in MHS, where patients may be severely unwell and lack mental capacity. As highlighted previously, the use of banter can go extremely badly when used with patients that have trauma histories.

Whilst carers are not unique to mental health settings, they often play a key role in the care of many patients with mental health problems and working collaboratively with them is an essential part of modern service delivery (Cleary et al., 2006); nevertheless, the findings demonstrate that all too often, carers are feeling left to the wayside and unsupported in their role. Given that a patient's mental capacity is often impaired in MHS, carers fill a vital role, *"if you are really doing the best for your patient, you have to be able to look at their social network, particularly in mental health"* (Carer).

Comparisons with the Existing Literature

As demonstrated previously (Green et al., 2009), the study found that patients and professionals do not always place the same regard on certain professional behaviours; there were differing views between a patient and professional regarding documentation.

In comparison to the medically derived PAF, the study identified a tenth attribute; working with carers. The importance of carers being involved in the treatment of

patients with mental health problems has previously been reported in the Triangle of Care guidance document (Worthington et al., 2013). Nevertheless, carers report being shunned by professionals, who won't even talk to them, because "*the patient has not given consent for the carer to be involved*" (Carer). This study adds further evidence to support the findings of Cleary et al. (2005); over 50% of carers reported that information, such as that regarding medication, illnesses and community resources is not provided to them.

Similarities are observed between these findings and the definition of professionalism proposed by Aylott et al. (2019). The attribute 'Commitment to Professionalism' is congruous with 'Intrapersonal Professionalism'; 'Patient Focus' and other attributes align to the concept of 'Interpersonal Professionalism'; and 'Coping with Pressure' is harmonious with 'Working Professionalism'–the ability to form judgements and act accordingly, thinking critically and using reflection in action.

The importance of contextual factors and situational judgement has previously been reported by Burford et al. (2014). Whilst their study was not conducted in MHS, their findings resonate with the current study; as a psychologist highlighted, "*You could say the same thing, but in different contexts. And some are professional and some are not professional*". Similarly, the use of humour and self-disclosure is dependent on contextual, patient related factors.

Humour is viewed as acceptable amongst professionals, but professionals must be aware of their audience, and they must use humour in a safe setting. Camaraderie and banter have been reported on in a study with medical undergraduates, whereby "*camaraderie was condoned as a legitimate means by which medics can diffuse stress*" (p.819; Finn et al., 2010). Nearly a decade later, the same issues still present. This is a controversial issue, as professionals argue that humour is vital for their wellbeing, but many alluded to its context dependence and sensitivity.

The need to be a positive role model was highlighted, echoing that previously reported with the 'hidden curriculum'; there are discrepancies between what students learn in textbooks and classrooms versus what they pick up on placement (Hafferty & Franks, 1994). Concerningly, the observation of and participation in unethical conduct has been shown to result in ethical erosion overtime (Satterwhite et al., 2000). Problematic behaviour during training has also been found to predict disciplinary action in later clinical practice (Papadakis et al., 2004).

In many ways, professionalism is more noticeable by its absence than its presence. Following their evaluation of a Situational Judgement Test measuring integrity, de Leng et al. (2018) suggested that there may be a greater consensus regarding what is considered inappropriate, as opposed to appropriate behaviour. Certainly, in the current study, participants frequently referred to professionalism as about 'what not to do', as opposed to more optimal behaviours. Healthcare is becoming increasingly bureaucratic and administrative, and in a litigious society, professions are more regulated than ever. Whilst participants seemed open

about their professional and unprofessional encounters, some behaviours may not have been disclosed. This is in keeping with what John McLachlan refers to as 'Pious Platitudes' about professionalism; that practitioners, when asked to define professionalism, may respond with what they think they ought to say, rather than with what they have actually observed (Monrouxe & Rees, 2017).

Interpretation of findings

Various medical specialties suggest that refinements to the concept of professionalism are needed for this to be pertinent to their practice (Woodruff et al., 2008). For instance, it has been argued that the Physician Charter does not take account of many elements of surgery practice, that pose specific ethical and professional questions for their specialty (Jones et al., 2006). Given the nature of MHS, it is not surprising that the study identified a tenth attribute–working with carers. Carers were not directly involved in the development of the PAF. However, there are nuances in MHS that are not present in other specialties; detained patients lose their patient autonomy–one of the three fundamental principles of medical professionalism (Project of the ABIM Foundation, 2002); and patients are vulnerable (Department of Health, 2000, updated in 2015), often relying on carers to protect and advocate for them. As stated in the Triangle of Care, "*Carers are usually the first to be aware of a developing crisis … They are often best placed to notice subtle changes in the person for whom they care, and usually the first to notice the early warning signs of a relapse*" (p.7, Worthington et al., 2013).

MHS pose many challenges for professionals, as well as professionalism in general. For example, patients may commit suicide, which undoubtedly impacts upon a professional's own wellbeing. Patients, also, may be criminals, yet the professional has a duty of care and must do their best by the patient. Discussing the ever-increasing challenges, a psychiatrist commented–"*isn't our status because we do, we make good a bad system … I think we are regarded at a level that we are regarded because we don't let bad things happen*". The current study explored the expectations placed on professionals working in MHS; however, we must acknowledge the requirements of the professions themselves, including a properly funded and value-driven healthcare system (Bhugra, 2008; Cruess, 2006). The social contract places expectations on patients too, such as 'a shared responsibility' for their own health and wellbeing (Bhugra, 2008; Cruess, 2006); it is important that patients and carers work with professionals, by, for example, returning calls in a timely manner, and turning up to appointments.

Strengths and Limitations

This is the first study to explore professionalism from the views of patients, carers, and professionals in MHS. The findings are enriched by the opinions of patients and carers; a voice underrepresented in the professionalism literature. Some researchers criticise the combining of interview and

focus group data; however, Lambert and Loiselle (2008) describe how the integration of focus group and individual interview data can enhance the trustworthiness of findings. Alike Lambert and Loiselle's study (2008), not all prompts were used during focus groups to elicit the information. Nonetheless, the topic guide enabled a semi-structured approach and minimised the likelihood of interviewer bias.

The study was conducted in a large NHS Trust and there were geographical variances in where professionals and patients, had trained, worked, and/or received healthcare previously. Professionals worked across a mixture of settings and a number of nationalities were represented; it is thus the opinion of the research team that these findings are transferable to and representative of mental health issues nationally. Acknowledging that the interviewer is an active participant in co-constructing meaning during interviews (Holloway & Galvin, 2017), two individuals analysed the data in order to minimise bias. A reflexive approach was taken throughout (Berger, 2015).

Implications for policy and practice

To promote collaboration between practitioners and carers, the authors support some recommendations proposed by Lloyd and King (2003); managers must openly ask staff about their involvement with carers during routine review meetings; collaboration should become a key feature of performance appraisals; and staff should receive appropriate training on the topic. The authors recommend that the study findings are used to guide training curricula. Unprofessional behaviours were highlighted during the study. It has previously been demonstrated that approximately a quarter of students found the school environment "not very conducive" or "not at all conducive" to the open discussion of ethical concerns (Satterwhite et al., 2000). The authors would like to reiterate the recommendations of Monrouxe et al. (2011); that more sense-making opportunities should be made available to students, under the supervision of clinical educators. The findings may be used to inform the development of assessments for the selection of staff into MHS. In addition, it may be fruitful to integrate these findings into the standards and inspection guides of health care regulators.

Implications for Future Research

The study did not set out, specifically, to identify differences amongst the professions, with regards to the desired professional attributes. Further research could utilise these findings to determine the extent that each profession endorses each attribute. Professionals face many challenges at work and rely on patients to engage with their treatment in order for the practitioner themselves, to fulfil their role. The authors suggest that future research explores the professionalism of patients that use MHS. A greater understanding of the impact that patient behaviour has on the clinical care delivered will facilitate the development of more tailored packages for patients, and more educational interventions

for staff. The study demonstrates that professionals need to fulfil expectations across 10 professional attributes in MHS; including working with carers. The authors urge practitioners to support carers and provide a forum for them to air their views. When a patient is in crisis, a carer's words could be lifesaving.

ORCID

Lauren M. E. Aylott http://orcid.org/0000-0002-2261-9360
Paul A. Tiffin http://orcid.org/0000-0003-1770-5034
Sally Brown http://orcid.org/0000-0001-5865-8858
Gabrielle M. Finn http://orcid.org/0000-0002-0419-694X

References

Aguilar, A., Stupans, L., & Scutter, S. (2011). Assessing students' professionalism: Considering professionalism's diverging definitions. *Education for Health*, 24(3), 599. http://www.educationforhealth.net/text.asp?2011/24/3/599/101417

Aylott, L. M., Tiffin, P. A., Saad, M., Llewellyn, A. R., & Finn, G. M. (2019). Defining professionalism for mental health services: A rapid systematic review. *Journal of Mental Health*, 28(5), 546–565. https://doi.org/10.1080/09638237.2018.1521933

Berger, R. (2015). Now I see it, now I don't: Researcher's position and reflexivity in qualitative research. *Qualitative Research*, 15(2), 219–234. https://doi.org/10.1177/1468794112468475

Bhugra, D. (2008). Renewing psychiatry's contract with society. *Psychiatric Bulletin*, 32(8), 281–283. https://doi.org/10.1192/pb.bp.108.020560

Bhugra, D., & Gupta, S. (2010). Medical professionalism in psychiatry. *Advances in Psychiatric Treatment*, 16(1), 10–13. https://doi.org/10.1192/apt.bp.108.005892

British Medical Association (2017). *Breaking down barriers: The challenge of improving mental health outcomes*. British Medical Association. https://www.bma.org.uk/-/media/files/pdfs/collective%20voice/policy%20research/public%20and%20population%20health/mental%20health/breaking-down-barriers-mental-health-briefing-apr2017.pdf?la=en.

Burford, B., Morrow, G., Rothwell, C., Carter, M., & Illing, J. (2014). Professionalism education should reflect reality: Findings from three health professions. *Medical Education*, 48(4), 361–374. https://doi.org/10.1111/medu.12368

Cleary, M., Freeman, A., & Walter, G. (2006). Carer participation in mental health service delivery. *International Journal of Mental Health Nursing*, 15(3), 189–194. https://doi.org/10.1111/j.1447-0349.2006.00422.x

Cleary, M., Freeman, A., Hunt, G. E., & Walter, G. (2005). What patients and carers want to know: An exploration of information and resource needs in adult mental health services. *The Australian*

and New Zealand Journal of Psychiatry, 39(6), 507–513. https://doi.org/10.1080/j.1440-1614.2005.01611.x

Cruess, S. R. (2006). Professionalism and medicine's social contract with society. *Clinical Orthopaedics and Related Research, 449*, 170–176. https://doi.org/10.1097/01.blo.0000229275.66570.97

de Leng, W. E., Stegers-Jager, K. M., Born, M. P., & Themmen, A. P. (2018). Integrity situational judgement test for medical school selection: Judging 'what to do' versus 'what not to do'. *Medical Education, 52*(4), 427–437. https://doi.org/10.1111/medu.13498

Department of Health. (2000). *No secrets: Guidance on developing and implementing multi-agency policies and procedures to protect vulnerable adults from abuse*. Department of Health. (updated in 2015). Retrieved from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/194272/No_secrets__guidance_on_developing_and_implementing_multi-agency_policies_and_procedures_to_protect_vulnerable_adults_from_abuse.pdf

Finn, G., Garner, J., & Sawdon, M. (2010). 'You're judged all the time!' Students' views on professionalism: A multicentre study. *Medical Education, 44*(8), 814–825. https://doi.org/10.1111/j.1365-2923.2010.03743.x

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 51*(4), 327–358. Retrieved from https://www.apa.org/pubs/databases/psycinfo/cit-article.pdf

Freeman, T. (2006). 'Best practice' in focus group research: Making sense of different views. *Journal of Advanced Nursing, 56*(5), 491–497. https://doi.org/10.1111/j.1365-2648.2006.04043.x

Green, M., Zick, A., & Makoul, G. (2009). Defining professionalism from the perspective of patients, physicians, and nurses. *Academic Medicine: Journal of the Association of American Medical Colleges, 84*(5), 566–573. https://doi.org/10.1097/ACM.0b013e31819fb7ad

Hafferty, F. W., & Franks, R. (1994). The hidden curriculum, ethics teaching, and the structure of medical education. *Academic Medicine: Journal of the Association of American Medical Colleges, 69*(11), 861–871. https://doi.org/10.1097/00001888-199411000-00001

Holloway, I., & Galvin, K. (2017). *Qualitative research in nursing and healthcare* (4th ed.). Wiley Blackwell.

Irvine, D. (1997). The performance of doctors. I: Professionalism and self regulation in a changing world. *BMJ, 314*(7093), 1540–1540. https://doi.org/10.1136/bmj.314.7093.1540

Jones, J. W., McCullough, L. B., & Richman, B. W. (2006). Ethics and professionalism: Do we need yet another surgeons' charter? *Journal of Vascular Surgery, 44*(4), 903–906. https://doi.org/10.1016/j.jvs.2006.07.028

Lambert, S., & Loiselle, C. (2008). Combining individual interviews and focus groups to enhance data richness. *Journal of Advanced Nursing, 62*(2), 228–237. https://doi.org/10.1111/j.1365-2648.2007.04559.x

Lloyd, C., & King, R. (2003). Consumer and carer participation in mental health services. *Australasian Psychiatry, 11*(2), 180–184. https://doi.org/10.1046/j.1039-8562.2003.00537.x

Medical Schools Council. (2011). *Improving selection to the foundation programme: Final Report*. https://isfp.orguk.files.wordpress.com/2017/04/isfp-final-report-appendices.pdf

Monrouxe, L. V., & Rees, C. E. (2017). *Healthcare professionalism: Improving practice through reflections on workplace dilemmas*. John Wiley & Sons.

Monrouxe, L. V., Rees, C. E., & Hu, W. (2011). Differences in medical students' explicit discourses of professionalism: Acting, representing, becoming. *Medical Education, 45*(6), 585–602. https://doi.org/10.1111/j.1365-2923.2010.03878.x

Papadakis, M. A., Hodgson, C. S., Teherani, A., & Kohatsu, N. D. (2004). Unprofessional behavior in medical school is associated with subsequent disciplinary action by a state medical board. *Academic Medicine: Journal of the Association of American Medical Colleges, 79*(3), 244–249. https://doi.org/10.1097/00001888-200403000-00011

Project of the ABIM Foundation, A. A. F., and European Federation of Internal Medicine. (2002). Medical professionalism in the new millennium: A physician charter. *Annals of Internal Medicine, 136*(3), 243–246. https://doi.org/10.7326/0003-4819-136-3-200202050-00012

Rees, C. E., & Knight, L. V. (2007). The trouble with assessing students' professionalism: Theoretical insights from sociocognitive psychology. *Academic Medicine: Journal of the Association of American Medical Colleges, 82*(1), 46–50. https://doi.org/10.1097/01.ACM.0000249931.85609.05

Ritchie, J., & Spencer, L. (1994). Qualitative data analysis for applied policy research. In A. Bryman & R. Burgess (Eds.), *Analizing Qualitative Data* (pp. 173–194). Routledge.

Ritchie, J., Spencer, L., & O'Connor, W. (2003). Carrying out qualitative analysis. In J. Ritchie & J. Lewis (Eds.), *Qualitative research practice: A guide for social science students and researchers*. (pp. 219–262). SAGE Publications Ltd.

Satterwhite, R. C., Satterwhite, W. M., & Enarson, C. (2000). An ethical paradox: The effect of unethical conduct on medical students' values. *Journal of Medical Ethics, 26*(6), 462–465. https://doi.org/10.1136/jme.26.6.462

van de Camp, K., Vernooij-Dassen, M., Grol, R., & Bottema, B. (2004). How to conceptualize professionalism: A qualitative study. *Medical Teacher, 26*(8), 696–702. https://doi.org/10.1080/01421590400019518

Woodruff, J., Angelos, P., & Valaitis, S. (2008). Medical professionalism: One size fits all? *Perspectives in Biology and Medicine, 51*(4), 525–534. https://doi.org/10.1353/pbm.0.0056

Worthington, A., Rooney, P., & Hannan, R. (2013). *The Triangle of Care. Carers Included: A Guide to Best Practice in Mental Health in England* (2nd ed.). Carers Trust. https://professionals.carers.org/sites/default/files/thetriangleofcare_guidetobestpracticeinmentalhealthcare_england.pdf

Appendix B: Results of literature search by database

(CINAHL Plus via EBSCO) <searched on 02/03/2017>

1.   (Professionali?m OR professionali?ation OR unprofessional*).m_titl. (2223)

2.   ("professional competenc*" or "professional skill*" or "professional value*" or "professional role*" or "professional attitude*" or "professional identit*" or "professional practice*" or "professional communication*" or "professional standard*" or "professional accountab*" or "professional dissonanc*" or "professional impair*" or "professional dysfunction*" or "professional malpractice*" or "professional misconduct*" or "professional omission*").m_titl. (3120)

3.   ((Professiona* ADJ3 (issue* OR behav* OR act* OR ethic* OR humanis*)) NOT (Professional ADJ3 activ*)).m_titl. (1984)

4.   *professionalism/ (2305)

5.   1 OR 2 OR 3 OR 4 (8287)

6.   ("mental health" or psychiatr* or "learning disabilit*" or "learning difficult*" or "learning disorder*" or "intellectual disabilit*").ti,ab. (113393)

7.   (AMHP* or counsell* or RMN* or psychotherap* or therap*).ti,ab. (352339)

8.   mental health/ (20385)

9.   psychiatry/ (7085)

10.   learning disorders/ (5999)

11.   intellectual disability/ (15732)

12.   6 OR 7 OR 8 OR 9 OR 10 OR 11 (477114)

13.   ("physical therap*" OR "occupational therap*").ti,ab. (27451)

14.   12 NOT 13 (449663)

15.   5 AND 14 (389)

16.   limit 15 to (english language and yr="2006 -Current") (219)

(Ovid MEDLINE(R) 1946) <searched on 02/03/2017>

1.   (Professionali?m OR professionali?ation OR unprofessional*).m_titl. (3024)

2.   ("professional competenc*" or "professional skill*" or "professional value*" or "professional role*" or "professional attitude*" or "professional identit*" or "professional practice*" or "professional communication*" or "professional standard*" or "professional accountab*" or "professional dissonanc*" or "professional impair*" or "professional dysfunction*" or "professional malpractice*" or "professional misconduct*" or "professional omission*").m_titl. (2724)

3.   ((Professiona* ADJ3 (issue* OR behav* OR act* OR ethic* OR humanis*)) NOT (Professional ADJ3 activ*)).m_titl. (1648)

4.   *professionalism/ (199)

5.   1 OR 2 OR 3 OR 4 (7173)

6.   ("mental health" or psychiatr* or "learning disabilit*" or "learning difficult*" or "learning disorder*" or "intellectual disabilit*").ti,ab. (268867)

7.   (AMHP* or counsell* or RMN* or psychotherap* or therap*).ti,ab. (2049188)

8.   mental health/ (28234)

9.   psychiatry/ (36705)

10.   learning disorders/ (13381)

11.   intellectual disability/ (50608)

12.   6 OR 7 OR 8 OR 9 OR 10 OR 11 (2350037)

13.   ("physical therap*" OR "occupational therap*").ti,ab. (23033)

14.   12 NOT 13 (2327004)

15.   5 AND 14 (401)

16.   limit 15 to (english language and yr="2006 -Current") (161)

Embase 1974 to 2017 <searched on 02/03/2017>

1.   (Professionali?m OR professionali?ation OR unprofessional*).m_titl. (3678)

2.   ("professional competenc*" or "professional skill*" or "professional value*" or "professional role*" or "professional attitude*" or "professional identit*" or "professional practice*" or "professional communication*" or "professional standard*" or "professional accountab*" or "professional dissonanc*" or "professional impair*" or "professional dysfunction*" or "professional malpractice*" or "professional misconduct*" or "professional omission*").m_titl. (3321)

3.   ((Professiona* ADJ3 (issue* OR behav* OR act* OR ethic* OR humanis*)) NOT (Professional ADJ3 activ*)).m_titl. (1905)

4.   *professionalism/ (2374)

5.   1 OR 2 OR 3 OR 4 (9437)

6.   ("mental health" or psychiatr* or "learning disabilit*" or "learning difficult*" or "learning disorder*" or "intellectual disabilit*").ti,ab. (406099)

7.   (AMHP* or counsell* or RMN* or psychotherap* or therap*).ti,ab. (3190765)

8.   mental health/ (131039)

9.   psychiatry/ (74336)

10.   learning disorders/ (12756)

11.   intellectual disability/ (1909)

12.   6 OR 7 OR 8 OR 9 OR 10 OR 11 (3605148)

13.   ("physical therap*" OR "occupational therap*").ti,ab. (39952)

14.   12 NOT 13 (3565196)

15.   5 AND 14 (614)

16.   limit 15 to (english language and yr="2006 -Current") (303)

PsycINFO 1806 <searched on 02/03/2017>

1. (Professionali?m OR professionali?ation OR unprofessional*).m_titl. (1291)

2. ("professional competenc*" or "professional skill*" or "professional value*" or "professional role*" or "professional attitude*" or "professional identit*" or "professional practice*" or "professional communication*" or "professional standard*" or "professional accountab*" or "professional dissonanc*" or "professional impair*" or "professional dysfunction*" or "professional malpractice*" or "professional misconduct*" or "professional omission*").m_titl. (2067)

3. ((Professiona* ADJ3 (issue* OR behav* OR act* OR ethic* OR humanis*)) NOT (Professional ADJ3 activ*)).m_titl. (1340)

4. *professionalism/ (2391)

5. 1 OR 2 OR 3 OR 4 (5919)

6. ("mental health" or psychiatr* or "learning disabilit*" or "learning difficult*" or "learning disorder*" or "intellectual disabilit*").ti,ab. (360463)

7. (AMHP* or counsell* or RMN* or psychotherap* or therap*).ti,ab. (386681)

8. mental health/ (51531)

9. psychiatry/ (23980)

10. learning disorders/ (2275)

11. intellectual disability/ (0)

12. 6 OR 7 OR 8 OR 9 OR 10 OR 11 (692834)

13. ("physical therap*" OR "occupational therap*").ti,ab. (11215)

14. 12 NOT 13 (681619)

15. 5 AND 14 (976)

16. limit 15 to (english language and yr="2006 -Current") (474)

HMIC Health Management Information Consortium 1979 <searched on 02/03/2017>

1.   (Professionali?m OR professionali?ation OR unprofessional*).m_titl. (184)

2.   ("professional competenc*" or "professional skill*" or "professional value*" or "professional role*" or "professional attitude*" or "professional identit*" or "professional practice*" or "professional communication*" or "professional standard*" or "professional accountab*" or "professional dissonanc*" or "professional impair*" or "professional dysfunction*" or "professional malpractice*" or "professional misconduct*" or "professional omission*").m_titl. (305)

3.   ((Professiona* ADJ3 (issue* OR behav* OR act* OR ethic* OR humanis*)) NOT (Professional ADJ3 activ*)).m_titl. (127)

4.   *professionalism/ (0)

5.   1 OR 2 OR 3 OR 4 (602)

6.   ("mental health" or psychiatr* or "learning disabilit*" or "learning difficult*" or "learning disorder*" or "intellectual disabilit*").ti,ab. (26793)

7.   (AMHP* or counsell* or RMN* or psychotherap* or therap*).ti,ab. (14278)

8.   mental health/ (5938)

9.   psychiatry/ (690)

10.   learning disorders/ (203)

11.   intellectual disability/ (0)

12.   6 OR 7 OR 8 OR 9 OR 10 OR 11 (40691)

13.   ("physical therap*" OR "occupational therap*").ti,ab. (1650)

14.   12 NOT 13 (39041)

15.   5 AND 14 (58)

16.   limit 15 to (english language and yr="2006 -Current") (19)

(CINAHL Plus via EBSCO) <searched on 27/07/2022>

1. TI Professionali?m OR professionali?ation OR unprofessional* (3027)

2. TI "professional competenc*" OR "professional skill*" OR "professional value*" OR "professional role*" OR "professional attitude*" OR "professional identit*" OR "professional practice*" OR "professional communication*" OR "professional standard*" OR "professional accountab*" OR "professional dissonanc*" OR "professional impair*" OR "professional dysfunction*" OR "professional malpractice*" OR "professional misconduct*" OR "professional omission*" (4354)

3. TI (Professiona* N3 (issue* OR behav* OR act* OR ethic* OR humanis*)) NOT (Professional N3 activ*) (2852)

4. MM *professionalism (3497)

5. S1 OR S2 OR S3 OR S4 (11742)

6. TI ("mental health" OR psychiatr* OR "learning disabilit*" OR "learning difficult*" OR "learning disorder*" OR "intellectual disabilit*") OR AB ("mental health" OR psychiatr* OR "learning disabilit*" OR "learning difficult*" OR "learning disorder*" OR "intellectual disabilit*") (223441)

7. TI (AMHP* OR counsell* OR RMN* OR psychotherap* OR therap*) OR AB (AMHP* OR counsell* OR RMN* OR psychotherap* OR therap*) (717013)

8. MM mental health (28069)

9. MM psychiatry (7754)

10. MM learning disorders (4990)

11. MM intellectual disability (18499)

12. S6 OR S7 OR S8 OR S9 OR S10 OR S11 (932224)

13. TI ("physical therap*" OR "occupational therap*") AND AB ("physical therap*" OR "occupational therap*") (49520)

14. S12 NOT S13 (882704)

15. S5 AND S14 (693)

16. S15 (Limiters = Published date: 20170301-; English Language; 204)

Appendix C: PRISMA checklist for the systematic review

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	84
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	-
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	85
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	85
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	87
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	87
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	90

Section/topic	#	Checklist item	Reported on page #
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	91
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	90
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	92
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	92
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	93
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	-
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2) for each meta-analysis.	93
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	-
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	-
RESULTS)			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	94
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	96

Section/topic	#	Checklist item	Reported on page #
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	-
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	106
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	106
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	-
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	116

DISCUSSION

Section/topic	#	Checklist item	Reported on page #
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	118
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	121
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	122

FUNDING

Section/topic	#	Checklist item	Reported on page #
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	21

Appendix D: University of York ethics approval / guidance – qualitative study

UNIVERSITY
of York

DEPARTMENT OF
HEALTH SCIENCES

c/o Department of Philosophy
Heslington
York YO10 5DD

Chair, Health Sciences Research
Governance Committee
www.york.ac.uk/healthsciences

9 February 2018

Miss L Aylott
Hull York Medical School
University of York
York
YO10 5DD

Dear Lauren

Understanding and exploring professionalism in a mental health service context: A qualitative study using stakeholder focus

Thank you for submitting the above project to the Health Sciences Research Governance Committee for approval. Your application was considered by the committee at its meeting on 5 February 2018.

I am pleased to report that the committee approved the submission, which can now go forward to the NHS REC. I was asked to feedback the following points.

IRAS Form

Q2 – is it correct to tick the mixed-method rather than qualitative methods category?

A13 – here (and elsewhere) there is mention of a follow-up study involving re-contacting participants who consented to this second stage (i.e., the scoring key), but the committee couldn't see where or how participants are informed of and consent to this follow-up.

A28 – posters and flyers are mentioned: a copy of the flyer is submitted but not a poster; if they are one and the same, just refer to flyers to avoid confusion.

A60 – a maximum of 65 participants is said to be recruited, but should this be 75 (60 max for the focus groups and 12-15 for the panel)?

A62 – a more detail account of how thematic analysis will be carried out, including a reference to an appropriate article or book, could be provided.

Flyer

Although professionals will probably understand terms such as, 'Situational Judgement Test' and 'Values Based Recruitment', patients/carers might not, so these terms could be made clear.

It is unclear whether the phone number on the Flyer is a personal mobile, but if so this must be changed to a work number.

1

Information Sheet

The sheet state:

> 'At the end of the project we would like to share data with others to facilitate future research. If you agree for us to do this we will send anonymised data to a repository so that it can be kept for a minimum of 10 years, in line with the University of York RDF policy. Data will be anonymised, and we will ensure it doesn't identify you.'

The committee felt that this was a little unclear; also, it is not represented on the consent form, so it's not clear how participants will agree to it.

'University of York Research Governance Committee' should read, 'Health Sciences Research Governance Committee at the University of York'

Consent form

Cross-reference to the information sheet requires the version and date of the information sheet.

The consent form should reiterate that participants can withdraw from the study, and clarify what will happen to data already collected.

The participant's name and participant ID number appear together on the same form but usually these are kept separate so as to preserve anonymity.

Participants are asked to circle the appropriate responses, but signatures (or at least initials) are better to ensure the participants have genuinely consented to each item.

The committee is happy for you to take up this feedback with your supervisors as you redraft the IRAS form, but if you have any queries regarding the decision or feedback, or make any substantial amendments to the study, please contact me. Finally if you intend to submit this letter or any other correspondence from the HSRGC as part of your assessed work (e.g., to demonstrate that your study has ethical approval) please make sure you edit the letter so as to maintain anonymity.

Yours sincerely

Chair: HSRGC

Cc: Dr Gabrielle Finn and Dr Paul Tiffin

Appendix E: Research ethics committee approval – qualitative study

NHS
Health Research
Authority

London - Camden & Kings Cross Research Ethics Committee

NHSBT Newcastle Blood Donor Centre
Holland Drive
Newcastle upon Tyne
NE2 4NQ

Telephone: 0207 104 8086

> Please note: This is the
> favourable opinion of the
> REC only and does not allow
> you to start your study at NHS
> sites in England until you
> receive HRA Approval

03 July 2018

Miss Lauren Aylott
Hull and York Medical School
University of York
Heslington
YO10 5DD

Dear Miss Aylott

Study title:	Understanding and Exploring Professionalism in a Mental Health Service Context: A Qualitative Study using Stakeholder Focus
REC reference:	18/LO/0630
IRAS project ID:	234296

Thank you for your letter of 07 June 2018, responding to the Committee's request for further information on the above research and submitting revised documentation.

The further information has been considered on behalf of the Committee by the Chair.

We plan to publish your research summary wording for the above study on the HRA website, together with your contact details. Publication will be no earlier than three months from the date of this opinion letter. Should you wish to provide a substitute contact point, require further information, or wish to make a request to postpone publication, please contact hra.studyregistration@nhs.net outlining the reasons for your request.

Confirmation of ethical opinion

On behalf of the Committee, I am pleased to confirm a favourable ethical opinion for the above research on the basis described in the application form, protocol and supporting documentation as revised, subject to the conditions specified below.

Conditions of the favourable opinion

The REC favourable opinion is subject to the following conditions being met prior to the start of the study.

Management permission must be obtained from each host organisation prior to the start of the

<u>study at the site concerned.</u>

Management permission should be sought from all NHS organisations involved in the study in accordance with NHS research governance arrangements. Each NHS organisation must confirm through the signing of agreements and/or other documents that it has given permission for the research to proceed (except where explicitly specified otherwise).

Guidance on applying for HRA and HCRW Approval (England and Wales)/ NHS permission for research is available in the Integrated Research Application System, at www.hra.nhs.uk or at http://www.rdforum.nhs.uk.

Where a NHS organisation's role in the study is limited to identifying and referring potential participants to research sites ("participant identification centre"), guidance should be sought from the R&D office on the information it requires to give permission for this activity.

For non-NHS sites, site management permission should be obtained in accordance with the procedures of the relevant host organisation.

Sponsors are not required to notify the Committee of management permissions from host organisations

<u>Registration of Clinical Trials</u>

All clinical trials (defined as the first four categories on the IRAS filter page) must be registered on a publically accessible database within 6 weeks of recruitment of the first participant (for medical device studies, within the timeline determined by the current registration and publication trees).

There is no requirement to separately notify the REC but you should do so at the earliest opportunity e.g. when submitting an amendment. We will audit the registration details as part of the annual progress reporting process.

To ensure transparency in research, we strongly recommend that all research is registered but for non-clinical trials this is not currently mandatory.

If a sponsor wishes to request a deferral for study registration within the required timeframe, they should contact hra.studyregistration@nhs.net. The expectation is that all clinical trials will be registered, however, in exceptional circumstances non registration may be permissible with prior agreement from the HRA. Guidance on where to register is provided on the HRA website.

It is the responsibility of the sponsor to ensure that all the conditions are complied with before the start of the study or its initiation at a particular site (as applicable).

Ethical review of research sites

NHS sites

The favourable opinion applies to all NHS sites taking part in the study, subject to management permission being obtained from the NHS/HSC R&D office prior to the start of the study (see "Conditions of the favourable opinion" below).

Non-NHS sites

5 Approved documents

The final list of documents reviewed and approved by the Committee is as follows:

Document	Version	Date
Copies of advertisement materials for research participants [Flyer - Focus Groups]	2	14 May 2018
Copies of advertisement materials for research participants [Flyer - Interviews]	2	14 May 2018
Covering letter on headed paper	version 1	19 March 2018
Evidence of Sponsor insurance or indemnity (non NHS Sponsors only)		
IRAS Application Form [IRAS_Form_27032018]		27 March 2018
Letter from funder [PhD Scholarship Confirmation]		06 September 2016
Letter from sponsor		
Non-validated questionnaire [Demographic Questionnaire - Focus Groups]	2	14 May 2018
Non-validated questionnaire [Demographic Questionnaire - Interviews]	2	13 May 2018
Participant consent form [Consent form for staff]	version 1	19 March 2018
Participant consent form [Consent form for carers]	version 1	19 March 2018
Participant consent form [Consent form for service users]	version 1	19 March 2018
Participant consent form [Service Users]	2	14 May 2018
Participant information sheet (PIS) [Allied Health Professionals]	2	14 May 2018
Participant information sheet (PIS) [Carers]	2	14 May 2018
Participant information sheet (PIS) [Nurses]	2	14 May 2018
Participant information sheet (PIS) [Psychiatrists]	2	14 May 2018
Participant information sheet (PIS) [Psychologists]	2	14 May 2018
Participant information sheet (PIS) [Service Users]	2	14 May 2018
Referee's report or other scientific critique report [Summary Review of wider NIHR project]		
Research protocol or project proposal	2	14 May 2018
Response to Request for Further Information [Response to REC]		07 June 2018
Summary CV for Chief Investigator (CI)	version 1	19 March 2018
Summary CV for supervisor (student research) [Gabrielle's CV]		
Summary CV for supervisor (student research) [Paul's CV]		

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

After ethical review

Reporting requirements

The attached document *"After ethical review – guidance for researchers"* gives detailed guidance on reporting requirements for studies with a favourable opinion, including:

- Notifying substantial amendments
- Adding new sites and investigators

- Notification of serious breaches of the protocol
- Progress and safety reports
- Notifying the end of the study

The HRA website also provides guidance on these topics, which is updated in the light of changes in reporting requirements or procedures.

User Feedback

The Health Research Authority is continually striving to provide a high quality service to all applicants and sponsors. You are invited to give your view of the service you have received and the application procedure. If you wish to make your views known please use the feedback form available on the HRA website:
http://www.hra.nhs.uk/about-the-hra/governance/quality-assurance/

HRA Training

We are pleased to welcome researchers and R&D staff at our training days – see details at
http://www.hra.nhs.uk/hra-training/

18/LO/0630	Please quote this number on all correspondence

With the Committee's best wishes for the success of this project.

Yours sincerely

pp

Mrs Rosie Glazebrook
Chair

Email:nrescommittee.london-camdenandkingscross@nhs.net

Enclosures:	"After ethical review – guidance for researchers"
Copy to:	*Dr Michael Barber* *Mrs Sarah Daniel, Tees, Esk and Wear Valleys NHS Foundation Trust*

Appendix F: Example participant information sheet – qualitative study

Professionalism in a Mental Health Context: A Qualitative Study, version 3.0

NHS
Tees, Esk and Wear Valleys
NHS Foundation Trust

H·Y·M·S
THE HULL YORK
MEDICAL SCHOOL

Participant Information Sheet – Psychologists

Study Title:

Understanding and exploring professionalism in a mental health service context: a qualitative study using stakeholder focus

Invitation to Participate

We would like to invite you to take part in a research project. It is important that you understand the purpose of this study and what it will involve before you decide whether you would like to participate. Please read this information sheet carefully and discuss with others if you wish. If you would like more information, or there is something you do not understand, please contact the research team. There will be no negative consequences should you choose to not participate.

Research team contact: Lauren Aylott, PhD Student, Hull York Medical School.

Email: lauren.aylott@nhs.net; Telephone: 07453 361566

What is the purpose of the study?

We would like to find out what professionalism means to individuals using or working within mental health services and what values, attitudes, skills and behaviours are desired of professionals working in a mental health service context.

A recent review of the literature that sought to define professionalism in mental health services highlighted the importance of the relationship between professionals and patients using this service. Service users, carers and staff desire certain attributes of professionals working in mental health services and we would like to further explore this area.

Once we have identified the behaviours that people desire of professionals working in mental health services we would like to develop a test that will assess healthcare workers' knowledge on this topic. Consequently, the test will support values based recruitment; promoting a workforce with the right values as well as the appropriate knowledge and skills for the role.

Version 3 21.08.18   IRAS ID: 234296

Why have I been invited to take part?

We are keen to hear the views of psychologists working in mental health or learning disability services. You are being asked to take part because you fulfil this role.

What will happen if I take part?

If you would like to participate we will ask you to sign a consent form.

You will be invited to a focus group with other psychologists that will last for approximately 60 to 90 minutes, conducted by the research team. There will be no need to divulge any clinical information about your own health, or that of any patients. We are simply interested in professional behaviours.

You will be asked about the values, attitudes, and behaviours that you would like to see in professionals working in mental health or learning disability services. We want to hear your views about professionalism, but request that all information provided is done so with anonymity. As noted in Tees, Esk and Wear Valleys 'Confidentiality and sharing information policy' (Ref CORP-0010-v9), both patient and staff information must be handled with regard to confidentiality (see p.4).

The interview will be audio recorded so we can remember what has been said.

The focus groups will be at a NHS trust property in a location close to you.

We will archive the transcribed data following the project, making it available to others for future research. This will be anonymised however and would not identify you. We will also retain consent forms for a period of 10 years to evidence your consent to participate. Over the course of the study, a file that links both your consent form and the data will be stored separately. This will only be accessible to the research team and will be destroyed as soon as the project ends (November 2021). Following the project, there will be no linkage between your consent form and the data and your data will remain anonymous.

Will you compensate me for my time?

Refreshments will be available at all focus groups.

Are there any advantages in taking part?

We hope that you will enjoy sharing your views and having the opportunity to highlight what's most important to you.

Findings from the focus groups will inform the development of a test to measure practitioners' knowledge of professionalism. This test will then be used during the

recruitment of workers into mental health and learning disability services. The aim is to better select staff with the highest levels of professionalism.

Your comments and the overall findings of this study may be considered in the teaching of professionalism and potentially influence practice.

Are there any disadvantages or risks in taking part?

We hope that you will find the opportunity to share your views empowering. In the unlikely event that you find the discussion distressing we will have clinical staff on hand to speak to. You are free to withdraw at any time.

We are grateful of the time people provide for the focus group.

What if I am unhappy or if there is a problem?

If you have any concerns or complaints about the study or about the way you've been approached and treated during it, you can contact the Research and Enterprise Office of the University of York: Dr Michael Barber, Innovation Centre, York Science Park, York YO10 5DG, UK; Tel: +44 (0)1904 435154.

The normal National Health Service complaints mechanisms will also be available to you.

What will happen to my answers?

Focus groups will be recorded. Audio recordings will be transcribed (typed) verbatim by a professional service and recordings will be destroyed following transcription. Transcripts will be stored on encrypted computer.

We will take precautions to make sure the information you share is kept safe. Audio recordings will be made on an encrypted Dictaphone and the data will be transferred to a laptop that is both encrypted and password protected. A backup of the data will be transferred via a virtual private network to a secure folder managed by the University of York.

Your name and details will be removed from any paper copies of data, which will instead contain an anonymous participant number. Hard copies of data will be scanned and stored on an encrypted and password protected computer. Hard copies will then be shredded.

Information collected during focus groups will be anonymised so that you cannot be identified in any way. To support analysis, your responses may be discussed within the research team. There are special circumstances when we may need to share information, including if there are concerns about yours or someone else's safety.

We would share information to keep you or others safe. If we believe this is necessary at any time we will inform you and explain why.

We will group your information with others to get an overall insight as to what's most important to patients, carers and staff members. We would like to write reports to share our findings and these may be published. In reports we may use direct quotes from the focus groups, but we will use pseudonyms (a name that is not your own) and you will never be named directly in any of our findings. We will make sure that others cannot identify you in any reports or publications. If you would like to receive a summary of the research findings, we will send you one once the study is completed.

Has this project received ethical approval?

This project has been endorsed by the Health Sciences Research Governance Committee at the University of York. Ethics approval has been received from the London – Camden and Kings Cross research ethics committee.

Do I have to take part?

No. It's your choice. If you choose to participate, you may withdraw at any time. You do not have to give a reason, and this will not disadvantage you in any way.

What happens next?

If you are interested in taking part or would like more information please contact:

Lauren Aylott, PhD Student, Hull York Medical School

Email: lauren.aylott@nhs.net

Telephone: 07453 361566

Thank you for taking the time to read this information sheet.

Appendix G: Consent form – qualitative study

Professionalism in a Mental Health Context: A Qualitative Study, version 2.0

H·Y·M·S
THE HULL YORK
MEDICAL SCHOOL

NHS
Tees, Esk and Wear Valleys
NHS Foundation Trust

Consent Form – Service Users

Title of Study:
Understanding and exploring professionalism in a mental health service context: a qualitative study using stakeholder focus

Principal Researcher   **Supervisors**
Miss Lauren Aylott   Dr Gabrielle Finn and Dr Paul Tiffin

If you wish to participate in this study please initial the appropriate responses and sign and date the declaration at the end of this document.

	Please initial box if in agreement	
	YES	NO
• I have read the participant information sheet (Version 3, 21.08.18) and had the opportunity to ask questions. I understand the answers provided and know what taking part in the study will involve.		
• I understand that participation is voluntary, and I am free to withdraw at any time. I do not have to give a reason, and this will not affect my current or future care.		
• I understand that if I withdraw from the study, data that has already been collected will be maintained and used for the purposes of this study.		
• I understand that the information I supply (unless anonymised) will be confidential and will not be shared with others **unless** circumstances arise that breach safeguarding practices (e.g. if myself or others are at risk of harm).		
• I understand that the information I provide may be published and this may include direct quotations. Pseudonyms will be used (a name that is not my own) to protect my identity.		
• I understand that the interview will be recorded using encrypted audio equipment. Audio recordings will be transferred and stored securely on an encrypted and password protected computer within 24 hours.		
• I agree to my anonymised data being stored for 10 years following the project so that it is available to others for future research.		
• I understand that my consent form will be retained for 10 years but any data linking my consent form to my data will be viewed by the research team only and will be destroyed as soon as the project ends (November 2021).		
• I understand that the Trust's records will be searched to confirm that I am current or recent patient and that my name, full date of birth and address will be required to access this information. I understand that my date of birth and address will be securely deleted from any research records as soon as my patient status has been confirmed.		
• I agree to participate in this research project.		

....

Name of Participant   Date   Signature

....

Name of Person taking consent   Date   Signature

Version 2 14.05.18   IRAS ID: 234296
When completed, 1 copy for participant; 1 for researcher

Appendix H: HYMS ethics committee approval – pilot study

Hull York Medical School

Hull
University of Hull
Hull, HU6 7RX, UK

York
University of York
York, YO10 5DD, UK

T 0870 1245500
info@hyms.ac.uk
www.hyms.ac.uk

11 November 2019

Lauren Aylott
PhD Student
Health Professions Education Unit
Hull York Medical School

Dear Lauren

19 47 – 'What would you do?' Evaluating a situational judgement test for use in mental health services.

I have reviewed this study on behalf of HYMS Ethical Committee with respect to the documents received on 25 October 2019 (Approval Letter_LA 11.11.19, Consent_Form_IACS_2019.10.10_FINAL, Email to IACS attendees_2019.10.10-FINAL, Email to Initial participants-2019.10.10-FINAL, Email to supervisors_peers-2019.10.10-FINAL, Flyer.2019.10.10-FINAL, IRAS-2019.10.10-FINAL, LA_Protocol_v1_2019.10.10-FINAL).

I am pleased to inform you that I do not have any HYMS-specific or ethical concerns, or additional requirements.

On behalf of the Ethics Committee, we wish you success with this study.

Kind regards

Yours sincerely

Professor Thozhukat Sathyapalan
Chair
HYMS Ethics Committee

UNIVERSITY OF HULL   UNIVERSITY of York   Athena SWAN Silver Award

Appendix I: Health Research Authority approval – pilot study

Ymchwil Iechyd
a Gofal Cymru
Health and Care
Research Wales

NHS
Health Research Authority

Miss Lauren Aylott
Hull and York Medical School
University of York
Heslington
YO10 5DD

Email: hra.approval@nhs.net
HCRW.approvals@wales.nhs.uk

25 November 2019

Dear Miss Aylott

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	'What would you do?' Evaluating a situational judgement test for use in mental health services
IRAS project ID:	270527
REC reference:	19/HRA/6403
Sponsor	University of York

I am pleased to confirm that HRA and Health and Care Research Wales (HCRW) Approval has been given for the above referenced study, on the basis described in the application form, protocol, supporting documentation and any clarifications received. You should not expect to receive anything further relating to this application.

Please now work with participating NHS organisations to confirm capacity and capability, in line with the instructions provided in the "Information to support study set up" section towards the end of this letter.

How should I work with participating NHS/HSC organisations in Northern Ireland and Scotland?
HRA and HCRW Approval does not apply to NHS/HSC organisations within Northern Ireland and Scotland.

If you indicated in your IRAS form that you do have participating organisations in either of these devolved administrations, the final document set and the study wide governance report (including this letter) have been sent to the coordinating centre of each participating nation. The relevant national coordinating function/s will contact you as appropriate.

Please see IRAS Help for information on working with NHS/HSC organisations in Northern Ireland and Scotland.

How should I work with participating non-NHS organisations?

HRA and HCRW Approval does not apply to non-NHS organisations. You should work with your non-NHS organisations to obtain local agreement in accordance with their procedures.

What are my notification responsibilities during the study?

The *"After HRA Approval – guidance for sponsors and investigators"* document on the HRA website gives detailed guidance on reporting expectations for studies with HRA and HCRW Approval, including:

- Registration of Research
- Notifying amendments
- Notifying the end of the study

The HRA website also provides guidance on these topics and is updated in the light of changes in reporting expectations or procedures.

Who should I contact for further information?

Please do not hesitate to contact me for assistance with this application. My contact details are below.

Your IRAS project ID is **270527**. Please quote this on all correspondence.

Yours sincerely

Isobel Lyle
HRA Approvals Manager
hra.approval@nhs.net

Copy to: *Dr Michael Barber*

Appendix J: SJT survey (a sample of version 1)

PARTICIPANT INFORMATION - SUMMARY

You should have received an information sheet by email. Please take some time to read the information - it's important you understand the purpose of this study and what's involved before you decide whether you would like to take part.

We're inviting you to take part in a research study called '*What would you do?' Evaluating a situational judgement test for use in mental health services*'. The aim of this project is to try out a type of questionnaire-based assessment that we have designed to measure professional judgement in mental health services. By developing this kind of assessment we hope to make 'values based recruitment' in the NHS both fairer and more effective at choosing the right staff to work with patients. The project is part of a Hull York Medical School funded PhD project being done as part of a programme of research on recruitment and selection of the health workforce. This overall programme is funded by the National Institute for Health Research (NIHR) - the research branch of the NHS.

We're asking you if you would like to take part in this study because you are a registered clinician working in mental health and/or learning disability services. Your participation is entirely voluntary and you can withdraw at any time. We are keen to make it clear that taking part in this project, and any answers you might give in the survey, will not affect your employment in any way.

The survey will only take around 30 minutes to complete.   The survey saves automatically, so you can leave the survey and come back later.

What do I need to do to take part?

Firstly, we need to ask you a few basic 'demographic' details about who you are. This is so we can later work out if people answer the questions differently depending on their age, gender, profession etc. We then show you a series of workplace situations. These are fictional, though mostly based on actual situations that have actually happened at some point. We will ask you to imagine that you are involved in these situations as they are happening. Based on these scenarios, you will be asked to rate a behaviour from a short list of options. We recommend that you don't take too long thinking about each question and that you simply base your answer on what comes to mind first. Some of the situations are repeated during the test. However, try and forget about previous answers you gave when answering any new questions.

Next, we will ask you to fill out a short, commonly used, personality questionnaire. This is to try and work out whether certain personality types are linked to the earlier answers to the situational judgement test. We'll send you back your scores on the personality test, with a brief explanation of what they might mean, unless you say otherwise. Finally, you will be asked to provide the details of your line manager or supervisor.

Why do you ask for the contact details of my line manager/supervisor and colleagues?

We'll ask you to provide the contact details of your line manager or supervisor, and up to three colleagues. This is so we can ask them to provide a rating in relation to workplace behaviours for you. They will also be asked to complete a personality questionnaire for

you. We need these details so we can work out if the test is likely to be useful in practice. That is- are the answers to the situational judgement test associated with the views of managers and colleagues about workplace behaviours? It is important to stress: **neither your line manager, supervisor or colleagues will see your responses to this survey**; also, the ratings they provide will be kept <u>confidential and anonymised</u> (not identifiable). They will be linked to your questionnaire answers only by using a unique study code for you. This means that no one will be able to work out who the ratings were about, even if they saw the information (though only the immediate research team have access to the anonymised information).

Please note: to receive the gift voucher, you must provide valid responses to all of the questions.

If you have any questions about the research study, please contact Lauren Aylott, PhD Student, Hull York Medical School.

Email: lmea500@york.ac.uk; Telephone: 07909 702929

This project has been endorsed by Hull York Medical School's Ethics Committee.

INFORMED CONSENT:

Completion of this survey indicates that:

- You have read and understood the participant information sheet (version 2, 12/02/2020)

- You have had the opportunity to consider whether you would like to participate, and to ask questions

- You have received enough information about the study

- You understand that your participation in the study is voluntary and that you are free to withdraw at any time.

- You may ask for your data to be withdrawn from the study, up until the data is anonymised, and linked only by a unique study number. This linked data will be securely deleted as soon as the project ends (November 2021).

- You understand that any information obtained during the study, including personal data, will be kept confidential, stored securely, and only accessed by those carrying out the study

- You understand that information you give will be anonymised, aggregated, and included in published documents

- You agree to take part in this study

If you do not wish to take part in the research study, please end the survey now.

DEMOGRAPHIC DETAILS:

We would now like to collect some 'demographic' details about who you are. Please note that all information captured for this project will be kept confidential, anonymised, and will be used for research purposes only.

What is your age? _____

How do you define your gender? (please select)

◯ Female

◯ Male

◯ Other (please state)

_____

◯ Prefer not to say

Which of the following best describes how you think of yourself? (please select)

◯ Heterosexual or Straight

◯ Gay or Lesbian

◯ Bisexual

◯ Other (please state)

_____

◯ Prefer not to say

Do you consider yourself to have a disability (ie. a physical or mental impairment that has a 'substantial' and 'long term' negative effect)? (please select)

◯ Yes

◯ No

◯ Prefer not to say

What is your ethnic group? (please select)

○ Asian / Asian British (e.g. Indian, Pakistani, Chinese)

○ Black / African / Caribbean / Black British

○ Mixed / Multiple ethnic groups

○ White

○ Other (please state)

_____

○ Prefer not to say

What is your religion? (please select)

○ No religion

○ Christian

○ Buddhist

○ Hindu

○ Jewish

○ Muslim

○ Sikh

○ Other (please state)

_____

○ Prefer not to say

Is English your first language? (please select)

○ Yes

○ No

○ Prefer not to say

PLEASE NOTE: TO COMPLETE THIS SURVEY, YOU MUST BE A QUALIFIED AND REGISTERED CLINICIAN

What type of professional are you? (please select)

○ Allied Health Professional (please state) _____

○ Nurse

○ Nursing Associate

○ Pharmacist

○ Psychiatrist

○ Psychologist

○ Social Worker

○ Other (please state)

_____

How many years have you been registered with your current profession? _____

What Agenda for Change band are you? (please select)

○ 4

○ 5

○ 6

○ 7

○ 8

○ Other (please state)

_____

Do you work in inpatient, community, or corporate services? (please select the option/s that most applies)

○ Inpatient services

○ Community services

○ Corporate services

What specialty do you work in? (please select the option/s that most applies)

○ Adult Mental Health

○ Child and Adolescent

○ Forensic Learning Disabilities

○ Forensic Mental Health

○ Learning Disabilities (adults)

○ Learning Disabilities (children)

○ Mental Health Services for Older People

○ All Services, Trustwide

From this point onwards, you will be shown a set of workplace situations. Please imagine that you are involved in these situations as they are happening. Please will you then select the response that you think is most appropriate for the situation. We recommend that you don't take too long thinking about each question and that you simply base your answer on what comes to mind first. Some of the situations are repeated during the test. However, try and forget about the previous answers you gave when answering any new questions.

SCENARIO - You are working for the crisis team. It is 10pm and you receive a call from a patient who is threatening to harm themselves, noting that they are on their own. Whilst you are on the phone, a colleague informs you that they are 'an attention seeker'.

How **appropriate** would it be to respond in the following manner:

BEHAVIOUR - To ask the manager to do some staff training regarding the appropriate use of language

○ Very appropriate

○ Appropriate, but not ideal

○ Inappropriate, but not awful

○ Very inappropriate

SCENARIO - You are working for the crisis team. It is 10pm and you receive a call from a patient who is threatening to harm themselves, noting that they are on their own. Whilst you are on the phone, a colleague informs you that they are 'an attention seeker'.

How **important** is it to take into account the following consideration when deciding how to respond to the situation?

CONSIDERATION - That the patient may have wasted your colleague's time previously

○ Very important

○ Important

○ Of minor importance

○ Not important at all

SCENARIO - You are working on an inpatient ward when you overhear two patients talking about Sarah, the daughter of a colleague of yours. The patients are laughing at a photo of Sarah, which she appears to have posted on social media when she was out in the town, drinking alcohol.

How **appropriate** would it be to respond in the following manner:

BEHAVIOUR - To tell the patients that it isn't very kind to be laughing at people's posts

○ Very appropriate

○ Appropriate, but not ideal

○ Inappropriate, but not awful

○ Very inappropriate

SCENARIO - You have been asked to do a shift on a forensic mental health ward. You notice that a man on the ward hasn't been for his breakfast or lunch and you inform a fellow colleague of this. Your colleague's response was that you need to lose some of your care and compassion as the man is a criminal.

How **important** is it to take into account the following consideration when deciding how to respond to the situation?

CONSIDERATION - That the man is a criminal

○ Very important

○ Important

○ Of minor importance

○ Not important at all

SCENARIO - You have been asked to do a shift on a forensic mental health ward. You notice that a man on the ward hasn't been for his breakfast or lunch and you inform a fellow colleague of this. Your colleague's response was that you need to lose some of your care and compassion as the man is a criminal.

How **important** is it to take into account the following consideration when deciding how to respond to the situation?

CONSIDERATION - That the patient may be hungry

- ○ Very important

- ○ Important

- ○ Of minor importance

- ○ Not important at all

SCENARIO - You work in a community mental health team. You decide that an inpatient admission may be necessary for one of your patients. You contact the ward to arrange a pre-admission assessment, but the ward staff are hesitant and comment 'oh no, not them again'. You are aware that your patient has previously been an inpatient. You feel an admission is necessary, but the ward staff are reluctant to accept this, saying that the patient 'will need to be out as soon as possible'.

How **important** is it to take into account the following consideration when deciding how to respond to the situation?

CONSIDERATION - That the ward staff may neglect your patient if he is on the ward

- ○ Very important

- ○ Important

- ○ Of minor importance

- ○ Not important at all

Thanks very much for your time so far. I would be grateful if you would please share your experience of completing the test by answering the following questions.

Approximately how long has it taken you to sit the test so far?
_____

Is the test relevant to your role? (please select)

　○ Yes

　○ No

Is the difficulty of the test appropriate for your grade? (please select)

　○ Yes

　○ No

Do you think this test would be suitable for the recruitment of staff into the mental health workforce? (please select)

　○ Yes

　○ No

Do you think the test would be fair to all job applicants, regardless of their profession, gender, age, race, and other characteristics? (please select)

　○ Yes

　○ No

If you have answered no to any of the above questions, please would you explain why?

If there is anything else you would like to add regarding this test, please would you use the space here?

Next, we would like you to complete a personality questionnaire. We will send you a copy of your results once everyone has completed the survey. The personality test should take between three and five minutes and we will share your scores with you.

We would like you to complete this questionnaire so that we can see whether certain personality types are linked to answers on the Situational Judgement Test. Your answers will be kept confidential; and will only be seen by you and the research team.

The Big Five Inventory–2 Short Form (BFI-2-S)

Here are a number of characteristics that may or may not apply to you. For example, do you agree that you are someone who *likes to spend time with others?* Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement.

1	2	3	4	5
Disagree Strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree strongly

I am someone who...

1. ___ Tends to be quiet.
2. ___ Is compassionate, has a soft heart.
3. ___ Tends to be disorganized.
4. ___ Worries a lot.
5. ___ Is fascinated by art, music, or literature.
6. ___ Is dominant, acts as a leader.
7. ___ Is sometimes rude to others.
8. ___ Has difficulty getting started on tasks.
9. ___ Tends to feel depressed, blue.
10. ___ Has little interest in abstract ideas.
11. ___ Is full of energy.
12. ___ Assumes the best about people.
13. ___ Is reliable, can always be counted on.
14. ___ Is emotionally stable, not easily upset.
15. ___ Is original, comes up with new ideas.
16. ___ Is outgoing, sociable.
17. ___ Can be cold and uncaring.
18. ___ Keeps things neat and tidy.
19. ___ Is relaxed, handles stress well.
20. ___ Has few artistic interests.
21. ___ Prefers to have others take charge.
22. ___ Is respectful, treats others with respect.
23. ___ Is persistent, works until the task is finished.
24. ___ Feels secure, comfortable with self.
25. ___ Is complex, a deep thinker.
26. ___ Is less active than other people.
27. ___ Tends to find fault with others.
28. ___ Can be somewhat careless.
29. ___ Is temperamental, gets emotional easily.
30. ___ Has little creativity.

Please write your full name here (this data will be anonymised as soon as it is linked to the workplace behaviour ratings):

_____

Please would you write your email address here:

_____

You have nearly completed the survey. Before you do, we would like to collect the details of your supervisor and up to three of your colleagues.

We need the name of your supervisor (or line manager) so that we can collect some feedback for you. We will contact them shortly for this. Your supervisor (and colleagues, if applicable) will be eligible for a certificate of participation if they take part. We will only collect this feedback if your supervisor and colleagues consent to doing so.

The ratings provided will be kept confidential. If you have been under your supervisor for less than three months, please would you provide details of your prior supervisor?

NB: Your supervisor will not have access to your responses on this survey. Both ratings and responses to this survey will be anonymised, by using a participant number, prior to analysing the data.

Please will you write your line manager's or supervisor's full name here?

_____

What is their email address?

_____

What relationship are they to you and how long have they fulfilled that position?
For example, current line manager, 1 year 2 months"

_____

It would be helpful to obtain feedback from more than one person. If you are happy to do so, please would you provide the names and email addresses of up to three colleagues in your organisation, so that we can contact them also?

Colleague 1

Colleague 2

Colleague 3

Thank you for taking the time to complete this survey. If you would like a certificate of participation, please request this from Lauren Aylott.

Email: lmea500@york.ac.uk; Telephone: 07909 702929

Should you wish to add anything regarding this survey, please use the space here.

Appendix K: A sample of the form sent for SME review and scoring

Scoring of Situational Judgement Test items

INSTRUCTIONS

Thanks for taking the time to help us create a scoring system for the new 'Situational Judgment Tests' we are developing.

Please would you treat the survey content confidentially.

The survey is expected to last no longer than two hours.

We will use your responses to develop a provisional scoring key.

Don't take too long to think about each question and submit the first response that comes to mind.

When you are rating responses, please score dependent on what would be your preferred choice, meaning that the option with the highest score would be the option you would most likely select on a test, whereas the option with the lowest score would be your last choice.

Please note that you will be asked to either 1) rate responses according to how appropriate they are, or 2) rate considerations as to how important you think they are. I provide two worked examples here to help you get the hang of this.

Example 1 - Appropriateness item

Scenario - You are working on a mental health ward when you find two staff members talking about what they did at the weekend. The staff members are unaware that a patient is stood next to them and can possibly overhear them. You are aware that the patient has attachment issues and has been known to follow staff around previously.

How **appropriate** would it be to respond in the following manner?

Behaviour - To interrupt the staff members and purposefully change the topic

Scoring

In this example, the behaviour was deemed appropriate, but not ideal, hence the option 'appropriate, but not ideal' receives a score of 4 (see below). You will observe that in this example the option 'very inappropriate' received a score of 1, as this was the least preferred option.

> Very appropriate = 3
>
> Appropriate, but not ideal = 4
>
> Inappropriate, but not awful = 2
>
> Very inappropriate = 1

Example 2 - Importance item

Scenario - You are attending an outpatient appointment with a colleague. Your colleague is joking along and swearing with the patient. They were both having fun, but the patient suddenly becomes upset and it appears that they may have reacted to one of your colleague's comments. The patient suffers from paranoia and occasionally becomes psychotic.

How **important** to take into account is the following consideration when deciding how to respond to the situation?

Consideration - That the staff member used swear words

Scoring

In this example, the consideration was deemed 'very important', hence the option 'very important' receives a score of 4 (as seen below). You will observe that in this example the option 'not important at all' received a score of 1 as this was the least preferred option.

> Very important = 4
>
> Important = 3
>
> Of minor importance = 2
>
> Not important at all = 1

Don't worry too much about this now. Some example questions will be provided to help you get used to the scoring system. You can also contact me on 07453 361566, should you wish to discuss this further.

It is worth noting that not all scenarios will apply to your current role. For example, you may not work in a community or inpatient setting, or the scenario may be more applicable to a staff member of an alternative profession. In these instances, please would you best imagine yourself in this role, whilst scoring each item.

Thanks!

Please input your name here so that I can provide you with a gift voucher following completion of the survey:

The next three pages incorporate practice items that are similar to the questions you will be asked throughout the survey.

Page 1 - This is a practice item

Please would you let us know whether the following scenario is 1) clear and unambiguous, and 2) realistic and relevant?

This is so we can make sure that our test is robust and has face validity

SCENARIO		
You are working on a ward when you observe a staff member and patient being playful with one another. You observe the staff member hit the patient with a slipper, albeit gently, to which the patient laughs out loud.		
Is the scenario clear and unambiguous? *(Please delete as appropriate)*		Y / N
Is the scenario realistic and relevant? *(Please delete as appropriate)*		Y / N

If you have any comments regarding this scenario, please would you let us know here?

If you think the scenario needs revising, please would you let us know what edits you would recommend?

Page 2 - This is a practice item

Please would you let us know how appropriate the following behaviour is, in relation to the given scenario?

Please score 4 points for the option that you would primarily choose, and 1 for your least preferred option. Please rate the remaining options on a likert scale according to what you think is the best fit (4 is the best response, 1 is the worst).

SCENARIO
You are working on a ward when you observe a staff member and patient being playful with one another. You observe the staff member hit the patient with a slipper, albeit gently, to which the patient laughs out loud.
How **appropriate** would it be to respond in the following manner?
BEHAVIOUR - To report the incident to the ward manager
A very appropriate thing to do   ___ Appropriate, but not ideal   ___ Inappropriate, but not awful   ___ A very inappropriate thing to do   ___

Is this behaviour plausible (i.e. is it possible that, at least a few, clinicians would respond in this manner)? *(Please delete as appropriate)*	Y / N

If you have any comments regarding this response option, please would you let us know here?

Page 3 - This is a practice item

Please would you let us know how important the following consideration is, in relation to the scenario given?

Please score 4 points for the option that you would primarily choose, and 1 for your least preferred option. Please rate the remaining options on a likert scale according to what you think is the best fit (4 is the best response, 1 is the worst).

SCENARIO You are working on a ward when you observe a staff member and patient being playful with one another. You observe the staff member hit the patient with a slipper, albeit gently, to which the patient laughs out loud.
How **important** to take into account is the following consideration when deciding how to respond to the situation?
CONSIDERATION - That the patient appears to be having fun
Very important   ___ Important   ___ Of minor importance   ___ Not important at all   ___

Is this consideration plausible (i.e. is it possible that, at least a few, clinicians would think the consideration noted is important)? *(Please delete as appropriate)*	Y / N

If you have any comments regarding this response option, please would you let us know here?

Thanks for completing the practice questions. We hope these questions were helpful in supporting you to complete the remaining parts of the survey. From this point onwards, all questions will be counted.

NB: the remainder of the survey has been removed for item security purposes.

Appendix L: Workplace behaviours rating tool – pilot study

STUDY INFORMATION - SUMMARY

You should have received an information sheet by email. Please take some time to read the information - it's important you understand the purpose of this study and what's involved before you decide whether you would like to take part.

We're inviting you to take part in a research study called '*What would you do?' Evaluating a situational judgement test for use in mental health services*'. The aim of this project is to try out a type of questionnaire-based assessment that we have designed to measure professional judgement in mental health services. By developing this kind of assessment we hope to make 'values based recruitment' in the NHS both fairer and more effective at choosing the right staff to work with patients. The project is part of a Hull York Medical School funded PhD project being done as part of a programme of research on recruitment and selection of the health workforce. This overall programme is funded by the National Institute for Health Research (NIHR) - the research branch of the NHS.

You are being asked to take part because we would like you to provide brief and confidential feedback for a colleague. We need this feedback to see if our new test is working as we would hope. Your participation is entirely voluntary and you can withdraw at any time.

The survey will take between 5 and 10 minutes to complete.

Please would you answer as honestly as possible? **The responses you provide will be strictly confidential; your colleague will have no access to these ratings.**

If you have any questions about the research study, please contact Lauren Aylott, PhD Student, Hull York Medical School.

Email: lmea500@york.ac.uk; Telephone: 07909 702929

This project has been endorsed by the Hull York Medical School's ethics committee.

ELECTRONIC CONSENT:

Completion of this survey indicates that:

- You have read and understood the participant information sheet (version 1, 10/10/2019)
- You have had the opportunity to consider whether you would like to participate, and to ask questions
- You have received enough information about the study
- You understand that your participation in the study is voluntary and that you are free to withdraw at any time.
- You may ask for your data to be withdrawn from the study, up until the data is anonymised, and linked only by a unique study number. This linked data will be securely deleted as soon as the project ends (November 2021).
- You understand that any information obtained during the study, including personal data, will be kept confidential, stored securely, and only accessed by those carrying out the study
- You understand that information you give will be anonymized, aggregated, and included in published documents
- You agree to take part in this study

If you do not wish to take part in the research study, please end the survey now.

What is the name of your colleague that are you completing this form about?

_____

What is your name?

_____

How do you define your gender? (please select)

○ Female

○ Male

○ Other (please state)

_____

○ Prefer not to say

What relationship do you hold with your colleague (for example, are you their line manager, supervisor or colleague)?

_____

How long have you known this colleague? (please select)

○ 0 – 3 months

○ 3 – 6 months

○ 6 – 12 months

○ More than one year

We will now provide you with a list of professional attributes. Following the description of each attribute, there will be three statements. Please would you select the most applicable statement for each attribute, based on your relationship with this staff member?

Please note: these ratings will be kept absolutely confidential and used for research purposes only. Please would you respond as honestly as possible.

Commitment to Professionalism - Individuals should be committed to honouring their profession by adhering to guidelines and challenging poor practice. Professionals must have integrity and be a responsible practitioner.

With regards to your colleague, please select the statement that most applies:

○ I am happy with this individual's commitment to professionalism

○ I have concerns regarding this individual's commitment to professionalism

○ I do not know the staff member well enough to comment

Coping with Pressure - Practitioners need to utilise their clinical judgement, particularly in times of uncertainty and ambiguity. Professionals must have resilience and be able to de-escalate situations when others are experiencing distress.

With regards to your colleague, please select the statement that most applies:

○ I am happy with this individual's ability to cope with pressure

○ I have concerns regarding this individual's ability to cope with pressure

○ I do not know the staff member well enough to comment

<u>Effective Communication</u> - Practitioners need to communicate effectively, using both verbal and non-verbal communication. Professionals must have the ability to build rapport with patients, as well as validate the thoughts and feelings of others.

With regards to your colleague, please select the statement that most applies:

○ I am happy with this individual's ability to communicate effectively

○ I have concerns regarding this individual's ability to communicate effectively

○ I do not know the staff member well enough to comment

<u>Patient Focus</u> - Practitioners must possess qualities that enable them to build therapeutic relationships with patients, such as altruism and humility. Professionals, also, must maintain an appropriate professional distance and not impose their own values on patients, whilst delivering person-centred care.

With regards to your colleague, please select the statement that most applies:

○ I am happy with this individual's ability to apply patient focus

○ I have concerns regarding this individual's ability to cope with pressure

○ I do not know the staff member well enough to comment

<u>Working Effectively as Part of a Team</u> - Practitioners are part of a team and must therefore work alongside colleagues effectively, acknowledging people's strengths and capabilities. Professionals must support fellow colleagues and act as a role model for others.

With regards to your colleague, please select the statement that most applies:

○ I am happy with this individual's ability to work as part of a team

○ I have concerns regarding this individual's ability to work as part of a team

○ I do not know the staff member well enough to comment

<u>Working with Carers</u> - Practitioners must involve carers where possible, whilst adhering to the bounds of confidentiality. Carers generally want to be involved, and should feel supported in their role.

With regards to your colleague, please select the statement that most applies:

○ I am happy with this individual's ability to work with carers

○ I have concerns regarding this individual's ability to work with carers

○ I do not know the staff member well enough to comment

Please will you now think very carefully about where this staff member sits on the following scales?

In comparison to other staff members of the same profession, <u>irrespective of their grade and experience</u>, please estimate the level of professionalism of this individual.

Please use the following definition to base your judgement; please indicate your decision, by moving the below marker to the desired percentage.

Professionalism allows practitioners to make appropriate judgements in times of need, applying critical thinking, reflection and situational judgement.

	Below Average	Average	Above Average
	0 10 20 30 40 50 60 70 80 90 100		
Professionalism of this individual. ()			

Please will you give a brief example of a behaviour that characterises the person in this domain (professionalism)?

_____

In comparison to other staff members of the same profession, <u>irrespective of their grade and experience</u>, please estimate how effective in their role, this staff member is.

Please indicate your decision, by moving the below marker to the desired percentage.
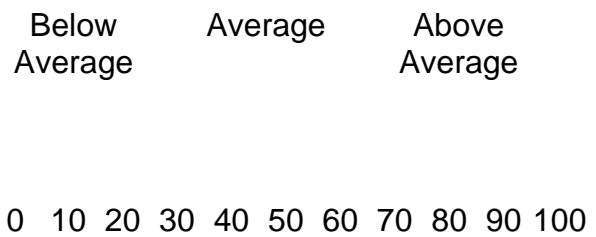
	Below Average	Average	Above Average
	0 10 20 30 40 50 60 70 80 90 100		
Effectiveness of this individual ()			

Please will you give a brief example of a behaviour that characterises the person in this domain (effectiveness)?

_____

Finally, we would like you to complete the following personality questionnaire, whilst rating **your colleague**.

The Big Five Inventory-2 Short Form (BFI-2-S)

Here are a number of characteristics that may or may not apply to your colleague. For example, do you agree that your colleague is someone who *likes to spend time with others*? Please select the most applicable option to indicate the extent to which you agree or disagree with that statement.

Please note: these ratings will be kept absolutely confidential and used for research purposes only. Please would you respond as honestly as possible.

	Disagree strongly	Disagree a little	Neutral; no opinion	Agree a little	Agree strongly
1. Tends to be quiet.	○	○	○	○	○
2. Is compassionate, has a soft heart.	○	○	○	○	○
3. Tends to be disorganized.	○	○	○	○	○
4. Worries a lot.	○	○	○	○	○
5. Is fascinated by art, music, or literature.	○	○	○	○	○
6. Is dominant, acts as a leader.	○	○	○	○	○
7. Is sometimes rude to others.	○	○	○	○	○
8. Has difficulty getting started on tasks.	○	○	○	○	○
9. Tends to feel depressed, blue.	○	○	○	○	○
10. Has little interest in abstract ideas.	○	○	○	○	○
11. Is full of energy.	○	○	○	○	○

12. Assumes the best about people.	○	○	○	○	○
13. Is reliable, can always be counted on.	○	○	○	○	○
14. Is emotionally stable, not easily upset.	○	○	○	○	○
15. Is original, comes up with new ideas.	○	○	○	○	○
16. Is outgoing, sociable.	○	○	○	○	○
17. Can be cold and uncaring.	○	○	○	○	○
18. Keeps things neat and tidy.	○	○	○	○	○
19. Is relaxed, handles stress well.	○	○	○	○	○
20. Has few artistic interests.	○	○	○	○	○
21. Prefers to have others take charge.	○	○	○	○	○
22. Is respectful, treats others with respect.	○	○	○	○	○

23. Is persistent, works until the task is finished.	○	○	○	○	○
24. Feels secure, comfortable with self.	○	○	○	○	○
25. Is complex, a deep thinker.	○	○	○	○	○
26. Is less active than other people.	○	○	○	○	○
27. Tends to find fault with others.	○	○	○	○	○
28. Can be somewhat careless.	○	○	○	○	○
29. Is temperamental, gets emotional easily.	○	○	○	○	○
30. Has little creativity.	○	○	○	○	○

Please click next if you would like to end the survey now. If there is anything else you would like to note, please would you add this here?

Appendix M: Participant information sheet – pilot study

NHS

Tees, Esk and Wear Valleys
NHS Foundation Trust

UNIVERSITY
of York

HULL
YORK
MEDICAL
SCHOOL

Participant Information Sheet for Clinicians
Online Survey

Study Title:

'What would you do?' Evaluating a situational judgement test for use in mental health services

Invitation to Participate

We are inviting you to take part in a research project. It's important that you understand the purpose of this study and what it involves before you decide whether you would like to take part. Please read this information sheet carefully and feel free to discuss it with others. If you would like more information, or if there is something you don't understand, please contact the research team. There will be no negative consequences if you choose not to take part.

Research team contact: Lauren Aylott, PhD Student, Hull York Medical School.

Email: lmea500@york.ac.uk; Telephone: 07909 702929

What's the purpose of the study?

This study is evaluating an assessment that has been created to test professional judgement in mental health services. We hope to use this test to improve 'values-based recruitment', making it fairer and more effective. First we need to check that the test works. Taking part will help us do this.

We need to check if the scores from the test are working as we would hope. To do this we are asking for brief and confidential feedback from supervisors and colleagues. The feedback ratings are confidential and any information linking either the test-taker, supervisor or colleague with the test scores will be removed by the research team. Colleagues will be offered a gift voucher for their participation.

Why have I been invited to take part?

You are being asked to take part because you are clinician, registered with a professional body, working in mental health and/or learning disability services.

What will happen if I take part?

If you decide to take part, you will be asked to complete an online survey. It will take about 30 to 35 minutes to complete the test. The test will save automatically and can be completed in more than one sitting. We hope that this will allow you to sit the test at a time that is convenient for you.

Please note that your taking part in this study, and any associated information, will be confidential and used for research purposes only.

When you begin the survey, you will be asked for some basic demographic details. These are needed so we can compare the questionnaire scores for different groups of people. You will then be shown a series of workplace situations and asked to imagine that they are actually happening. There will be a series of behaviours shown linked to these scenarios. You will be asked to rate the behaviour using a list of multiple choice options. Some of the scenarios are repeated during the test and there are 50 questions in total. We suggest not taking too long thinking about each question and answering based on whatever comes to your mind first. We have included a few further questions about how you found the test. Next, we have included a short personality questionnaire to fill in. This should only take between 3 and 5 minutes to complete. We will send you the results of the personality test, with a brief explanation of what the scores may mean. We will send these once everyone has taken part in the survey. Finally, you will be asked to give the name and contact details for your line manager or supervisor, as well as up to three of your colleagues, so they can provide brief workplace ratings for you. We need these brief ratings so we can see if our new test is working as we would hope. Your colleagues and supervisor will not see your responses and we can assure you the ratings they provide will be kept confidential, anonymised, and used for research purposes only. We are only collecting these data because it essential to understand how the new test is working.

Will you compensate me for my time?

Participants will offered a gift voucher to the value of £10. A certificate of participation can also be provided on request.

Are there any advantages in taking part?

All professionals understand that research and development is important to the NHS; taking part in this study can be used as evidence to support appraisal and revalidation.

We hope that you'll enjoy doing our survey and be able to reflect on how you might respond to certain work situations. We won't be able to send you the scores on the situational judgment test, as this is still being developed. However we will send you the results from the personality questionnaire (unless you don't want us to).

The results from this research project will be used to fine-tune our situational judgment test and the scoring system used. These assessments are already used to recruit and select staff elsewhere in healthcare. We hope that by creating a situational judgment test to use in mental health services we will improve 'values

based recruitment' and staff development in mental health and learning disability services. Therefore, it is possible that you might have to sit a situational judgement test in the future, if you apply for a new job. Taking part in this study should help you become familiar with the type of questions you may be asked.

Are there any disadvantages or risks in taking part?

We hope that you will find doing the survey interesting and don't expect any problems with taking part. The survey will take approximately 30 to 35 minutes to complete and you're free to withdraw at any time. You don't have to give a reason, and this won't disadvantage you in any way. Answers will be deleted if you don't fully complete the test.

Please be aware that there are some risks with transferring data online. We will take every precaution to make sure the information you share is kept safe.

What if I am unhappy or if there is a problem?

If you have any concerns or complaints about the study or about the way you've been approached and treated during it, you can contact the Research and Enterprise Office at the University of York: Dr Michael Barber, University of York, RCH/120, Ron Cooke Hub, Heslington, York YO10 5GE; Tel: +44 (0)1904 435154.

What will happen to my answers?

Data will be saved via the survey platform – Qualtrics. Qualtrics is recommended by the University of York and is GDPR (General Data Protection Regulation) compliant.

Data will be transferred from the online survey to an encrypted computer drive. Data will be anonymised, linked only by a unique study number, as soon as we have received the feedback from your colleagues. Anonymised data will be stored on a secure university hosted drive, only allowing access to other team members (Gabrielle Finn and Paul Tiffin).

Once the data has been anonymised, the identifiable personal details linking the unique participant study number will be stored as on a separate secure drive to that of any of the other study data and will only be accessible to the principal investigator (Lauren Aylott). This linked data and will be securely deleted as soon as the project ends (November 2021).

Feedback from colleagues will be collected by the principal investigator and will be de-identified upon collection.

We will archive the data following the study. This will be anonymised and would not identify you. Anonymised data will be kept for up to 10 years, and will be stored in a repository approved by the University of York.

There are special circumstances when we may need to share information, including if there are concerns about yours or someone else's safety. We would only share information to keep you or others safe. If we believe this is necessary at any time we will inform you and explain why.

We would like to write reports to share our findings and these may be published. We will report on the data as a whole however and will not report on any single individual. If you would like to receive a summary of the research findings, we will send you one once the study is finished.

Please see the privacy notice for additional information.

Has this project received ethical approval?

This project has been endorsed by Hull York Medical School's ethics committee.

Do I have to take part?

No. It's your choice. If you choose to take part you can withdraw at any time. You don't have to give a reason, and this won't disadvantage you in any way.

What happens next?

If you have any questions about the research study, please contact Lauren Aylott. A link to the survey will be sent to you over the next couple of days. If you would rather not take part just delete the link when it arrives.

Lauren Aylott, PhD Student, Hull York Medical School

Email: lmea500@york.ac.uk

Telephone: 07909 702929

Thank you for taking the time to read this information sheet.

Appendix N: Privacy notice – pilot study

Evaluating a Situational Judgement Test for mental health services, version 1.0

NHS
Tees, Esk and Wear Valleys
NHS Foundation Trust

UNIVERSITY of York

HULL YORK MEDICAL SCHOOL

Supplementary Privacy Notice

Study Title:

'What would you do?' Evaluating a situational judgement test for use in mental health services

On what basis will you process my data?

Under the General Data Protection Regulation (GDPR), the University has to identify a legal basis for processing personal data and, where appropriate, an additional condition for processing special category data.

In line with our charter which states that we advance learning and knowledge by teaching and research, the University processes personal data for research purposes under Article 6 (1) (e) of the GDPR:

Processing is necessary for the performance of a task carried out in the public interest

Special category data is processed under Article 9 (2) (j):

Processing is necessary for archiving purposes in the public interest, or scientific and historical research purposes or statistical purposes

Research will only be undertaken where ethical approval has been obtained, where there is a clear public interest and where appropriate safeguards have been put in place to protect data.

In line with ethical expectations and in order to comply with common law duty of confidentiality, we will seek your consent to participate where appropriate. This consent will not, however, be our legal basis for processing your data under the GDPR.

How will you use my data?

Scores on the survey will be evaluated and compared to ratings provided by colleagues. In some instances, where you have agreed to this, survey scores will be compared to scores on the Independent Assessment of Clinical Skills. Anonymised data will be stored in a secure folder on the university shared drive, allowing access to other team members (Gabrielle Finn and Paul Tiffin).

Will you share my data with 3rd parties?

We will archive the data following the study. This will be anonymised and would not identify you. Anonymised data will be kept for up to 10 years, and will be stored in a repository approved by the University.

Version 1 10.10.2019   IRAS ID: 270527

Anonymised data may be reused by the research team or other third parties for secondary research purposes.

How will you keep my data secure?

The University will put in place appropriate technical and organisational measures to protect your personal data and/or special category data. For the purposes of this project we will transfer data to an encrypted computer drive. Data will be anonymised, linked only by a unique study number, as soon as we have received the feedback from colleagues. Anonymised data will be stored on a secure university hosted drive, only allowing access to other team members (Gabrielle Finn and Paul Tiffin).

Once the data has been anonymised, the identifiable personal details linking the unique participant study number will be stored as on a separate secure drive to that of any of the other study data and will only be accessible to the principal investigator (Lauren Aylott). This linked data and will be securely deleted as soon as the project ends (November 2021).

Information will be treated confidentiality and shared on a need-to-know basis only. The University is committed to the principle of data protection by design and default and will collect the minimum amount of data necessary for the project. In addition, we will anonymise or pseudonymise data wherever possible.

Will you transfer my data internationally?

No. Data will be held within the European Economic Area in full compliance with data protection legislation.

Will I be identified in any research outputs?

We would like to write reports to share our findings and these may be published. We will report on the data as a whole however and will not report on any single individual. If you would like to receive a summary of the research findings, we will send you one once the study is completed.

How long will you keep my data?

Data will be retained in line with legal requirements or where there is a business need. Retention timeframes will be determined in line with the University's Records Retention Schedule.

What rights do I have in relation to my data?

Under the GDPR, you have a general right of access to your data, a right to rectification, erasure, restriction, objection or portability. You also have a right to withdrawal. Please note, not all rights apply where data is processed purely for research purposes. For further information see, https://www.york.ac.uk/records-management/generaldataprotectionregulation/individualsrights/.

Participants will be unable to access their supervisor and peer ratings as this would seriously impair the aims of the study. In addition, we won't be able to send you the scores on the situational judgment test, as this is still being developed. Individuals will have a right to access the demographic details they provided, as well as the data provided for the personality test. Participants may request for their data to be erased, up until this data has been deidentified, at which point we would be unable to delete this.

Questions or concerns

If you have any questions about this privacy notice or concerns about how your data is being processed, please contact Lauren Aylott in the first instance. If you are still dissatisfied, please contact the University of York's Data Protection Officer at dataprotection@york.ac.uk.

Right to complain

If you are unhappy with the way in which the University has handled your personal data, you have a right to complain to the Information Commissioner's Office. For information on reporting a concern to the Information Commissioner's Office, see www.ico.org.uk/concerns.

Lauren Aylott, PhD Student, Hull York Medical School

Email: lmea500@york.ac.uk

Telephone: 07909 702929

Thank you for taking the time to read this privacy notice

Appendix O: Distribution of SJT scores across all four scoring approaches for the entire sample

Histogram displaying the distribution of SJT scores using raw SME polytomous scoring
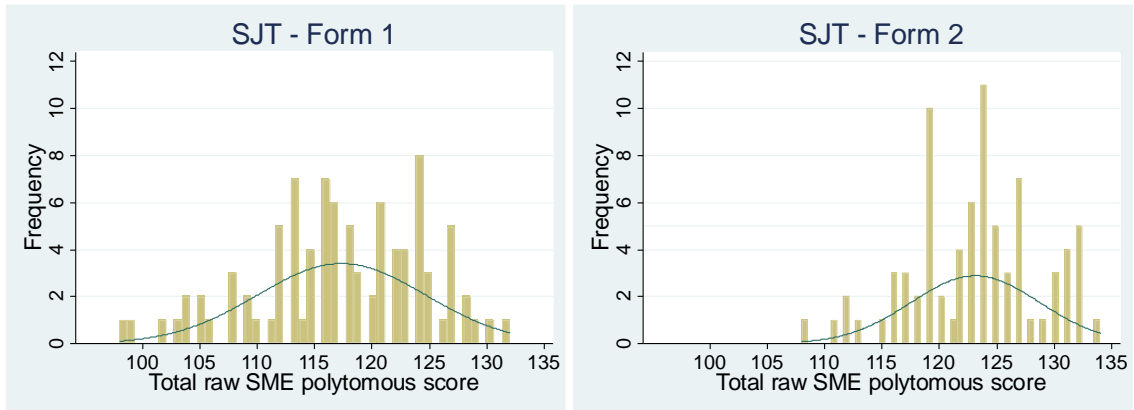
Histogram displaying the distribution of SJT scores using raw SME binary scoring
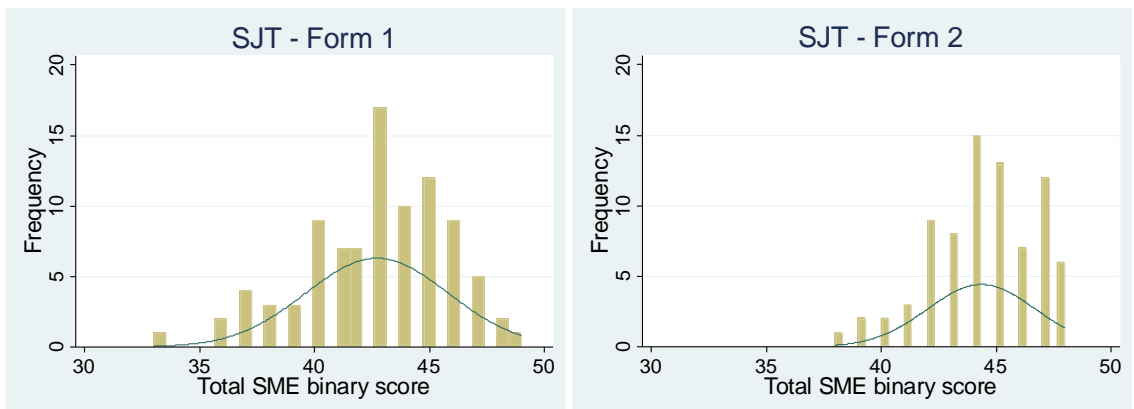
Histogram displaying the distribution of SJT scores using dichotomous modal consensus scoring
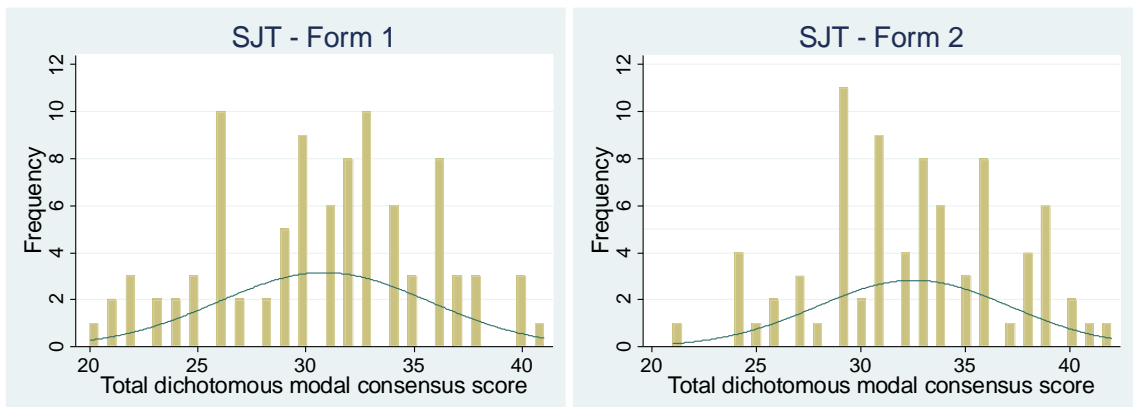
Histogram displaying the distribution of SJT scores using proportion modal consensus scoring
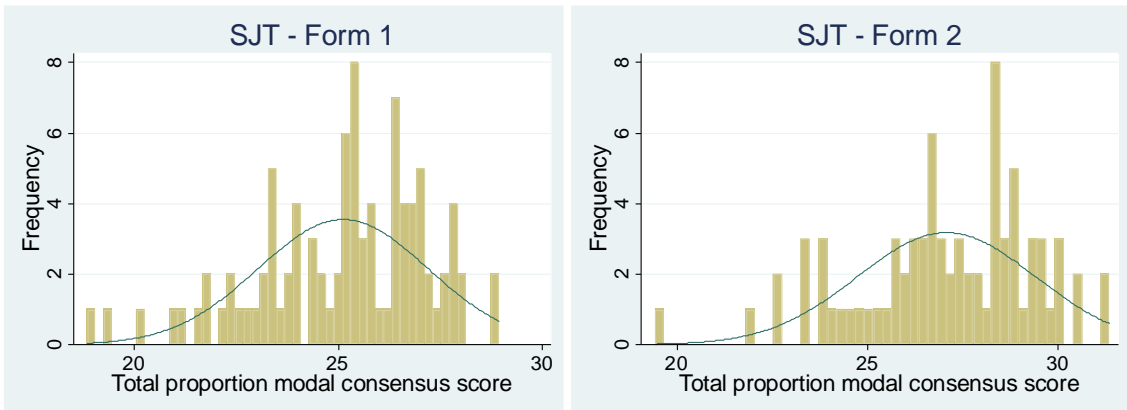
Appendix P: Distribution of SJT scores in the nurse and allied health professional sample

Histogram displaying the distribution of SJT scores obtained by nurses and allied health professionals using discipline specific dichotomous modal consensus scoring
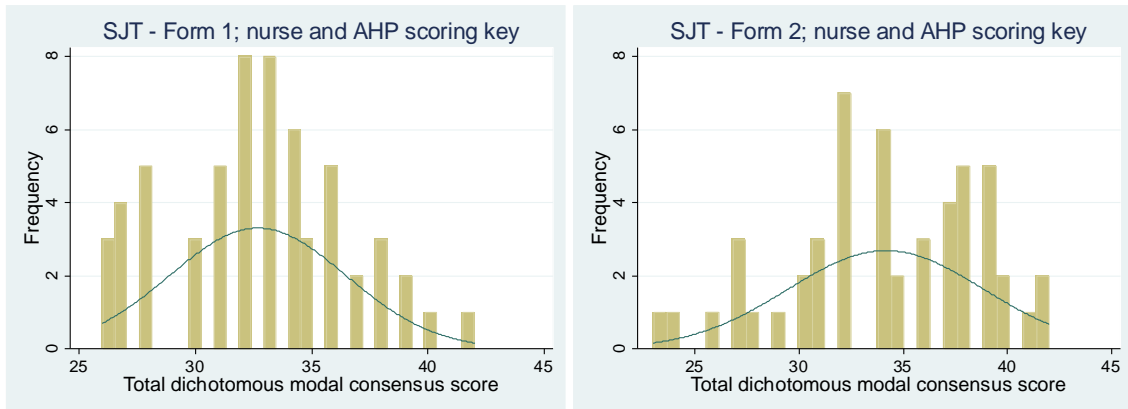
Histogram displaying the distribution of SJT scores obtained by nurses and allied health professionals using discipline specific proportion modal consensus scoring
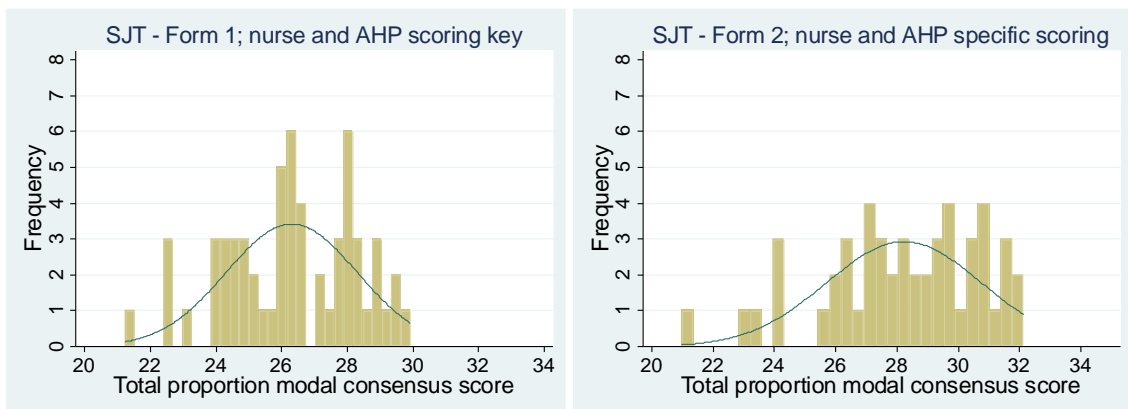
Appendix Q: Face validity of the SJT, across the entire sample, by form

	Overall (n=170) N (%)	Form 1 (n=92) N (%)	Form 2 (n=78) N (%)
Relevant to role			
- Yes	160 (94.1)	89 (96.7)	71 (91.0)
- No	10 (5.9)	3 (3.3)	7 (9.0)
Appropriate difficulty for grade			
- Yes	160 (94.1)	89 (96.7)	71 (91.0)
- No	10 (5.9)	3 (3.3)	7 (9.0)
Suitable for recruitment			
- Yes	155 (91.2)	82 (89.1)	73 (93.6)
- No	15 (8.8)	10 (10.9)	5 (6.4)
Fair to applicants			
- Yes	143 (84.1)	79 (85.9)	64 (82.1)
- No	27 (15.9)	13 (14.1)	14 (18.0)

Appendix R: SJT items relationship with professionalism for the nurse and allied health professional sample using dichotomous modal consensus scoring

SJT item	N	Odds ratio
Item 1	30	1.052
Item 2	30	1.014
*Item 3	30	0.972
Item 4	30	1.143
*Item 5	30	0.893
Item 6	30	1.025
Item 7	30	1.007
*Item 8	30	0.938
Item 9	69	1.033
Item 10	69	0.988
Item 11	69	1.001
Item 12	69	0.982
Item 13	69	1.000
Item 14	69	0.986
Item 15	69	1.027
Item 16	69	1.003
Item 17	69	1.056
Item 18	69	1.014
Item 19	30	1.039
Item 20	30	1.044
Item 21	30	1.270
Item 22	30	1.009
Item 23	30	1.078
Item 24	30	1.076
Item 25	30	1.001
*Item 26	30	0.912
*Item 27	30	0.963
Item 28	30	1.072
Item 29	30	0.995
Item 30	30	1.073
Item 31	30	1.036
*Item 32	30	0.978
Item 33	30	1.008
Item 34	30	1.017
Item 35	30	0.983
Item 36	30	1.078
Item 37	30	1.080
Item 38	30	1.070
Item 39	30	1.020
Item 40	30	1.026
Item 41	30	0.989
Item 42	30	1.030
Item 43	30	0.981
Item 44	30	1.000
Item 45	30	1.035

Item 46	30	1.005
Item 47	30	0.997
Item 48	30	1.009
Item 49	30	0.990
Item 50	30	1.025
Item 51	39	1.034
Item 52	39	1.053
*Item 53	39	0.890
*Item 54	39	0.942
Item 55	39	1.065
Item 56	39	1.002
Item 57	39	0.994
Item 58	39	1.011
*Item 59	39	0.971
Item 60	39	1.044
*Item 61	39	0.977
*Item 62	39	0.943
*Item 63	39	0.964
Item 64	39	1.012
*Item 65	39	0.971
Item 66	39	1.011
Item 67	39	1.047
Item 68	39	0.998
Item 69	39	1.019
Item 70	39	1.004
Item 71	39	1.007
Item 72	39	0.982
*Item 73	39	*Does not vary*
Item 74	39	1.052
*Item 75	39	0.966
Item 76	39	1.007
Item 77	39	1.008
Item 78	39	1.039
Item 79	39	1.025
Item 80	39	1.023
Item 81	39	0.992
Item 82	39	1.009
Item 83	39	1.034
Item 84	39	1.066
Item 85	39	1.009
Item 86	39	1.002
Item 87	39	1.007
Item 88	39	1.006
Item 89	39	1.001
Item 90	39	0.999

*Items that are not scored in the final SJT assessment are denoted by an Asterix

Appendix S: Factor analysis output for the trimmed forms

FACTOR OUTPUT _ FORM 1 (FINAL ITEMS ONLY)

```
                          F A C T O R

                  Unrestricted Factor Analysis


                  Release Version 10.5.01  x64bits
                          April, 2017
                  Rovira i Virgili University
                      Tarragona, SPAIN


                  Date: Monday, June 27, 2022
                  Time: 11:30:32
-----------------------------------------------------------------------


DETAILS OF ANALYSIS

Method to handle missing values                 : Hot-Deck Multiple Imputation in Exploratory
Factor Analysis (Lorenzo-Seva & Van Ginkel, 2016)
Missing code value                              : 999
Number of participants                          : 59
Number of variables                             : 44
Variables included in the analysis              : ALL
Variables excluded in the analysis              : NONE
Number of factors                               : 1
Number of second order factors                  : 0
Procedure for determining the number of dimensions  : Optimal implementation of Parallel Analysis
(PA) (Timmerman, & Lorenzo-Seva, 2011)
```

```
Dispersion matrix                              : Polychoric Correlations
Robust analyses based on bootstrap             : None
Method for factor extraction                   : Unweighted Least Squares (ULS)
Rotation to achieve factor simplicity          : Promin (Lorenzo-Seva, 1999)
Clever rotation start                          : Weighted Varimax
Number of random starts                        : 100
Maximum number of iterations                   : 1000
Convergence value                              : 0.00001000


-----------------------------------------------------------------------------


ADEQUACY OF THE CORRELATION MATRIX


Determinant of the matrix      = 0.000000000003568
Bartlett's statistic           =  1120.3 (df =   946; P = 0.000073)
Kaiser-Meyer-Olkin (KMO) test = 0.28156 (inacceptable)


-----------------------------------------------------------------------------


EXPLAINED VARIANCE BASED ON EIGENVALUES
Variable  Eigenvalue   Proportion of   Cumulative Proportion
                       Variance        of Variance


    1      2.43313      0.05530         0.05530
    2      1.95901      0.04452
    3      1.82332      0.04144
    4      1.71279      0.03893
    5      1.68369      0.03827
    6      1.57940      0.03590
    7      1.55013      0.03523
    8      1.49323      0.03394
```

```
 9      1.39353      0.03167
10      1.33894      0.03043
11      1.28881      0.02929
12      1.23613      0.02809
13      1.21113      0.02753
14      1.19589      0.02718
15      1.15306      0.02621
16      1.09664      0.02492
17      1.07183      0.02436
18      1.05628      0.02401
19      1.03044      0.02342
20      1.01788      0.02313
21      1.00739      0.02290
22      0.98688      0.02243
23      0.98177      0.02231
24      0.94978      0.02159
25      0.93343      0.02121
26      0.91482      0.02079
27      0.88068      0.02002
28      0.84599      0.01923
29      0.82652      0.01878
30      0.80403      0.01827
31      0.75342      0.01712
32      0.74335      0.01689
33      0.71770      0.01631
34      0.65188      0.01482
35      0.60724      0.01380
36      0.56571      0.01286
37      0.48935      0.01112
38      0.42879      0.00975
39      0.40052      0.00910
```

```
40        0.36613        0.00832
41        0.33605        0.00764
42        0.30440        0.00692
43        0.16669        0.00379
44        0.01219        0.00028
```

PARALLEL ANALYSIS (PA) BASED ON MINIMUM RANK FACTOR ANALYSIS
(Timmerman & Lorenzo-Seva, 2011)

Implementation details:

    Correlation matrices analysed: Polychoric correlation matrices
    Number of random correlation matrices: 500
    Method to obtain random correlation matrices: Permutation of the raw data (Buja & Eyuboglu,
1992)

Variable	Real-data % of variance	Mean of random % of variance	95 percentile of random % of variance
1	5.9**	5.2	5.7
2	4.8*	4.7	5.2
3	4.5*	4.4	4.8
4	4.2*	4.2	4.5
5	3.9	4.0	4.3
6	3.8	3.8	4.1
7	3.6	3.6	3.9
8	3.5	3.5	3.7
9	3.4	3.3	3.6
10	3.1	3.2	3.4

** Advised number of dimensions when 95 percentile is considered: 1
* Advised number of dimensions when mean is considered: 4

GOODNESS OF FIT STATISTICS

Goodness of Fit Index (GFI) =   0.846

UNROTATED LOADING MATRIX

Variable   F   1   Communality

V   1   -0.003   0.000
V   2   -0.000   0.000
V   3   -0.507   0.257
V   4   -0.112   0.012
V   5   0.075   0.006
V   6   0.007   0.000
V   7   -0.040   0.002
V   8   -0.366   0.134
V   9   0.006   0.000
V   10   0.001   0.000

MPLUS OUTPUT _ FORM 1 (FINAL ITEMS ONLY)

```
Mplus VERSION 8.6
MUTHEN & MUTHEN
06/27/2022   2:42 PM


SUMMARY OF ANALYSIS


Number of groups                                                1
Number of observations                                         59


Number of dependent variables                                 44
Number of independent variables                                0
Number of continuous latent variables                          0


Estimator                                                  WLSMV
Rotation                                                  GEOMIN
Row standardization                                  CORRELATION
Type of rotation                                         OBLIQUE
Epsilon value                                             Varies
Maximum number of iterations                                1000
Convergence criterion                                  0.500D-04
Maximum number of steepest descent iterations                 20
Maximum number of iterations for H1                         2000
Convergence criterion for H1                           0.100D-03
Optimization Specifications for the Exploratory Factor Analysis
Rotation Algorithm
  Number of random starts                                     30
  Maximum number of iterations                             10000
  Derivative convergence criterion                      0.100D-04
Link                                                      PROBIT
```

```
SUMMARY OF MODEL FIT INFORMATION


                     Number of                         Degrees of
          Model      Parameters      Chi-Square        Freedom      P-Value


          1-factor       44           961.631            902        0.0823
          2-factor       87           902.352            859        0.1481
          3-factor      129           847.170            817        0.2255
          4-factor      170           796.694            776        0.2954
          5-factor      210           754.484            736        0.3103


                                                      Degrees of
          Models Compared             Chi-Square      Freedom      P-Value


          1-factor against 2-factor    62.659            43        0.0267
          2-factor against 3-factor    57.043            42        0.0607
          3-factor against 4-factor    52.125            41        0.1142
          4-factor against 5-factor    44.247            40        0.2970



EXPLORATORY FACTOR ANALYSIS WITH 1 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters                    44

Chi-Square Test of Model Fit

        Value                            961.631*
        Degrees of Freedom                  902
        P-Value                          0.0823
```

RMSEA (Root Mean Square Error Of Approximation)

       Estimate   0.033
       90 Percent C.I.   0.000 0.051
       Probability RMSEA <= .05   0.938

CFI/TLI

       CFI   0.567
       TLI   0.546

Chi-Square Test of Model Fit for the Baseline Model

       Value   1083.786
       Degrees of Freedom   946
       P-Value   0.0000

SRMR (Standardized Root Mean Square Residual)

       Value   0.218

MINIMUM ROTATION FUNCTION VALUE   6.99182

       GEOMIN ROTATED LOADINGS (* significant at 5% level)
                  1

	1
S_MODAL_1	0.018
S_MODAL_2	0.023
S_MODAL_4	0.526*
S_MODAL_6	0.136

```
S_MODAL_7       -0.233
S_MODAL_9       -0.141
S_MODAL_10       0.051
S_MODAL_11       0.635*
S_MODAL_12       0.275
S_MODAL_13      -0.179
S_MODAL_14       0.159
S_MODAL_15       0.363*
S_MODAL_16       0.252
S_MODAL_17       0.598*
S_MODAL_18       0.162
S_MODAL_19       0.478*
S_MODAL_20      -0.042
S_MODAL_21       0.146
S_MODAL_22      -0.078
S_MODAL_23       0.559*
S_MODAL_24      -0.108
S_MODAL_25       0.310*
S_MODAL_28       0.348
S_MODAL_29      -0.339*
S_MODAL_30      -0.002
S_MODAL_31       0.379*
S_MODAL_33       0.010
S_MODAL_34       0.658*
S_MODAL_35       0.113
S_MODAL_36       0.084
S_MODAL_37       0.274
S_MODAL_38       0.816*
S_MODAL_39       0.260
S_MODAL_40       0.409*
S_MODAL_41       0.489*
```

```
S_MODAL_42        -0.384*
S_MODAL_43        -0.442*
S_MODAL_44        -0.629*
S_MODAL_45        -0.596*
S_MODAL_46         0.715*
S_MODAL_47         0.587*
S_MODAL_48        -0.564*
S_MODAL_49         0.402*
S_MODAL_50         0.588*


EXPLORATORY FACTOR ANALYSIS WITH 2 FACTOR(S):


MODEL FIT INFORMATION
Number of Free Parameters                      87


Chi-Square Test of Model Fit


        Value                            902.352*
        Degrees of Freedom                    859
        P-Value                            0.1481


RMSEA (Root Mean Square Error Of Approximation)


        Estimate                            0.029
        90 Percent C.I.            0.000   0.048
        Probability RMSEA <= .05            0.964


CFI/TLI


        CFI                                 0.685
        TLI                                 0.654
```

Chi-Square Test of Model Fit for the Baseline Model

          Value   1083.786
          Degrees of Freedom   946
          P-Value   0.0000

SRMR (Standardized Root Mean Square Residual)

          Value   0.200

MINIMUM ROTATION FUNCTION VALUE   3.32025

               GEOMIN ROTATED LOADINGS (* significant at 5% level)
                    1   2

                 _____   _____
 S_MODAL_1   0.006   -0.002
 S_MODAL_2   0.022   0.017
 S_MODAL_4   -0.345   0.578*
 S_MODAL_6   -0.629*   0.264
 S_MODAL_7   0.552*   -0.348
 S_MODAL_9   -0.232   -0.099
 S_MODAL_10   -0.256   0.096
 S_MODAL_11   0.143   0.621*
 S_MODAL_12   0.128   0.256
 S_MODAL_13   -0.124   -0.158
 S_MODAL_14   0.066   0.144
 S_MODAL_15   0.325   0.301*
 S_MODAL_16   -0.698*   0.358
 S_MODAL_17   0.163   0.580*
 S_MODAL_18   0.704*   0.003

```
S_MODAL_19          0.353           0.423*
S_MODAL_20         -0.494*          0.040
S_MODAL_21          0.240           0.121
S_MODAL_22         -0.247          -0.034
S_MODAL_23         -0.200           0.608*
S_MODAL_24         -0.279          -0.055
S_MODAL_25          0.141           0.289
S_MODAL_28         -0.169           0.389*
S_MODAL_29          0.322*         -0.405*
S_MODAL_30         -0.169           0.032
S_MODAL_31         -0.066           0.399*
S_MODAL_33         -0.287           0.053
S_MODAL_34          0.086           0.656*
S_MODAL_35          0.249           0.067
S_MODAL_36         -0.068           0.095
S_MODAL_37          0.360           0.211
S_MODAL_38         -0.010           0.830*
S_MODAL_39          0.562*          0.149
S_MODAL_40          0.691*          0.290
S_MODAL_41          0.042           0.488*
S_MODAL_42         -0.187          -0.351*
S_MODAL_43         -0.078          -0.436*
S_MODAL_44         -0.230          -0.591*
S_MODAL_45         -0.384*         -0.528*
S_MODAL_46         -0.028           0.726*
S_MODAL_47          0.421*          0.512*
S_MODAL_48          0.133          -0.598*
S_MODAL_49          0.044           0.394*
S_MODAL_50         -0.356           0.659*
```

```
EXPLORATORY FACTOR ANALYSIS WITH 3 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters                        129

Chi-Square Test of Model Fit

        Value                            847.170*
        Degrees of Freedom                   817
        P-Value                           0.2255

RMSEA (Root Mean Square Error Of Approximation)

        Estimate                           0.025
        90 Percent C.I.              0.000  0.046
        Probability RMSEA <= .05           0.978

CFI/TLI

        CFI                                0.781
        TLI                                0.746

Chi-Square Test of Model Fit for the Baseline Model

        Value                           1083.786
        Degrees of Freedom                   946
        P-Value                           0.0000
```

SRMR (Standardized Root Mean Square Residual)

            Value   0.185

MINIMUM ROTATION FUNCTION VALUE   2.15986

              GEOMIN ROTATED LOADINGS (* significant at 5% level)
                      1   2   3

                  _____   _____   _____
  S_MODAL_1   -0.006   -0.044   0.075
  S_MODAL_2   -0.020   0.013   0.039
  S_MODAL_4   0.667*   -0.018   0.166
  S_MODAL_6   0.648*   0.036   -0.356*
  S_MODAL_7   -0.686*   0.249   0.019
  S_MODAL_9   0.070   0.093   -0.355*
  S_MODAL_10   0.270   -0.138   -0.053
  S_MODAL_11   0.311   0.079   0.538*
  S_MODAL_12   0.058   0.217   0.155
  S_MODAL_13   0.025   -0.356*   0.018
  S_MODAL_14   0.059   -0.205   0.327
  S_MODAL_15   -0.032   0.031   0.474*
  S_MODAL_16   0.778*   -0.105   -0.280
  S_MODAL_17   0.222   0.362*   0.352*
  S_MODAL_18   -0.459   -0.002   0.539*
  S_MODAL_19   -0.079   0.773*   0.051
  S_MODAL_20   0.434*   -0.376   -0.107
  S_MODAL_21   -0.006   0.535*   -0.296
  S_MODAL_22   0.163   -0.450*   0.189
  S_MODAL_23   0.508*   0.411*   0.005
  S_MODAL_24   0.090   0.421*   -0.642*
  S_MODAL_25   0.000   0.499*   -0.028

```
S_MODAL_28          0.318           0.612*          -0.259
S_MODAL_29         -0.556*          0.177           -0.162
S_MODAL_30          0.152          -0.029           -0.105
S_MODAL_31          0.269           0.361*          -0.020
S_MODAL_33          0.297          -0.547*           0.251
S_MODAL_34          0.374           0.588*           0.018
S_MODAL_35         -0.091          -0.200            0.421*
S_MODAL_36          0.124          -0.075            0.066
S_MODAL_37         -0.211           0.589*          -0.003
S_MODAL_38          0.535*          0.236            0.478*
S_MODAL_39         -0.259          -0.010            0.581*
S_MODAL_40         -0.362           0.253            0.690*
S_MODAL_41          0.255           0.101            0.387*
S_MODAL_42         -0.117           0.236           -0.614*
S_MODAL_43         -0.188          -0.319*          -0.197
S_MODAL_44         -0.190          -0.393*          -0.379*
S_MODAL_45         -0.022          -0.299           -0.543*
S_MODAL_46          0.483*         -0.012            0.596*
S_MODAL_47         -0.014           0.392*           0.482*
S_MODAL_48         -0.475*         -0.136           -0.280
S_MODAL_49          0.273          -0.089            0.409*
S_MODAL_50          0.686*          0.080            0.171
```

EXPLORATORY FACTOR ANALYSIS WITH 4 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters   170

Chi-Square Test of Model Fit

        Value   796.694*
        Degrees of Freedom   776
        P-Value   0.2954

RMSEA (Root Mean Square Error Of Approximation)

        Estimate   0.021
        90 Percent C.I.   0.000   0.045
        Probability RMSEA <= .05   0.984

CFI/TLI

        CFI   0.850
        TLI   0.817

Chi-Square Test of Model Fit for the Baseline Model

        Value   1083.786
        Degrees of Freedom   946
        P-Value   0.0000

SRMR (Standardized Root Mean Square Residual)

        Value   0.169

MINIMUM ROTATION FUNCTION VALUE   2.54901

```
          GEOMIN ROTATED LOADINGS (* significant at 5% level)
                   1              2              3              4
              _____     _____     _____     _____
S_MODAL_1        -0.183         -0.092         -0.129          0.253
S_MODAL_2        -0.168         -0.115         -0.033          0.220
S_MODAL_4         0.490          0.037         -0.046          0.430
S_MODAL_6         0.730*        -0.229          0.202         -0.025
S_MODAL_7        -0.640*         0.052          0.201         -0.162
S_MODAL_9        -0.043         -0.446*         0.141          0.110
S_MODAL_10        0.274         -0.065         -0.073          0.014
S_MODAL_11        0.435          0.692*         0.072          0.043
S_MODAL_12       -0.035          0.120          0.135          0.214
S_MODAL_13        0.038         -0.038         -0.331         -0.072
S_MODAL_14        0.040          0.270         -0.239          0.077
S_MODAL_15        0.093          0.550*         0.029         -0.032
S_MODAL_16        0.706*        -0.308         -0.020          0.172
S_MODAL_17       -0.023          0.232          0.178          0.506
S_MODAL_18       -0.020          0.852*         0.059         -0.555
S_MODAL_19       -0.068          0.193          0.733*         0.111
S_MODAL_20        0.364         -0.209         -0.336          0.104
S_MODAL_21        0.298          0.077          0.547*        -0.395
S_MODAL_22        0.018         -0.033         -0.434*         0.204
S_MODAL_23        0.278         -0.104          0.353          0.515*
S_MODAL_24        0.054         -0.583*         0.506*        -0.003
S_MODAL_25       -0.153         -0.096          0.482*         0.295
S_MODAL_28        0.106         -0.302          0.558*         0.389
S_MODAL_29       -0.629*        -0.239          0.138         -0.020
S_MODAL_30        0.312          0.053          0.100         -0.203
S_MODAL_31        0.315          0.086          0.422*         0.064
S_MODAL_33        0.114         -0.014         -0.580*         0.280
S_MODAL_34        0.468*         0.262          0.578*        -0.007
```

```
S_MODAL_35        -0.242         0.236        -0.349*         0.252
S_MODAL_36         0.025        -0.023        -0.122          0.152
S_MODAL_37        -0.299        -0.016         0.561*         0.189
S_MODAL_38         0.297         0.345         0.090          0.587*
S_MODAL_39         0.008         0.772*       -0.044         -0.290
S_MODAL_40        -0.636         0.470*        0.002          0.488
S_MODAL_41        -0.086         0.110        -0.069          0.601*
S_MODAL_42        -0.177        -0.610*        0.267         -0.060
S_MODAL_43         0.163         0.037        -0.104         -0.587*
S_MODAL_44        -0.056        -0.342        -0.271         -0.381
S_MODAL_45         0.324        -0.284        -0.024         -0.634
S_MODAL_46         0.139         0.321        -0.225          0.686*
S_MODAL_47        -0.113         0.455         0.262          0.314
S_MODAL_48        -0.090        -0.002         0.078         -0.691*
S_MODAL_49         0.399         0.488*       -0.081         -0.025
S_MODAL_50         0.511         0.082         0.041          0.434
```

EXPLORATORY FACTOR ANALYSIS WITH 5 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters   210

Chi-Square Test of Model Fit

        Value   754.484*
        Degrees of Freedom   736
        P-Value   0.3103

RMSEA (Root Mean Square Error Of Approximation)

      Estimate   0.021
      90 Percent C.I.   0.000 0.045
      Probability RMSEA <= .05   0.983

CFI/TLI

      CFI   0.866
      TLI   0.828

Chi-Square Test of Model Fit for the Baseline Model

      Value   1083.786
      Degrees of Freedom   946
      P-Value   0.0000

SRMR (Standardized Root Mean Square Residual)

      Value   0.157

MINIMUM ROTATION FUNCTION VALUE   2.14381

```
          GEOMIN ROTATED LOADINGS (* significant at 5% level)
                1           2           3           4           5
            _____    _____    _____    _____    _____

 S_MODAL_1    -0.125       0.009      -0.146       0.236      -0.202
 S_MODAL_2     0.227       0.217      -0.088      -0.028      -0.225
 S_MODAL_4     0.849*      0.025       0.049       0.149      -0.015
 S_MODAL_6     0.367*     -0.556*      0.329      -0.033       0.004
 S_MODAL_7     0.056       0.818*      0.017      -0.534*     -0.118
```

S_MODAL_9	0.154	-0.028	0.099	-0.204	-0.410*
S_MODAL_10	0.724*	0.069	-0.078	-0.306	0.027
S_MODAL_11	0.081	-0.072	0.215	0.458	0.623*
S_MODAL_12	-0.298	-0.031	0.176	0.398*	-0.013
S_MODAL_13	-0.071	-0.214	-0.308	0.038	0.052
S_MODAL_14	0.455*	0.231	-0.270	-0.008	0.204
S_MODAL_15	0.042	0.204	0.071	0.186	0.454*
S_MODAL_16	0.785*	-0.416	0.079	-0.105	-0.112
S_MODAL_17	-0.106	0.104	0.242	0.615*	-0.043
S_MODAL_18	-0.098	0.446*	0.053	-0.213	0.847*
S_MODAL_19	-0.064	0.411	0.716*	0.035	0.002
S_MODAL_20	-0.055	-0.611*	-0.232	0.285	-0.063
S_MODAL_21	-0.041	0.010	0.593*	-0.349	0.243
S_MODAL_22	0.421*	-0.015	-0.456*	0.057	-0.064
S_MODAL_23	-0.029	-0.224	0.463*	0.549*	-0.250
S_MODAL_24	0.136	-0.042	0.465*	-0.397	-0.492*
S_MODAL_25	0.026	0.288	0.439*	0.079	-0.287
S_MODAL_28	0.171	0.041	0.560*	0.111	-0.420*
S_MODAL_29	-0.337	0.395*	-0.004	-0.236	-0.346
S_MODAL_30	0.090	-0.192	0.148	-0.105	0.186
S_MODAL_31	0.271	0.029	0.470*	-0.005	0.078
S_MODAL_33	0.632*	-0.017	-0.607*	0.078	-0.034
S_MODAL_34	0.056	-0.077	0.687*	0.164	0.258
S_MODAL_35	0.002	0.209	-0.363	0.297	0.053
S_MODAL_36	0.397*	0.139	-0.154	-0.057	-0.079
S_MODAL_37	-0.130	0.442	0.497*	0.000	-0.228
S_MODAL_38	0.336*	0.038	0.193	0.641*	0.095
S_MODAL_39	0.037	0.380*	-0.033	-0.008	0.702*
S_MODAL_40	-0.035	0.800*	-0.093	0.396	-0.011
S_MODAL_41	0.123	0.125	-0.047	0.551*	-0.165
S_MODAL_42	-0.002	0.017	0.178	-0.434	-0.530*

```
S_MODAL_43      0.070      -0.106      -0.118      -0.519*      0.306
S_MODAL_44     -0.185      -0.280      -0.307      -0.372      -0.094
S_MODAL_45     -0.008      -0.423*     -0.012      -0.580*      0.113
S_MODAL_46      0.283       0.021      -0.137       0.758*      0.036
S_MODAL_47      0.054       0.427*      0.267       0.350       0.166
S_MODAL_48     -0.204       0.062       0.015      -0.633*      0.237
S_MODAL_49      0.331      -0.020       0.007       0.159       0.502*
S_MODAL_50      0.592*     -0.169       0.139       0.348       0.012
```

FACTOR OUTPUT _ FORM 2 (FINAL ITEMS ONLY)

```
                        F A C T O R

              Unrestricted Factor Analysis

              Release Version 10.5.01  x64bits
                      April, 2017
              Rovira i Virgili University
                   Tarragona, SPAIN

              Date: Friday, July 01, 2022
              Time: 16:12:8
--------------------------------------------------------------------------

DETAILS OF ANALYSIS

Number of participants                            : 50
Number of variables                               : 41
Variables included in the analysis                : ALL
Variables excluded in the analysis                : NONE
Number of factors                                 : 1
Number of second order factors                    : 0
Procedure for determining the number of dimensions : Optimal implementation of Parallel Analysis
(PA) (Timmerman, & Lorenzo-Seva, 2011)
Dispersion matrix                                 : Polychoric Correlations
Robust analyses based on bootstrap                : None
Method for factor extraction                      : Unweighted Least Squares (ULS)
Rotation to achieve factor simplicity             : Promin (Lorenzo-Seva, 1999)
Clever rotation start                             : Weighted Varimax
Number of random starts                           : 100
```

```
Maximum number of iterations                              : 1000
Convergence value                                         : 0.00001000


----------------------------------------------------------------------

ADEQUACY OF THE CORRELATION MATRIX

Determinant of the matrix     = 0.000000000000341
Bartlett's statistic          =    990.3 (df =    820; P = 0.000037)
Kaiser-Meyer-Olkin (KMO) test = 0.19263 (inacceptable)


----------------------------------------------------------------------

EXPLAINED VARIANCE BASED ON EIGENVALUES

Variable  Eigenvalue   Proportion of    Cumulative Proportion
                       Variance         of Variance

   1        2.30859      0.05631         0.05631
   2        1.67697      0.04090
   3        1.55995      0.03805
   4        1.50178      0.03663
   5        1.46615      0.03576
   6        1.42249      0.03469
   7        1.33832      0.03264
   8        1.24762      0.03043
   9        1.23611      0.03015
  10        1.18511      0.02891
  11        1.17233      0.02859
  12        1.14650      0.02796
  13        1.08338      0.02642
```

14	1.07649	0.02626
15	1.05723	0.02579
16	1.03277	0.02519
17	1.00972	0.02463
18	1.00731	0.02457
19	1.00156	0.02443
20	1.00002	0.02439
21	1.00000	0.02439
22	0.99866	0.02436
23	0.99563	0.02428
24	0.98145	0.02394
25	0.97926	0.02388
26	0.97040	0.02367
27	0.93665	0.02285
28	0.91662	0.02236
29	0.87962	0.02145
30	0.85275	0.02080
31	0.81311	0.01983
32	0.78052	0.01904
33	0.69585	0.01697
34	0.66850	0.01630
35	0.64693	0.01578
36	0.57628	0.01406
37	0.52195	0.01273
38	0.49507	0.01207
39	0.41697	0.01017
40	0.28806	0.00703
41	0.05533	0.00135

```
PARALLEL ANALYSIS (PA) BASED ON MINIMUM RANK FACTOR ANALYSIS
(Timmerman & Lorenzo-Seva, 2011)

Implementation details:

    Correlation matrices analized:                    Polychoric correlation matrices
    Number of random correlation matrices:        500
    Method to obtain random correlation matrices: Permutation of the raw data (Buja & Eyuboglu,
1992)

Variable  Real-data       Mean of random    95 percentile of random
          % of variance   % of variance     % of variance

    1       6.0*            5.7               6.3
    2       4.4             5.1               5.6
    3       4.1             4.8               5.2
    4       4.0             4.5               4.9
    5       3.9             4.3               4.6
    6       3.5             4.1               4.4
    7       3.3             3.9               4.2
    8       3.2             3.7               4.0
    9       3.2             3.6               3.8
   10       3.1             3.4               3.6
*  Advised number of dimensions when mean is considered:           1

--------------------------------------------------------------------------


GOODNESS OF FIT STATISTICS
Goodness of Fit Index (GFI) =   0.912
```

```
UNROTATED LOADING MATRIX

Variable     F   1     Communality

V   1        -0.084       0.007
V   2        -0.020       0.000
V   3        -0.310       0.096
V   4        -0.006       0.000
V   5         0.001       0.000
V   6         0.005       0.000
V   7        -0.029       0.001
V   8        -0.001       0.000
V   9        -0.043       0.002
V  10        -0.348       0.121


----------------------------------------------------------------------------
```

MPLUS OUTPUT _ FORM 2 (FINAL ITEMS ONLY)

```
Mplus VERSION 8.6
MUTHEN & MUTHEN
06/27/2022   2:38 PM

SUMMARY OF ANALYSIS

Number of groups                                               1
Number of observations                                        50


Number of dependent variables                                 41
Number of independent variables                                0
Number of continuous latent variables                          0


Estimator                                                  WLSMV
Rotation                                                  GEOMIN
Row standardization                                  CORRELATION
Type of rotation                                         OBLIQUE
Epsilon value                                             Varies
Maximum number of iterations                                1000
Convergence criterion                                  0.500D-04
Maximum number of steepest descent iterations                 20
Maximum number of iterations for H1                         2000
Convergence criterion for H1                           0.100D-03
Optimization Specifications for the Exploratory Factor Analysis
Rotation Algorithm
  Number of random starts                                     30
  Maximum number of iterations                             10000
  Derivative convergence criterion                      0.100D-04
Link                                                      PROBIT
```

```
SUMMARY OF MODEL FIT INFORMATION


                      Number of                    Degrees of
          Model       Parameters     Chi-Square     Freedom      P-Value

          1-factor        41          827.105         779        0.1129
          2-factor        81          772.580         739        0.1900
          3-factor       120          718.475         700        0.3060
          4-factor       158          678.362         662        0.3212
          5-factor       195          636.685         625        0.3642


                                                    Degrees of
          Models Compared              Chi-Square    Freedom      P-Value

          1-factor against 2-factor      55.062        40        0.0568
          2-factor against 3-factor      53.608        39        0.0597
          3-factor against 4-factor      41.558        38        0.3185
          4-factor against 5-factor      41.670        37        0.2748



EXPLORATORY FACTOR ANALYSIS WITH 1 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters                    41

Chi-Square Test of Model Fit

        Value                           827.105*
        Degrees of Freedom                 779
        P-Value                          0.1129
```

RMSEA (Root Mean Square Error Of Approximation)

```
        Estimate                              0.035
        90 Percent C.I.                0.000   0.056
        Probability RMSEA <= .05              0.869
```

CFI/TLI

```
        CFI                                   0.614
        TLI                                   0.593
```

Chi-Square Test of Model Fit for the Baseline Model

```
        Value                              944.491
        Degrees of Freedom                     820
        P-Value                             0.0000
```

SRMR (Standardized Root Mean Square Residual)

```
        Value                                0.230
```

GEOMIN ROTATED LOADINGS (* significant at 5% level)
                    1

```
                _____
S_MODAL_9        0.379*
S_MODAL_10       0.072
S_MODAL_11       0.611*
S_MODAL_12       0.042
S_MODAL_13      -0.022
S_MODAL_14      -0.184
```

```
S_MODAL_15      0.285
S_MODAL_16      0.248
S_MODAL_17      0.472*
S_MODAL_18      0.735*
S_MODAL_51      0.594*
S_MODAL_52      0.565*
S_MODAL_55      0.525*
S_MODAL_56     -0.421*
S_MODAL_57      0.086
S_MODAL_58      0.409*
S_MODAL_60      0.236
S_MODAL_64      0.511*
S_MODAL_66      0.438*
S_MODAL_67      0.505*
S_MODAL_68     -0.154
S_MODAL_69      0.527*
S_MODAL_70     -0.341*
S_MODAL_71      0.625*
S_MODAL_72     -0.348*
S_MODAL_74      0.529*
S_MODAL_76      0.073
S_MODAL_77      0.259
S_MODAL_78      0.583*
S_MODAL_79      0.378*
S_MODAL_80      0.707*
S_MODAL_81      0.521*
S_MODAL_82      0.782*
S_MODAL_83      0.328
S_MODAL_84      0.946*
S_MODAL_85      0.216
S_MODAL_86      0.227
```

```
 S_MODAL_87         0.129
 S_MODAL_88         0.194
 S_MODAL_89        -0.255
 S_MODAL_90         0.290
```

EXPLORATORY FACTOR ANALYSIS WITH 2 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters   81

Chi-Square Test of Model Fit

         Value   772.580*
         Degrees of Freedom   739
         P-Value   0.1900

RMSEA (Root Mean Square Error Of Approximation)

         Estimate   0.030
         90 Percent C.I.   0.000   0.053
         Probability RMSEA <= .05   0.917

CFI/TLI

         CFI   0.730
         TLI   0.701

Chi-Square Test of Model Fit for the Baseline Model

         Value   944.491
         Degrees of Freedom   820
         P-Value   0.0000

SRMR (Standardized Root Mean Square Residual)

         Value   0.205

MINIMUM ROTATION FUNCTION VALUE   3.58174

              GEOMIN ROTATED LOADINGS (* significant at 5% level)
                    1   2

                 _____   _____
 S_MODAL_9   0.041   0.488*
 S_MODAL_10   -0.455*   0.589*
 S_MODAL_11   0.848*   -0.134
 S_MODAL_12   0.420*   -0.412*
 S_MODAL_13   -0.170   0.177
 S_MODAL_14   -0.163   -0.073
 S_MODAL_15   0.089   0.301
 S_MODAL_16   0.044   0.320
 S_MODAL_17   0.297   0.335
 S_MODAL_18   0.640*   0.277
 S_MODAL_51   0.486*   0.293
 S_MODAL_52   0.262   0.498*
 S_MODAL_55   0.705*   -0.045
 S_MODAL_56   -0.475*   -0.069
 S_MODAL_57   -0.354   0.512*
 S_MODAL_58   0.351   0.185

```
S_MODAL_60          0.573*         -0.304*
S_MODAL_64          0.317           0.357
S_MODAL_66         -0.048           0.708*
S_MODAL_67          0.263           0.434*
S_MODAL_68         -0.381*          0.217
S_MODAL_69          0.000           0.814*
S_MODAL_70         -0.279          -0.168
S_MODAL_71          0.759*         -0.026
S_MODAL_72         -0.421*         -0.004
S_MODAL_74          0.620*          0.017
S_MODAL_76         -0.225           0.360*
S_MODAL_77          0.232           0.113
S_MODAL_78          0.439*          0.331
S_MODAL_79          0.101           0.435*
S_MODAL_80          0.240           0.787*
S_MODAL_81          0.228           0.491*
S_MODAL_82          0.600*          0.431
S_MODAL_83         -0.185           0.665*
S_MODAL_84          0.835*          0.377
S_MODAL_85         -0.008           0.329*
S_MODAL_86         -0.156           0.502*
S_MODAL_87          0.282          -0.142
S_MODAL_88          0.033           0.241
S_MODAL_89         -0.306          -0.015
S_MODAL_90          0.320           0.045
```

```
EXPLORATORY FACTOR ANALYSIS WITH 3 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters                        120

Chi-Square Test of Model Fit

        Value                            718.475*
        Degrees of Freedom                   700
        P-Value                           0.3060

RMSEA (Root Mean Square Error Of Approximation)

        Estimate                           0.023
        90 Percent C.I.              0.000   0.050
        Probability RMSEA <= .05          0.954
CFI/TLI

        CFI                                0.852
        TLI                                0.826

Chi-Square Test of Model Fit for the Baseline Model

        Value                            944.491
        Degrees of Freedom                   820
        P-Value                           0.0000

SRMR (Standardized Root Mean Square Residual)

        Value                              0.185
```

```
MINIMUM ROTATION FUNCTION VALUE        2.33126

            GEOMIN ROTATED LOADINGS (* significant at 5% level)
                  1                2                3


                _____        _____        _____
 S_MODAL_9        0.347*           0.535*           0.085
 S_MODAL_10       0.011            0.786*          -0.117
 S_MODAL_11       0.925*          -0.126           0.000
 S_MODAL_12       0.496*          -0.308          -0.273
 S_MODAL_13      -0.417*          -0.022           0.321
 S_MODAL_14      -0.215          -0.053          -0.039
 S_MODAL_15       0.467*           0.437*          -0.119
 S_MODAL_16      -0.338          -0.107           0.584*
 S_MODAL_17       0.307            0.192           0.302
 S_MODAL_18       0.593*           0.103           0.361
 S_MODAL_51       0.347            0.015           0.447*
 S_MODAL_52       0.414*           0.412*           0.279
 S_MODAL_55       0.732*          -0.107           0.099
 S_MODAL_56      -0.512*           0.009          -0.126
 S_MODAL_57      -0.304           0.455*           0.245
 S_MODAL_58      -0.146          -0.293           0.659*
 S_MODAL_60       0.204          -0.557*           0.240
 S_MODAL_64       0.701*           0.452*          -0.059
 S_MODAL_66      -0.049           0.484*           0.502
 S_MODAL_67      -0.112          -0.011           0.713*
 S_MODAL_68      -0.090           0.417*          -0.192
 S_MODAL_69       0.037            0.583*           0.559*
 S_MODAL_70      -0.053           0.097          -0.380*
 S_MODAL_71       0.457          -0.373           0.429*
 S_MODAL_72      -0.079           0.301          -0.402*
 S_MODAL_74       0.469          -0.187           0.287
```

```
S_MODAL_76        -0.625*          0.028          0.560*
S_MODAL_77        -0.081         -0.191          0.407*
S_MODAL_78         0.390          0.117          0.380
S_MODAL_79         0.040          0.221          0.391*
S_MODAL_80        -0.007          0.303          0.860*
S_MODAL_81         0.265          0.311          0.373
S_MODAL_82         0.252         -0.055          0.754*
S_MODAL_83         0.281          0.815*         0.016
S_MODAL_84         0.661*         0.003          0.617*
S_MODAL_85        -0.207          0.100          0.414*
S_MODAL_86        -0.040          0.449*         0.225
S_MODAL_87         0.209         -0.205          0.026
S_MODAL_88        -0.020          0.128          0.230
S_MODAL_89        -0.185          0.143         -0.190
S_MODAL_90         0.040         -0.304          0.364
```

EXPLORATORY FACTOR ANALYSIS WITH 4 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters   158

Chi-Square Test of Model Fit

        Value   678.362*
        Degrees of Freedom   662
        P-Value   0.3212

```
RMSEA (Root Mean Square Error Of Approximation)

        Estimate                              0.022
        90 Percent C.I.                 0.000  0.050
        Probability RMSEA <= .05              0.952

CFI/TLI

        CFI                                   0.869
        TLI                                   0.837

Chi-Square Test of Model Fit for the Baseline Model

        Value                               944.491
        Degrees of Freedom                      820
        P-Value                              0.0000

SRMR (Standardized Root Mean Square Residual)

        Value                                 0.171

MINIMUM ROTATION FUNCTION VALUE        2.63232


            GEOMIN ROTATED LOADINGS (* significant at 5% level)
                   1            2            3            4

                _____     _____     _____     _____
 S_MODAL_9        0.033        0.282        0.551*      -0.031
 S_MODAL_10       0.032       -0.125        0.774*      -0.199
 S_MODAL_11      -0.059        0.967*      -0.094       -0.168
 S_MODAL_12      -0.407*       0.451*      -0.241       -0.110
 S_MODAL_13       0.754*      -0.171       -0.210       -0.109
```

S_MODAL_14	0.158	-0.149	-0.125	-0.123
S_MODAL_15	-0.004	0.414	0.440*	-0.275
S_MODAL_16	-0.027	-0.388	0.024	0.777*
S_MODAL_17	0.214	0.370	0.189	0.090
S_MODAL_18	0.325	0.688*	0.090	0.021
S_MODAL_51	0.066	0.378	0.088	0.389
S_MODAL_52	-0.077	0.324	0.456*	0.282
S_MODAL_55	-0.186	0.742*	-0.029	0.085
S_MODAL_56	0.403*	-0.394	-0.103	-0.356
S_MODAL_57	0.500*	-0.240	0.373	-0.039
S_MODAL_58	0.234	-0.038	-0.255	0.638*
S_MODAL_60	0.299	0.431	-0.572*	-0.020
S_MODAL_64	0.030	0.668*	0.468*	-0.287
S_MODAL_66	0.528*	0.032	0.437	0.151
S_MODAL_67	0.595*	0.081	-0.046	0.372*
S_MODAL_68	-0.133	-0.235	0.423*	-0.093
S_MODAL_69	0.080	-0.065	0.638*	0.594*
S_MODAL_70	-0.103	-0.108	0.058	-0.350
S_MODAL_71	0.010	0.560*	-0.295	0.386
S_MODAL_72	-0.567*	-0.343	0.385*	0.036
S_MODAL_74	-0.022	0.523*	-0.141	0.253
S_MODAL_76	0.427	-0.543*	-0.013	0.475
S_MODAL_77	0.063	-0.048	-0.141	0.463*
S_MODAL_78	0.001	0.392	0.189	0.345
S_MODAL_79	-0.111	-0.072	0.335	0.552*
S_MODAL_80	0.525	0.106	0.306	0.570
S_MODAL_81	-0.141	0.159	0.434*	0.481*
S_MODAL_82	0.278	0.367	-0.004	0.604*
S_MODAL_83	0.035	0.158	0.819*	-0.102
S_MODAL_84	0.464	0.866*	0.011	0.158
S_MODAL_85	0.494*	-0.069	0.035	0.133

```
S_MODAL_86        0.281         -0.026          0.412*          0.028
S_MODAL_87       -0.310          0.148         -0.102          0.226
S_MODAL_88        0.106         -0.017          0.137          0.193
S_MODAL_89        0.285         -0.107          0.036         -0.418
S_MODAL_90       -0.456         -0.023         -0.108          0.727*
```

EXPLORATORY FACTOR ANALYSIS WITH 5 FACTOR(S):

MODEL FIT INFORMATION

Number of Free Parameters   195

Chi-Square Test of Model Fit

        Value   636.685*
        Degrees of Freedom   625
        P-Value   0.3642

RMSEA (Root Mean Square Error Of Approximation)

        Estimate   0.019
        90 Percent C.I.   0.000   0.049
        Probability RMSEA <= .05   0.957
CFI/TLI

        CFI   0.906
        TLI   0.877

Chi-Square Test of Model Fit for the Baseline Model

          Value   944.491
          Degrees of Freedom   820
          P-Value   0.0000

SRMR (Standardized Root Mean Square Residual)

          Value   0.154

MINIMUM ROTATION FUNCTION VALUE   2.23609

              GEOMIN ROTATED LOADINGS (* significant at 5% level)
                   1   2   3   4   5

                _____   _____   _____   _____   _____
  S_MODAL_9   0.468*   0.016   0.432*   -0.005   0.042
  S_MODAL_10   0.869*   0.313   -0.071   0.008   0.122
  S_MODAL_11   -0.004   -0.632*   0.584   0.360   -0.049
  S_MODAL_12   -0.081   -0.571*   0.038   0.121   0.232
  S_MODAL_13   -0.283   0.409   0.082   0.116   -0.724*
  S_MODAL_14   -0.223   -0.054   0.103   -0.265   -0.381*
  S_MODAL_15   0.695*   -0.066   0.026   0.438   0.146
  S_MODAL_16   -0.222   0.608*   -0.095   0.002   0.570*
  S_MODAL_17   0.115   0.056   0.434*   0.198   -0.046
  S_MODAL_18   0.012   -0.118   0.673*   0.327   -0.226
  S_MODAL_51   0.123   0.194   0.141   0.552*   0.473*
  S_MODAL_52   0.086   -0.137   0.766*   -0.263   0.054
  S_MODAL_55   -0.038   -0.465*   0.525   0.223   0.185
  S_MODAL_56   0.044   0.285   -0.396   0.029   -0.482*
  S_MODAL_57   0.254   0.500*   0.121   -0.016   -0.354
  S_MODAL_58   -0.581*   0.319   0.309   -0.001   0.097

S_MODAL_60	-0.277	-0.016	-0.141	0.721*	-0.019
S_MODAL_64	0.589*	-0.242	0.466	0.288	-0.013
S_MODAL_66	0.469	0.701*	0.023	0.553	0.029
S_MODAL_67	-0.030	0.686*	0.000	0.646*	0.064
S_MODAL_68	0.369*	0.076	-0.042	-0.266	0.045
S_MODAL_69	-0.024	0.312	0.785	-0.485	0.036
S_MODAL_70	0.314	-0.055	-0.409*	0.103	-0.009
S_MODAL_71	-0.319	-0.124	0.335	0.434	0.293
S_MODAL_72	0.131	-0.228	0.082	-0.768*	0.209
S_MODAL_74	-0.318	-0.292	0.559	0.077	0.049
S_MODAL_76	-0.322	0.699*	-0.005	-0.135	-0.112
S_MODAL_77	-0.288	0.257	0.076	0.089	0.251
S_MODAL_78	-0.007	-0.010	0.548*	0.075	0.216
S_MODAL_79	0.152	0.368	0.136	0.022	0.556*
S_MODAL_80	-0.050	0.593*	0.635*	0.117	-0.007
S_MODAL_81	0.160	0.146	0.485*	-0.097	0.421*
S_MODAL_82	-0.412*	0.130	0.829*	-0.004	0.006
S_MODAL_83	0.636*	0.031	0.575*	-0.265	-0.104
S_MODAL_84	0.032	0.029	0.671	0.699*	-0.057
S_MODAL_85	-0.014	0.455*	0.058	0.255	-0.192
S_MODAL_86	0.223	0.251	0.366*	-0.132	-0.227
S_MODAL_87	-0.106	-0.165	0.024	0.029	0.387*
S_MODAL_88	-0.113	0.088	0.373*	-0.238	-0.095
S_MODAL_89	0.118	0.014	-0.068	-0.055	-0.514*
S_MODAL_90	-0.278	0.054	0.039	-0.023	0.767*

List of abbreviations

AFC	Agenda for Change
AHP	Allied Health Professional
AKT	Applied Knowledge Test
BAME	Black, Asian, and Minority Ethnic
BFI-2-S	Big Five Inventory–2 Short Form
COVID-19	Coronavirus disease
CQC	The Care Quality Commission
CSA	Clinical Skills Assessment
ERS	Extreme Response Style
HCPC	The Health and Care Professions Council, formerly
HEE	Health Education England
HYMS	Hull York Medical School
IACS	Independent Assessment of Clinical Skills Programme
ITP	Implicit Trait Policy
MHS	Mental Health Services
MSC	Medical Schools Council
MMI	Multiple Mini Interview
NHS	National Health Service

NHSFT	National Health Service Foundation Trust
NMC	The Nursing and Midwifery Council
OT	Occupational Therapist
PhD	Doctor of Philosophy
PICOS	Population, Intervention, Comparison, Outcome, Study Design
PIS	Participant Information Sheet
PROSPERO	International prospective register of systematic reviews
R&D	Research and Development
REC	Research Ethics Committee
SJT	Situational Judgement Test
SME	Subject Matter Expert
SPIDER	Sample, Phenomenon of Interest, Design, Evaluation and Research type
TEWV NHSFT	Tees, Esk and Wear Valleys NHS Foundation Trust
UCAT	University Clinical Aptitude Test (formerly UKCAT)
UK	United Kingdom
UKCAT	UK Clinical Aptitude Test (currently UCAT)
USA	United States of America