

Evaluating diversity education in the National Health Service: the development and piloting of a new situational judgement test

Riya Elizabeth George (✉ r.george@qmul.ac.uk)

Queen Mary University of London Barts and The London School of Medicine and Dentistry

<https://orcid.org/0000-0001-5385-8391>

Pat Dugard

University of Dundee

Nisha Dogra

University of Leicester

Research article

Keywords: Diversity education, cultural competence, National Health Service, healthcare, evaluation and situational judgement test

Posted Date: December 11th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-124415/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

As a subject, diversity education poses inherent challenges in its delivery and evaluation. Current evaluation tools are largely inadequate and often rely on participants' self-reported assessments of their knowledge, attitudes and skills. These tools rarely ask participants to demonstrate these attributes. This study describes the development and piloting of a new situational judgement test (SJT) for use in evaluating diversity education in the National Health Service (NHS) and healthcare educational institutions.

Methods

We started by adapting scenarios developed from a series of participatory workshops. Scenario based questions, response items, response formats and scoring methods were devised, tested and refined through an iterative process of three stages of piloting with a total of 198 new participants, either attending NHS diversity trainings or as part of a control group.

Results

SJTs are not generally used to evaluate diversity education, but our findings suggest they have considerable untapped potential. The findings from this study SJT shows how an individual can demonstrate differing attitudes to a range of diversity factors.

Conclusion

This tool has important applications beyond research. An unexpected finding was the enthusiasm of the participants for using the SJT as an educational resource to foster clinically relevant discussions on diversity issues. The scenarios can be tailored to different contexts, further developed and refined for future use. This preliminary attempt at devising and piloting a SJT is a key innovative step in the development of improved evaluation of diversity education.

Background

The need to provide diversity education is driven by the changing demographic realities of healthcare populations internationally. A multitude of policies demonstrate an expectation that healthcare professionals can work effectively with an increasingly diverse population (General Medical Council, 2009; Association of American Medical Colleges, 2019; Nursing and Midwifery Council, 2010), and have suggested diversity education as part of a strategy to reduce racial/ethnic disparities in healthcare. In the National Health Service (NHS) diversity education is mandatory for all healthcare professionals

(Department of Health, 2012), however wide variability exists in how this type of teaching is developed, delivered and evaluated across healthcare settings (George et al, 2015). Diversity education, often used interchangeably with the term cultural competence training, has become a widespread practice in an active attempt to improve the knowledge, awareness, skills and attitudes of healthcare professionals in serving culturally diverse populations (Dogra et al, 2015).

Despite the frequent use of and progression towards the term 'diversity', particularly in the United Kingdom, much of the literature on the evaluation of these types of teaching uses the term cultural competence (Price et al, 2005; Shen et al, 2015; Lotin et al, 2013). Several definitions of cultural competence exist, the most commonly cited being *"a set of congruent behaviours, attitudes and policies that come together in a system, agency or among professionals and enables that system, agency or those professionals to work effectively in cross-cultural situations"* (Cross et al, 1989). However, cultural competence as a concept or strategy is not yet fully developed and often vaguely defined, poorly understood, and used synonymously with a range of terms (George et al, 2019). The variability in how this subject is understood, taught and evaluated is attributable to the lack of conceptual clarity of the terms diversity education and cultural competence (Dogra et al, 2016).

Current tools or measures to evaluate diversity education are largely inadequate and often rely on participants' self-reported assessments of their knowledge, attitudes and skills (Shen et al, 2015; Lotin et al, 2013). These tools rarely ask participants to demonstrate these attributes themselves. Furthermore, most tools designed to evaluate diversity education, conflate the terms diversity and cultural competence and often use these terms synonymously without clarity in how they are distinct from each other.

This study aims to develop and pilot a Situational Judgement Test (SJT) that can be used to evaluate diversity education in the NHS and in health educational institutions. We begin by firstly clarifying the definitions of diversity and cultural competence to provide conceptual clarity and discuss the critiques of each of the constructs and its implications on evaluation. We provide an overview of the limitations of current evaluative methods and pertinent findings from a series of participatory workshops are also shared as these led to the decision to develop a SJT.

Defining cultural competence and diversity

Since its introduction in the 1980s, a range of theoretical frameworks and models for cultural competence have flourished, and the conceptualisation of the subject has evolved through several phases from a reductionist approach focusing on culture knowledge, towards a more dynamic approach underpinned by critical social theory (Kumagai et al, 2009; Dogra et al, 2015). Traditionally, cultural competence frameworks or models emphasized the importance of developing cultural knowledge or 'cultural expertise' as a way of becoming proficient in serving culturally diverse populations. These notions of cultural competence were reflected in definitions of culture which favoured group based distinctions, categorising sets of individuals based on factors such as religion, race or ethnicity (Bhui et al, 2012). Over the years, cultural competence models have increasingly attracted negative criticisms of

their assumptions that a healthcare professional can learn or know enough about other cultural groups and develop a full understanding of cultures different from their own. Criticisms were made of the inherent disregard for the complexity of culture and the assumption that one could become culturally competent by simply learning general facts about certain cultural groups (Kai et al, 2007). Recent research demonstrates a theoretical progression away from knowledge based cultural competence models to process oriented models, where understanding oneself takes precedence over gaining knowledge and expertise about others (Bintley and George, 2019).

Throughout the decades, the theoretical frameworks for cultural competence show a clear convergence to developing one's self-awareness and reflection to facilitate a better understanding of the complexity of cultural differences in others. Diversity became a term that better described the complexity of cultural differences, however achieving consensus on the definition has proved challenging. In its broadest sense, any individual difference can be regarded as diversity (Dogra et al, 2015). The manner and variation upon which shared meanings of culture are internalised, understood and practised in individuals gives rise to diversity in populations and creates differences in the understandings of health and illness.

Various theoretical frameworks to achieve cultural competence or guide diversity education have been applied in healthcare education, but the literature infrequently refers to a clear theoretical position. Recent reviews illustrate the lack of conceptual clarity and rigour in identifying a sound, evidence-based theoretical framework upon which to base cultural competence education (Reitmanova, 2011; Hobgood et al, 2006; Betancourt et al, 2003). Different educational philosophies and theoretical frameworks view culture and diversity very differently, leading to programmes with differing intentions (Dogra et al, 2015). The distinctions between these theoretical frameworks and models is somewhat blurred, resulting in terms such as 'cultural competence' and 'cultural sensitivity' being used interchangeably and synonymously. Despite multiple recent efforts to establish common competencies and standards for cultural competence/diversity education, the overall learning objectives are still unclear.

George et al (2015) conducted a critical interpretive review of the literature exploring cultural competency trainings in UK healthcare settings; concluding that cultural competence is '*underdeveloped, undertheorised and piecemeal in nature*' and is predominately empirically rather than theoretically driven. This finding concurs with the quality of literature on cultural competence in the United States, Canada and Europe (Dogra et al, 2010, Betancourt, 2005; Sorenson et al, 2019), which frequently refers to the lack of conceptual and theoretical clarity in the field. Recent reviews have tended to focus on synthesizing interventions to improve cultural competence (Truong et al, 2014; Alizadeh et al, 2016; Price et al, 2008) or evaluative/ outcome measures used to determine the effectiveness of cultural competence programmes (Shen et al, 2015; Renzaho et al, 2013; Kumas-Tan, 2007; Anderson et al, 2003; Bhui et al, 2007), generally reporting that the effectiveness of these types of teaching is inconclusive and in some cases not even measured.

Limitations in evaluation methods for cultural competence

Given that most of the literature on the evaluation is on the evaluation of cultural competence training, this term is used for this subsection of the paper. Cultural competence education poses inherent challenges in its evaluation and measurement largely because its meaning is so varied, broad, nuanced and complex (Suzuki et al, 2001). Many cultural competence evaluation tools and measures rely heavily on the three domains of knowledge, attitude and skills, with a large emphasis on cultural knowledge. However, there is still much dispute as to whether these domains can fully and reliably capture the complexity of diversity and cultural issues (Malet et al, 2013).

The widespread reliance on self-reported measures (generally questionnaires) to evaluate cultural competence raises issues around social desirability (the tendency to answer questions in a manner that is viewed as favourable by others). Multiple studies have shown that respondents may over or underestimate their cultural competence and may be inclined to report what they anticipate to be their cultural competence rather than their actual behaviours and attitudes (Lotin et al, 2013; Price et al, 2015; Shen, 2016). Furthermore, it may be challenging for participants to assess their cultural competence without a clear understanding of how these terms are defined and understood in practice. Many existing measures oversimplify the concepts of culture and cultural competence as a means to conform to the narrow domains of measurement scales (Burcham, 2005).

Another challenge that is frequently raised is relevance and usability; existing measures of self-reported psychometric questionnaires can be lengthy and not entirely relevant to those of different healthcare professions (Suzuki et al, 2001). Many of the evaluation tools and measures for cultural competence were developed in the context of counselling psychology and used internationally, and questions of transferability of these tools in the UK context still need further exploration. Additionally, the NHS now requires all healthcare professionals to undergo diversity/cultural competence training, therefore it is important to ensure that evaluation instruments are broad enough to apply to all healthcare professionals.

Kumas-Tan (2007) conducted a critical analysis of the ten most widely used measures for evaluating cultural competence teaching. Kumas-Tan (2007) defined six underlying assumptions about culture and cultural competence that these measures exemplified. These were 1.) Culture is a matter of ethnicity and race, 2.) Culture is possessed by the 'other' and the 'other' has the problem, 3.) The problem of cultural incompetence lies in practitioners' lack of familiarity with the 'other', 4.) The problem of cultural incompetence lies in practitioners' discriminatory attitudes towards the 'other', 5.) Cross-cultural healthcare is about White practitioners working with patients from ethnic and racial minority groups and 6.) Cultural competence is about being confident in oneself and comfortable with others. These underlying assumptions reflect the traditional notion of cultural competence which assumes it can be attained by acquiring specialised knowledge of other cultures. Many of the measures disproportionately emphasise issues of race and ethnicity in comparison to other individual differences and resonate with group based definitions of culture that tend to homogenise groups of individuals. Despite the array of other patient differences that can affect healthcare interactions, race and ethnicity remain the most frequently discussed differences to the near exclusion of other individual differences. This may reflect a

larger conceptual problem about defining cultural competence and its parameters; echoing current debates about how 'diversity' is a better term to capture any individual difference. Given that institutional and healthcare educational policies are emphasising a wider range of desired attributes, aptitudes and skills beyond the typical knowledge and attitudes in relation to cultural competence/diversity, further research is needed to explore whether these different constructs are best measured separately, whether multiple methods should be used or whether new instruments can expand the conceptualisation of cultural competence to account for diversity.

Towards a new evaluation tool for diversity education: a situational judgement test

Findings from participatory workshops

This study was the final phase of a PhD research programme conducted by the first author (RG) which explored how to better teach and evaluate diversity education in the NHS. The former phases of the PhD involved conducting a series of participatory workshops with groups of key stakeholders (a total of 94); these were patients with mental illness (George et al, 2019), healthcare professionals and medical educators. The workshops explored the expectations of key stakeholders on healthcare professionals deemed culturally competent and what an evaluation tool for diversity education should be seeking to measure.

Overall, the findings from the participatory workshops endorsed the ideas and concepts of the relationship-centred care (RCC) (Tresolini, 1994) model and in essence reproduced the relationships of the practitioner – patient and the practitioner – practitioner, but also added something new, namely the 'practitioner – self relationship', which is central to our effort to devise a new method of evaluating diversity education. Collectively the findings revealed that diversity education should focus on the nuances and dynamics of clinical relationships, where the influence of both the patient and the practitioner are acknowledged and explored. The relationship considered the most important to examine with respect to diversity education was the '*practitioner-self relationship*'. This requires health professionals to explore, unpack and reflect upon the meaning of diversity on an individual level and in relation to colleagues, peers and patients, to facilitate an appreciation of and value for diversity in others. A reconstructed RCC model incorporating the four relationships; practitioner-self, practitioner-patient, practitioner-practitioner and practitioner-organisation was developed.

The participatory workshop findings showed that an evaluation tool for diversity education should be seeking to measure health professionals' demonstration of the attributes (i.e. an increase in knowledge or change in attitudes) developed after receiving diversity education. Participants desired an evaluation tool that made health professionals *think* about diversity issues from multiple perspectives, that was contextualised, clinically relevant, meaningful to every day practice and simple to administer. These considerations, alongside the attributes included in the reconstructed RCC model strongly suggested that a situational judgement test (SJT) may be the most appropriate evaluation tool.

This study aims to develop and pilot a SJT that can be used to evaluate diversity education in the NHS and healthcare educational institutions. First we describe SJTs, their current uses and applicability to evaluating diversity education. Ethical approval for this study was granted by the University of Leicester and informed consent was obtained from all individual participants included in this study.

Defining a situational judgement test

Situational judgement tests (SJTs) are designed to evaluate an individual's reactions to or judgements of several hypothetical scenario based questions that reflect situations they are likely to encounter in clinical practice (Patterson, 2016). These scenarios are developed based on a rigorous and detailed analysis of pertinent attributes and traits of the desired role and are constructed collaboratively with a range of stakeholder experts. This robust developmental process ensures the test is able to evaluate the key attributes that are associated with competent performance.

Situational judgement tests are classed as a *measurement methodology* (Chan et al, 1998) as opposed to a single style of evaluation or assessment. This is due to the variability in scenario content, response formats and approaches to scoring. Typically, candidates are presented with a likely scenario which is accompanied by a series of possible responses (known as 'items'), and are asked to identify the appropriateness or effectiveness of these responses. The response items are developed in the same rigorous fashion as the scenarios and a predefined scoring key is agreed by stakeholder experts. Several scenarios are likely to be included in a SJT as it allows broad and complex constructs to be measured efficiently.

Best practice for developing a situational judgement test

SJTs which are designed for selection, assessment, evaluation or developmental purposes should follow best practice outlined in the Patterson et al (2016) review to ensure psychometric quality. The recommended steps as they apply to this study are outlined below, followed by a full description of the developmental process:

1. Role analysis based on qualitative research to ascertain the key attributes and capabilities associated with competent performance.
2. Test specification in collaboration with key stakeholders: this means first identifying a set of critical incidents (salient or challenging situations that reflect every day scenarios that are likely to arise in the work of those in the targeted role). These critical incidents can then be used to draft appropriate scenarios and response items, followed by a rigorous review and editing process.
3. Piloting and further review of scenarios, response items, response formats and scoring to produce a final version.

Methods

Development of the situational judgement test

Role analysis

The first stage in designing a SJT is to conduct a role analysis. Typically, a role analysis includes conducting interviews with key stakeholders, however in the context of this study the qualitative findings from the participatory workshops provided the ideal platform of essential information on what an evaluation tool for diversity education should be seeking to measure. It outlined the desired attributes and skills of a culturally competent healthcare professional and clarified the meaning of 'cultural competence' and its constituent components.

Test specification

The next step in the process involves identifying and collecting several 'critical incidents'. The participants in all of the participatory workshops illustrated their ideas and concerns with stories of situations where diversity issues arose and they found themselves at a loss as to the best response. These difficult situations covered a wide range of diversity issues: religion, gender, and disability as well as race and ethnicity. A total of 90 critical incidents were identified and developed into scenario based questions and responses that conformed to the dimensions of the reconstructed relationship centred care model. Response items were developed in discussion with members of the research team (PD and ND) as well as participants.

Scenario and response item development

SJT scenarios and response items are best developed in collaboration with key stakeholders i.e. those who have an expertise in diversity issues and are familiar with the target role (Patterson et al, 2016). This process of development is essential to ensure SJT scenarios and response items are realistic, appropriate and plausible.

In this study, the critical incidents which were the starting point for developing the SJT were all provided by key stakeholders who took part in the participatory workshops and all were derived from anecdotes from their own personal and clinical experiences.

In collaboration with different NHS diversity leads, a list of three to six response items for each scenario were constructed. Each response item needs to depict a realistic response an individual might make in that scenario, with the list of response items reflecting a mixture of good and poor actions. How the response items are framed, the language used and connotations associated with each response were considered throughout the piloting stages.

Response format and scoring

Response instructions, format and scoring take various forms depending on how the SJT is being used. Response instructions are typically grouped in two categories: 1.) Knowledge based (i.e. what is the best option) or 2.) Behavioural tendency (what would you be most likely to do). Within these two categories, a

variety of response formats can be used such as *ranking* all the response items independently or ranking possible actions in order. An alternative format is that of *multiple choice*, where candidates are asked to choose the best/worst response items. Other researchers have opted for *single response* formats where only one response is chosen (Motowidlo et al, 2009). The type of response format depends on the role analysis and test specification and the context or level in the education the SJT is targeting. The formats piloted in this study included ranking in order (from 1 for the best option), choosing the two best and choosing the best option. For choosing the best option or the best two, it is necessary to devise a scoring key which assigns a numerical score for each response item. This scoring key is defined and agreed upon by an in-depth review process collaboratively with key stakeholders or in-depth interviews with stakeholder experts. Our scoring key was developed in collaboration with the same NHS diversity leads who contributed to the development of response items.

Piloting and review

Three stages of piloting and review were undertaken and these are detailed below.

Stage I pilots

An overview of the different pilot stages is shown in Figure 1.

Stage I consisted of three small pilot sessions, each with a key stakeholder group (medical educators in the field of diversity (n=7), NHS diversity leads (n=10) and patients with mental illness (n=10), and a final fourth pilot session (n = 18) with NHS healthcare professionals. Participants were asked to provide feedback on the design, authenticity (i.e. how realistic are the scenarios), clarity, relevance and fairness (content validity) of scenarios and to identify appropriate responses. They could also provide other feedback if they wished.

The initial three pilot sessions lasted an hour each. Participants worked in pairs to complete four example scenarios each with a list of responses to rank in order of appropriateness. Each pair had a set of four unique scenarios, so that feedback was available for 20 scenarios from each pilot. Participants were encouraged to discuss their scenarios before responding. The response sheets asked respondents: a) to explain their reasoning for their responses and b) suggestions for improvement for this scenario and possible response items. An example of the feedback received is shown in Table 1. The feedback from the first three pilot sessions led to the selection of 20 scenarios which were then further refined

For the final fourth pilot session we developed a pre and post SJT using some of the 20 scenarios, ten for each of the pre and post SJT versions. The post SJT version used five of the pre SJT version scenarios plus five new scenarios. This fourth pilot session was necessary to explore the challenges of response formats and scoring and also to assess participants' reactions and perspectives on the SJT in the context of a typical diversity education session in the NHS, with the sample being representative of NHS staff.

Stage II pilots

Stage II pilots involved conducting two pilot sessions (n = 23 and n = 80) where a trial pre and post SJT was distributed; each pilot session was with NHS healthcare professionals attending a diversity education. The NHS trust sites collaborating in this research provided written confirmation of their consent to participate in the study. Prospective attendees were informed beforehand about the research study and provided with additional information if they wished to contact the research team. Key information about the research study was re-iterated at the beginning of the diversity education session and participants were asked to read and sign a consent form, which all attendees completed. As before they willingly agreed to complete a trial pre and post SJT to help with the research. For the first stage II pilot session we used six scenarios each with four possible responses. Participants had to choose the best response. The same six scenarios were used pre and post training, but for the post version they were randomly reordered. The pre SJT version can be seen in Figure 2.

For the second stage II pilot session, two of the scenarios were discarded (1 and 3 shown in Figure 2), the decision to reduce the length is discussed in the results section below. The remaining four scenarios represented the relationships, practitioner-self, practitioner-practitioner and practitioner-organisation in the reconstructed RCC model and covered diversity issues such as religion, disability, language and sexual orientation. These diversity issues in particular were covered in the diversity education session NHS healthcare professionals were attending.

Stage III pilots

The participants for this stage were from the finance and investment banking sector and mathematics (n = 50). They completed the post SJT with four scenarios that was used for the earlier Stage II pilot session. These two groups comprised a single control group who had confirmed they had not received NHS diversity education.

The wording of the scenarios, the lists of response items and their scores, and the response format were repeatedly revised through discussion and review of the pilots. Table 2 shows how one example scenario, the response items and format were developed over the pilot stages.

Results

The findings of all the pilot stages are discussed below and illustrate the necessity for an iterative and rigorous piloting process in the development of a SJT to aid and refine scenario and item development, response format and scoring methods.

Stage I Pilots

Comments from participants

Participants in these pilot sessions provided valuable feedback on many aspects of the draft SJT, from details of the scenarios and response formats to comments on time constraints and suggestions about

further uses of the SJT. Overall, participants were very positive about the scenarios, which they found interesting, thought provoking and relevant to their practice. Details are given below.

- **Scenario refinement:** some participants thought the scenarios should clearly state the role (e.g. as a manager or nurse) the participant plays in each scenario. However others suggested that scenarios that allowed the respondents to answer from the perspective of their current role would elicit more authentic and plausible responses. Some participants suggested a blank response item where participants could provide a qualitative personal response.
- **Specificity of scenarios:** mixed responses were reported among the pilot groups as to whether the scenarios should be broad or specific in nature. Patients emphasised the need for broad scenarios in order for them to be applicable to a broad range of healthcare professionals that attended diversity education teaching. Conversely medical educators and NHS healthcare professionals suggested more specific details to be included in the scenarios for example, patient information and details of the context.
- **Response category:** all stakeholders proposed that items should be behavioural (what would you be likely to do) responses, as opposed to knowledge based (what is the best option).
- **Response format:** some participants commented that it was confusing to be asked to respond in different ways to different scenarios, therefore a single response format needed to be established for future versions of the SJT.
- **Time constraints:** all the pilot groups said more time was needed to do justice to the task. They suggested that some scenarios could be made more concise and also that fewer scenarios should be used (perhaps four to six).
- **Tailored versions of the SJT:** since some diversity education sessions last for 3 hours or more, others last only an hour, a SJT with a larger number of scenarios (perhaps eight to ten) could be used for longer teaching sessions and a smaller number (perhaps four to six) could be used for shorter teaching sessions.
- **Format versatility:** all four groups raised the different ways the SJT could be presented and used. Medical educators as well as NHS healthcare professionals suggested that the SJT can be used simultaneously as an educational resource and as an evaluation tool. The patient pilot group suggested that the SJT could be converted into a video based format and used as part of the e-learning package for diversity education.

Analysis of responses

After each of the three Stage I pilot sessions the frequencies with which each item was chosen as most or least appropriate were listed. In some cases, participants were almost unanimous in their choice. An example being in Scenario 1 in Table 3 (the Muslim doctor who says bare below the elbows is against her religion), almost everyone chose option B, 'discuss the issue with the rest of the team and ask for advice from a senior member of the team'. Such cases suggest that the response was too obvious. On closer examination, selecting option B excused the individual from making a decision. All scenarios and their

responses were collated and reviewed to assess whether certain items were almost unanimously chosen or rejected. Response items were reviewed and refined to assess whether they could be more discriminating. The response items needed to cover a range of plausible, mutually exclusive responses which recognise the subtlety of a situation such as the 'bare below the elbows' example, where a case could be made for either flexibility or insistence that professional expectations must be upheld.

Ranking (1 to 4 or 5 for best to worst) was used for some scenarios and others 'choose the best two' response formats in stage I pilot. Parallel to the participants' comments that different response formats were confusing, analysis of responses raised several other concerns:

- **Tied rankings:** many participants opted for the use of tied ranking (identical responses for different items). These required calculations to be performed (find the average rank) to make their responses comparable with others. Typically scoring should be based on the difference from the optimal ranking but with tied responses it is unclear how this should be calculated. Correlation between the participant's and the optimal ranking may be the best score, but this involves calculations which would be very time consuming and perhaps error prone in routine use.
- **Omitted items:** some participants omitted options from their ranking, and it is unclear how such responses should be scored. After the stage I pilots we did not use the ranking format.
- **Consistency in the number of response items:** As stage I pilot sessions involved the initial stages of item development, the scenarios requesting ranking all had different numbers of possible items, so the same weights could not be applied when totalling the scores. The range of possible values/scores varied in accordance with the number of possible items for each scenario.
- **Multiple choice scoring:** scenarios that required participants to choose the two most appropriate responses raised concerns. Some participants ignored the instructions, opting for ranking the items instead. Other participants responded by choosing only one option. In these cases it is not at all clear how responses should be scored.

In the light of the results above, the 'best single response' was deemed the most appropriate and a total of 4 response items per scenario were used. This offered some uniformity to the task presented to the participants and also allowed a uniform scoring system to be used. If the SJT is to be suitable for routine use, a simple scoring system is most suitable.

Stage II Pilots

Comments from participants

The feedback for the stage II pilot sessions mirrored that of stage I. As the participants in first group in this stage reported needing more time, for the second group we reduced the test to four scenarios and this gave sufficient time to complete the task.

Analysis of responses: session 1

Our scoring key assigned 1 to the best response item and higher scores to worse response items. Each participant had a pre and post score for each scenario and a total pre and post score was obtained by adding scores across the scenarios. Lower scores and totals indicated responses showing more awareness of diversity issues.

The correlation between the pre and post totals is quite large (0.41), however not so large as to indicate that there were no changes in participants responses after the diversity education session. A cross tabulation of participants' responses for each scenario was performed (see Table 4) which indicates how participants changed their responses from pre to post.

The difference in mean total score from pre (9.95) to post (10.24) is quite small (0.29) and certainly not approaching significance (paired samples t test = 0.44, df = 20). Interestingly the findings showed that there were no consistent changes towards more appropriate responses after receiving diversity education. The total scores tentatively suggest participants perform worse on the SJT after the receiving diversity education (although this is a small change and not near significance).

The findings also suggest that good performance on one scenario concerning a specific diversity issue does not equate to good performance on another scenario concerning a different diversity issue. For example, of the 15 participants who chose the best response to the disability scenario after receiving diversity education, 10 participants chose one of the worst two responses to the healthcare values and beliefs scenario. However there are some positive correlations among the pre scores and among the post scores, which suggests there is some commonality present (Table 5 shows the pre training correlations). Given that the sample size for session 1 is small, a larger sample size would provide a clearer picture of correlations among the scenarios.

Analysis of responses: session 2

Only four of the six scenarios used in pilot session 1 were used in session 2, in response to participants' comments about time constraints. The final version of the four scenarios, their response items and scoring key are shown in Figure 3.

The same statistical tests performed on session 1 data were also performed for this session 2, but the sample was larger (n = 80). The correlation between the pre and post totals was smaller than session 1 (0.165), suggesting more change in participants' responses after receiving diversity education.

A cross tabulation of participants' responses for each scenario was performed (shown in Table 6) which indicates how participants changed their responses from pre to post. As in session 1, the findings of session 2 indicated no consistent changes towards more appropriate responses after receiving diversity education. The difference in mean total scores from pre (7.08) to post (7.45) was small (0.37) and as in session 1 did not approach significance (paired sample t test = 0.1537, df = 75 because of some missing responses). Though the difference is small and not significant, it is in the wrong direction, suggesting that participants performed slightly worse on the SJT after receiving diversity education, as in session 1.

In addition, participants who performed well on one scenario about a particular diversity issue did not consistently perform well in other scenarios exploring different diversity issues. This is illustrated in Table 7, which shows that, of the 57 who gave one of the best responses to the transgender scenario (A or B), 40 (70%) gave one of the two worst responses to the disability scenario (A or D). The same table shows that more than a quarter of those who gave one of the best responses to the disability scenario (B or C) gave one of the worst two responses to the transgender scenario (C or D). Table 8 shows the correlations among the post scores, the largest of which is negative and none of which are significant. This suggests that these four scenarios are exploring largely independent dimensions of diversity awareness.

Stage III Pilot

Comments from participants

The control group (non-NHS) were also positive about the SJT, indicating that the scenarios can be easily understood despite being tailored for a NHS context. This suggests that large scale validation testing should not be problem in the future.

Analysis of responses

The difference between the mean total score for the non-NHS groups (7.50) and the post diversity education mean total for the stage II pilot session 2 group (7.44) was very small (0.06) and did not approach significance, (independent t test= 0.185, df =125).

Discussion

The aim of diversity education is to improve the knowledge, awareness, skills and attitudes of healthcare professionals in serving culturally diverse populations. However without a robust evaluation tool, it is challenging to determine the effectiveness of such teaching. The findings of the pilot stages provide provisional support for the claims of many NHS staff describing the frequent challenges they face in trying to discern the most appropriate response to take when dealing with different diversity issues, and their struggle appears not to be helped by current diversity education. Both the pilot sessions in stage II suggested that NHS professionals performed slightly worse after receiving diversity education and the stage III findings showed a non-NHS group performing almost the same as the NHS groups who received diversity education. This provides tentative support to the consistently reported limitations of diversity education in the findings of the participatory workshops and in the wider literature (Wear et al, 2003; George et al, 2015; Sorenson et al, 2019).

Participants reported that diversity education in the NHS is too prescriptive and attempts to provide fixed answers to complex questions. They suggested that diversity education should focus on exploring and discussing the *questions* around how to effectively deal with diversity issues rather than defining fixed *answers*. The findings revealed that what participants considered appropriate for dealing with one diversity issue may be inappropriate for dealing with another diversity issue, and that responding to

diversity issues requires an active acknowledgement of the context and the individuals involved, an open dialogue and a safe and supportive environment to explore and discuss diversity issues. It suggests diversity education needs to be interactive, relational, participatory and exploratory in helping healthcare professionals first understand the complexity of what diversity means to them in order to then be competent to explore and appreciate the complexity of diversity in others. This is consistent with the findings of the participatory workshops (George et al, 2019) and the reconstructed RCC model, and also with the feedback comments from participants in the stage I pilots.

The findings of the pilot sessions also suggest that a single evaluation tool may be insufficient to measure the effectiveness of diversity education and to capture the complexity of diversity issues. The use of a summative evaluation tool may be more practically feasible in the context of NHS diversity education, but using a combination of formative and summative assessment/evaluation tools may provide more useful, insightful information about how to improve diversity education and its long term impact on professional development and patient outcomes. However, this may be considered time inefficient within NHS contexts. Reliance on a single tool may fail to provide sufficient information on how to improve diversity education and which parts are not adequate in meeting the needs of healthcare professionals and the expectations outlined in health educational policies. A variety of question types such as social vignettes, Likert type questions or true/false statements would ensure the evaluative tool is gathering sufficient information about the respondents' degree of competence in knowledge, attitudes and skills, rather than previous instruments that relied on self-judgements of these levels of competence (George et al, 2019).

Using the scenarios as an educational resource

Participants were consistently positive about the SJT, frequently expressing that it was "*thought provoking*" and very helpful in enabling discussion on diversity issues that were relevant to their clinical practice. All scenarios and proposed response items were distilled from 'critical incidents/stories' related by participants in the discussions in the participatory workshops. Many participants contrasted these scenarios and the quality of the discussions elicited by them with the current diversity education content, which they often found was unhelpful, clinically irrelevant and devoid of open discussions. This suggests that some of the original 90 scenarios could be further developed as an educational resource. This would aid an expansion of the typical fixed blueprint of answers for dealing with diversity issues to an open and respectful discussion. Participants found the discussions most helpful, indicating that the diversity issues they encounter in practice are too varied and complex for fixed answers, and also that the most appropriate response to one diversity issue may not be applicable to another in a different context.

The challenges for healthcare professionals are firstly, recognising diversity issues within themselves and then in others, then secondly learning how to demonstrate mutual respect, flexibility and tolerance, and how to be confident in initiating discussions that are uncomfortable and uncertain (George and Lowe, 2019). This also raises the need for faculty development and for trainers in diversity education to be comfortable in facilitating difficult and sensitive discussions. From the feedback received, using the

scenarios as an educational resource for both the trainers and trainees could potentially be a helpful starting point.

Further development of the SJT

The aim of this study was to develop and pilot an evaluation tool for NHS diversity education and the findings demonstrate significant progress in the development of an effective and plausible SJT, but substantial further development is needed. This study shows that an SJT may be an effective tool for diversity evaluation. However, the case scenarios used may need to reflect local challenges and contexts. Much of their use will also depend on the facilitator's skills in their application if they are to be used for supporting learning. Evaluation may not be dependent merely on the response but the process undergone to reach a conclusion.

Ideally a comparison between a larger sample of NHS health professionals (e.g. $n \geq 500$) and non-NHS individuals who have not undertaken diversity education and are perhaps less aware of diversity issues, is needed to establish whether the SJT can discriminate between such groups, before it can be used to evaluate diversity education. All the pilot sessions showed that people can perform well (selecting the best response) on one scenario concerning a specific diversity issue, but score poorly on another. This resonates with research that suggests different diversity issues require different skills and management (Shen et al, 2016; Sorenson et al, 2019). It accords with our daily experiences that people may be very diversity aware in some areas but quite blind to other diversity factors. In the future it would be valuable to further develop the original bank of 90 scenarios from the participatory workshops into usable scenarios with plausible sets of response items. A larger sample of participants would allow a factor analysis to be conducted to explore the different dimensions of diversity awareness and competence.

It will be apparent from this study that identifying a good set of response items and determining a scoring key is challenging, and even after several iterations, further pilots are needed to refine the response items and scoring keys. At the start, the response formats that were piloted involved ranking in order and choosing the best or the best two options. A format that was not tested in this study was the ranking of each of the items independently, and this should be considered for future developments. To do this, a grid would have to be provided (see Table 9 as an example) with space for participants to rank each item according to how likely they would be to choose it. This approach may mitigate the difficulty some respondents have in deciding on the best option, resulting in them ignoring the instruction and choosing two instead of one. It would also remove the difficulty of deciding on the correct scores for the key, though items would still have to be ranked as better or worse responses in order to calculate a score for each participant. It may be useful to consider the responses to each individual response item, and for participants it may allow a more subtle response if they think there is some merit in more than one of the response items.

Study limitations

There are certainly limitations to our SJT in its current state of development. These include the restriction of responses to a single choice, particularly where the best response may differ depending on the context.

Furthermore, the response items may not provide an adequate range of responses, resulting in candidates feeling forced to select a response that is not necessarily compatible with their values and views. Sometimes they may wish to make a more subtle response that allows them to acknowledge the merits of more than one option. Some of these issues are considered above in the section on further development of the SJT. A further limitation of the work so far is that we have not attempted to investigate reliability, and assess validity. This means that our conclusions about implications for current NHS diversity education can only be tentative. Moving forward, we are engaging in on-going efforts to continuously refine the SJT, we believe this preliminary attempt in developing and piloting a SJT is key for beginning meaningful research on evaluating diversity education in health educational settings. Furthermore, developing different versions of SJTs (including shorter and longer versions) may in future be more useful and practical.

Conclusions

This paper highlights insightful findings and implications for future development of a SJT to evaluate diversity education. The development of a SJT that can be used for routine evaluation of diversity education still needs considerable work, but in the process of bringing it to this level, important issues about the needs of learners and educators and the shortcomings of diversity education were exposed. The scenario based questions in their current format can be used as an educational resource to initiate open and constructive discussions around different diversity issues in healthcare. The tentative scenarios and response items can be further developed to explore how diversity education can be evaluated. While SJTs have rarely been used to evaluate diversity education, the findings of this study show they have considerable potential and versatility.

Abbreviations

CC - Cultural competence

NHS - National Health Service

RCC - Relationship centred care

SJT - Situational Judgement Test

Declarations

Ethics approval and consent to participate

Ethical approval was granted by the University of Leicester. All participants were provided written consent to participate

Consent for publication

All participants and involved parties consented to the findings being published

Availability of data and material

Raw data can be provided on request

Competing interests

The authors note no competing interests

Funding

This study was not funded

Authors' contributions

This study was part of the first author's PhD research project. The second and third author were the PhD supervisors and contributed in a supervisory and advisory capacity

Acknowledgements

The authors would like thank Ms. Harjit Bansal (Equality and Diversity Manager) and Mr. Michael Beackon (Assistant Equality and Diversity Manager) at the North East London Foundation Trust (NELFT) and all the participants who were involved in this study for their support in completing this piece of research.

References

Alizadeh, S., & Chavan, M. (2016). Cultural competence dimensions and outcomes: a systematic review of the literature. *Health & Social Care in the Community*. 24(6):e117-30.

Anderson, L.M., Scrimshaw, S.C., Fullilove, M.T., Fielding, J.E., Normand, J., & Task Force on Community Preventive Services. (2003). Culturally Competent Healthcare Systems: A Systematic Review. *American Journal of Preventive Medicine*. 24 (3): 68-79.

Association of American Medical Colleges (AAMC). (2019). Core Competencies for Entering Medical Students. *Medical Education. Workforce*. Available at: <https://www.aamc.org/services/admissions-lifecycle/competencies-entering-medical-students>

Betancourt, J. (2003). Cross-Cultural Medical Education: Conceptual Approaches and Frameworks for Evaluation. *Academic Medicine*, 78 (6):.560-569.

Betancourt, J.R., Green, A.R., Emilio Carrillo, J. Park, E.R. (2005). Cultural competence and healthcare disparities: Key perspectives and trends. *Health Affairs*, 24 (2): 499-505.

Bhui, K., Ascoli, M., Nuamn O. (2012). The place of race and racism in cultural competence; what can we learn from the English experience about the narratives of evidence and argument? *Transcultural Psychiatry*, 49 (2).

Bhui, K., Warfa, N., Edonya, P., McKenzie, K., & Bhugra, D. (2007). Cultural Competence in Mental Health Care: A Review of Model Evaluations. *BMC Health Services Research*, 7 (15): 1-10.

Bintley, H., **George, R.E.** (2019). (In Press). Teaching diversity in healthcare education: Conceptual clarity and the need for an intersectional trans-disciplinary approach. *Clinical Education for the Health Professions: Theory and Practice*. Springer Nature Elviseer.

Burchum, J.L.R. (2000). Cultural Competence: An Evolutionary Perspective. *Nursing Forum*, 37, pp.5-15.

Chan, D., Schmitt, N., Sacco, J.M., De-Shon, R.P. (1998). Understanding Pre-Test and Post-Test Reactions to Cognitive Ability and Personality Tests. *Journal of Applied Psychology*, 83 (3): 471-485.

Cross, T.L., Bazron, B.J., Isaacs, M.R., & Dennis, K.W. (1989). Towards a Culturally Competent System of Care. A Monograph on Effective Services for Minority Children Who are Severely Emotionally Disturbed. Washington, DC: Georgetown University Centre Child Health and Mental Health Policy, CASSP Technical Assistance Center. Available from: http://csmha.umaryland.edu/how/cultural_competency_2001.

Department of Health. 2012. Equality Objectives Action Plan. Available at: <https://www.gov.uk/government/publications/department-of-health-equality-objectives-2012-to-2016>

Dogra, N., Reitmanova, S., & Carter-Pokras, O. (2010). Teaching Cultural Diversity: Current Status in UK, US, and Canadian Medical Schools. *Journal of General Internal Medicine*. 25 Suppl 2. S164-8. 10.1007/s11606-009-1202-7.

Dogra, N., Bhatti, F., Ertubey, C., Kelly, M., Rowlands, A., Singh, D. & Turner, M. (2015). Teaching Diversity to Medical Undergraduate; Curriculum Development, Delivery and Assessment. *AMEE Guide No. 103*. *Medical Teacher*, 38 (4): 323-337.

General Medical Council (GMC). (2009). *Medical Students: Professional Values and Fitness to Practice*. London: General Medical Council.

George, R.E., Dogra, N., & Thornicroft, G. (2015). Exploration of Cultural Competency Training in UK Healthcare Settings. A Critical Interpretive Review of the Literature. *Diversity and Equality in Health and Care*. 12 (3), pp. 104-115.

George, R.E., Smith, K., O'Reilly, M., Dogra, N. (2019). Perspectives of Patients with Mental Illness on How to Better Teach and Evaluate Diversity Education in the National Health Service. *Journal of Continuing Education in Healthcare Professions*. 39 (2): 92-102.

George, R.E., Lowe, W.A. (2019). Wellbeing and Uncertainty in Healthcare Practice. Faculty Development Review. Special Issue: Healthcare Professionals' and Students' Wellbeing. The Clinical Teacher. 16: 298-305.

Gozu, A., Beach, M.C., Price, E., Gary, T.L., Robinson, K., Palacio, A., Smarth, C., Jenckes, M., Feuerstein, C., Bass, E.B., Powe, N.R., & Cooper, L.A. (2007). Self-Administered Instruments to Measure Cultural Competence of Health Professionals: A Systematic Review. Teaching and Learning in Medicine, 19 (2): 180-190.

Hobgood, C., Sawing, S., Bowen, J., Savage, K. (2006). Teaching Culturally Appropriate Care: A Review of Educational Models and Methods. Educational Advances. Academic Emergency Medicine, 13 (12): 1288-1295.

Kai, J., Beavan, J., Faull, C., Dodson, L., Gill, P., & Beighton, A. (2007). Professional Uncertainty and Disempowerment Responding to Ethnic Diversity in Health Care: A Qualitative Study. PLoS Medicine. 4(11): e323.

Kumagai, A. & Lyson, M.A. (2009). Beyond Cultural Competence: Critical Consciousness, Social Justice and Multi-cultural Education. Academic Medicine. 84 (6): 782-787.

Kumas-Tan, Z., Beagan, B., Loppie, C. MacLeod, A., & Frank, B. (2007). Measures of Cultural Competence: Examining Hidden Assumptions. Academic Medicine. 82(6): 548-557.

Loftin, C., Hartin, V. Branson, M. & Reyes, H. (2013). Measures of Cultural Competence in Nurses: An Integrative Review. The Scientific World Journal. Vol. 2013: 1-10.

Malat, J. (2013). The Appeal and Problems of a Cultural Competence Approach to Reducing Racial Disparities. Journal General Internal Medicine, 28 (5): 605-7.

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 British Psychology. 24(3):281–88.

Nursing Midwifery Council (NMC). (2010). Standards for Competence for Registered Nurses. London. Available at: <https://www.nmc.org.uk/globalassets/sitedocuments/standards/nmc-standards-for-competence-for-registered-nurses.pdf>

Patterson, F. Zibarras, L. & Ashworth, V. (2016). Situational Judgement Tests in Medical Education and Training: Research, Theory and Practice: AMEE Guide No.100. Medical Teacher, 38 (1): 3-17.

Price, E. Beach, M.C., Gary, T.L., Robinson, K.A., Gozu, A., Palacio, A., Smarth, C., Jenckes, M., Feuerstein, C., Bass, E.B., Powe, N. & Cooper, L.A. (2005). A Systematic Review of the Methodological Rigor of Studies Evaluating Cultural Competence Training of Health Professionals. Academic Medicine. 80 (6): 578-586.

Reitmanova, S. (2011). Cross-Cultural Undergraduate Medical Education in North America: Theoretical Concepts and Educational Approaches, *Teaching and Learning in Medicine*. *Teaching and Learning in Medicine*, 23 (2): 197-203.

Renzaho, A.M.N., Romios, P., Crock, C., & Sonderlund, A.L. (2013). The Effectiveness of Cultural Competence Programs in Ethnic Minority Patient-Centred Health Care – A Systematic Review of the Literature. *International Journal for Quality in HealthCare*. 25 (3): 261-269.

Shen, Z. (2015). Cultural Competence Models and Cultural Competence Assessment Instruments in Nursing: A Literature Review. *Journal of Transcultural Nursing*, 26 (3):308-321.

Sorenson, J., Norredam, M., Suurmond, J., Carter-Pokras, O., Garcia-Ramirez, M., & Krasnik, A., (2019). Need for Ensuring Cultural Competence in Medical Programmes of European Universities. *BMC Medical Education*. 19(21): 1-8.

Stanhope, V., Solomon, P., Pernell-Arnold, A., Sands, R.G., & Bourjolly, J.N. (2005). Evaluating Cultural Competence Among Behavioural Health Professionals. *Psychiatric Rehabilitation Journal*, 28 (3): 225-233.

Suzuki, L.A., McRae, M.B. & Short, E.L. (2001). "The Facets of Cultural Competence: Searching Outside the Box." *The Counselling Psychologist*, 29 (6): 842-9.

Tresolini, R.M. (1994) Pew-Fetzer Task Force on Advancing Psychological Health Education. *Health Professions Education and Relationship-Centred Care*. San Francisco, CA: Pew Health Professions Communication.

Truong, M., Paradies, Y. & Priest, N. (2014). Interventions to Improve Cultural Competency in Healthcare: A Systematic Review of Reviews. *BioMed Central Health Services Research*, 14 (99): 1-17

Wear, D. (2003). Insurgent multiculturalism: rethinking how and why we teach culture in medical education. *Academic Medicine*, 78:549– 554.

Tables

Table 1: Example of feedback from Stage I Pilots

STAGE I PILOTING			
EXAMPLE SCENARIO	Example Mental health Patient Feedback	Example Mental health Patient Feedback	Example Mental health Patient Feedback
<p>Scenario: You are based in a GP surgery and have been asked to take a history from Chris Briggs. When he comes into the room to talk to you, you think he is a man dressed as a woman.</p>	<ul style="list-style-type: none"> Reasoning for proposed choice: "Cross dressing is probably normal for them, not clinically relevant. If not clinically relevant there is no need to ask the patient about their gender." "This will make the professionals think about these issues from different perspectives." Suggested improvements: "Some questions were easier to understand than other, would use less responses, maybe 4." 	<ul style="list-style-type: none"> Reasoning for proposed choice: "If a person decides to dress like a lady that is a personal choice that is a freedom. The person should just be treated as an individual. A person's sexual or dress sense should not be an obstacle to receiving care, we need to be comfortable with differences." 	<ul style="list-style-type: none"> Reasoning for proposed choice: "To consider the patient's culture which includes their gender should be covered any way in part of his care? It is a good idea for a GP to know the patient their culture and language to build rapport." "Interesting education resource for the training."
<p>Scoring: Choose the most appropriate actions to take in this situation.</p>			
<p>Proposed Responses:</p> <p>A. Ignore and do not ask about the person's gender as you feel it is not clinically relevant.</p> <p>B. Continue with the consultation as normal.</p> <p>C. Ask about the patient's gender to clarify if they identify as a man or a woman.</p> <p>D. Only ask about the patient's gender if it is deemed clinically relevant.</p> <p>E. Do not ask about the patient's gender as this may</p>	<p>Answers:</p> <p>A. Ignore and do not ask about the person's gender as you feel it is not clinically relevant.</p> <p>B. Continue with the consultation as normal. (ü)</p> <p>C. Ask about the patient's gender to clarify if they identify as a man or a woman.</p> <p>D. Only ask about the patient's gender if it is deemed clinically relevant.</p> <p>E. Do not ask about the patient's gender as this may offend the patient.</p>	<p>Answers:</p> <p>A. Ignore and do not ask about the person's gender as you feel it is not clinically relevant.</p> <p>B. Continue with the consultation as normal.</p> <p>C. Ask about the patient's gender to clarify if they identify as a man or a woman.</p> <p>D. Only ask about the patient's gender if it is deemed clinically relevant. (ü)</p> <p>E. Do not ask about the patient's gender as this may offend the patient.</p>	

offend the patient.

clarify if they identify as a man or a woman.

D. Only ask about the patient's gender if it is deemed clinically relevant. (ü)

E. Do not ask about the patient's gender as this may offend the patient.

Table 2: Examples of Different Response Format and Changes in Item Scoring

MULTIPLE CHOICE FORMAT	RANKING RESPONSE FORMAT	BEST SINGLE RESPONSE FORMAT
<p>Scenario: Rachel formally known as Richard is a transgender nurse staff and has made you, the care manager aware that a number of her colleagues are bullying her and are preventing her from using the female toilets. Rachel presents herself as female throughout her employment, yet many of the staff as well as occasionally patients feel uncomfortable with her using the female toilets. How would you respond to this situation?</p>	<p>Scenario: Rachel formally known as Richard is a transgender nurse staff and has made you, the care manager aware that a number of her colleagues are bullying her and are preventing her from using the female toilets. Rachel presents herself as female throughout her employment, yet many of the staff as well as occasionally patients feel uncomfortable with her using the female toilets.</p>	<p>Scenario: Rachel formerly known as Richard is a transgender staff nurse. She has made you aware as a colleague that she is being bullied by several colleagues and that she is being prevented from using the female toilets.</p>
<p>Scoring: Choose the MOST appropriate ACTIONS you would take in this situation.</p>	<p>Scoring: Please RANK in order from 1 to 5 (1 being the most inappropriate and 5 being the most appropriate) the following responses to this scenario:</p> <p>Please note ranks can be tied.</p>	<p>Scoring: Choose the ONE most appropriate action to take in this situation.</p>
<p>Proposed Responses:</p> <ul style="list-style-type: none"> A. Discuss with the Trust Board the need to provide separate toilets specifically for transgender members of staff. B. Explore Rachel’s concerns and ask other staff for their experiences and feedback on this matter. C. Would not consider this a major issue of bullying. D. Reassure Rachel that she would discuss the issue with the staff team lead to ensure she does have access to the female toilets. E. Review ‘Equality & Diversity’ policy on managing transgender issues. F. Suggest an option of using the unisex disabled toilets. 	<p>Proposed Responses:</p> <ul style="list-style-type: none"> A. Discuss with the Trust Board the need to provide separate toilets specifically for transgender members of staff. B. Explore Rachel’s concerns and ask staff for their experiences and feedback on this matter. C. Would not consider this a major issue of bullying and suggest an option of using the unisex disabled toilets. D. Reassure Rachel that you would discuss the issue with the staff team lead to ensure she does have access to the female toilets. E. Review ‘Equality & Diversity’ policy on managing transgender issues. 	<p>Proposed Responses:</p> <ul style="list-style-type: none"> A. Raise the issue in a team meeting in a way that everyone can safely and openly discuss the matter and resolve together. B. Explore Rachel’s concerns and ask other staff for their experiences and feedback on this matter but highlight to all colleagues that bullying is not acceptable under any circumstances. C. Discuss with other staff their concerns and if legitimate ask Rachel to use the unisex disabled toilets. D. Reassure Rachel that you would discuss the issue

with the staff team leader to ensure she does have access to the female toilets.

Reasons for Item Scoring:

The three responses deemed the most appropriate are in italics. These responses and scoring were explored in phase I pilot.

A. Discuss with the Trust Board the need to provide separate toilets specifically for transgender members of staff.

Justification: Although this provides a resolution to the problem, it does not attempt to actively engage and address Rachel and staff concerns and may further marginalise/stigmatise transgender members of staff.

B. Explore Rachel's concerns and ask other staff for their experiences and feedback on this matter.

Justification: This response takes into account both Rachel's and staff perspectives and experiences before making any actions.

C. Would not consider this a major issue of bullying.

Justification: This response fails to recognise this issue as 'bullying'.

D. Reassure Rachel that she would discuss the issue with the staff team lead to ensure she does have access to the female toilets.

Justification: This response attempts to address and resolve Rachel's concerns and legally this may appear to be a popular response as Rachel is protected. However it does not attempt to gain the perspectives of other staff members and does not necessarily change the culture.

E. Review 'Equality & Diversity' policy on managing transgender issues.

Justification: This response highlights the importance of

Reasons for Item Scoring:

Responses were given a predefined score of the following. These scores were determined collaboratively by a range of stakeholder perspectives:

(1=most inappropriate and 5=most appropriate)

A. Discuss with the Trust Board the need to provide separate toilets specifically for transgender members of staff. (3)

Justification: Although this provides a resolution to the problem, it does not attempt to actively engage and address Rachel and staff concerns and may further marginalise/stigmatise transgender members of staff.

B. Explore Rachel's concerns and ask staff for their experiences and feedback on this matter. (5)

Justification: This response takes into account both Rachel's and staff perspectives and experiences before making any actions.

C. Would not consider this a major issue of bullying and suggest an option of using the unisex disabled toilets. (1)

Justification: This response fails to recognise the issue as 'bullying' and although suggests the 'unisex' disabled toilets as an appropriate option, it does not actively explore or discuss Rachel's or staff concerns.

D. Reassure Rachel that you would discuss the issue with the staff team lead to ensure she does have access to the female toilets. (2)

Reasons for Item Scoring:

Responses were given a predefined score of the following. These scores were determined collaboratively by a range of stakeholder perspectives:

1= most appropriate and 4=inappropriate

A. Raise the issue in a team meeting in a way that everyone can safely and openly discuss the matter and resolve together. (1)

Justification: This response takes into account different perspectives and tries to lead to respect and understanding. It also highlights the issue of 'safety' and ensuring a supportive environment to discuss these sensitive issues.

B. Explore Rachel's concerns and ask other staff for their experiences and feedback on this matter but highlight to all colleagues that bullying is not acceptable under any circumstances. (2)

Justification: This response is also appropriate in that it makes clear that bullying is not permitted and raises the issue but there may be no resolution.

healthcare professionals being educating and aware of policies and guidelines around 'equality and diversity' on managing transgender issues. Although it does not actively seek staff perspectives.

F. Suggest an option of using the unisex disabled toilets.

Justification: This option does provide a practical resolution to the issue, although it does not explore Rachel's and staff concerns to prevent this from happening again.

Justification: This response attempts to address and resolve Rachel's concerns and legally this may appear to be a popular response as Rachel is protected. However it does not attempt to gain the perspectives of other staff members and does not necessarily change the culture.

E. Review 'Equality & Diversity' policy on managing transgender issues.(4)

Justification: This response highlights the importance of healthcare professionals being educating and aware of policies and guidelines around 'equality and diversity' on managing transgender issues. Although it does not actively seek staff perspectives.

C. Discuss with other staff their concerns and if legitimate ask Rachel to use the unisex disabled toilets. (4)

Justification: This response is the least acceptable as it appears to favour one the side of colleagues and may make Rachel feel that her issues are being talked about. Using the unisex toilets may leave the issue unresolved, although may not in itself be inappropriate.

D. Reassure Rachel that you would discuss the issue with the staff team leader to ensure she does have access to the female toilets. (3)

Justification: This response attempts to address and resolve Rachel's concerns and legally this may appear to be a popular response as Rachel is protected. However it does not attempt to gain the perspectives of other staff members and does not necessarily change the culture.

Table 3: Stage II Pilot Session 1: Cross Tabulation

Scenario 1: Religion and Health Policy

	After A	B	C	D	
Before A	1	2	2	0	5
B	2	8	2	0	12
C	1	1	4	0	6
D	0	0	0	0	0
	4	11	8	0	

Scenario 2: Health beliefs

	After A	B	C	D	
Before A	2	0	1	0	3
B	0	10	1	0	11
C	0	1	5	0	6
D	0	0	1	0	1
	2	11	8	0	

2 spoiled

Scenario 3: Sexual orientation

	After A	B	C	D	
Before A	0	0	1	0	1
B	0	11	1	3	15
C	0	1	1	0	2
D	0	2	1	2	5
	0	14	4	5	

Scenario 4: Language

	After A	B	C	D	
Before A	0	1	3	0	4
B	0	0	0	0	0
C	1	3	13	0	17
D	0	0	0	0	0
	1	4	16	0	

2 spoiled

Scenario 5: Disability

	After A	B	C	D	
Before A	1	0	2	0	3
B	1	2	2	0	5
C	0	4	11	0	15
D	0	0	0	0	0
	2	6	15	0	

Scenario 6: Religion and Service Delivery

	After A	B	C	D	
Before A	2	0	2	0	4
B	0	0	0	0	0
C	0	0	13	1	14
D	1	0	2	1	4
	3	0	17	2	

1 spoiled

Table 4: Stage II Pilot Session 1 Pre training Correlations

	Religion & Health Policy	Values & Health Beliefs	Sexual Orientation	Language	Disability	Religion & Health Service Delivery
Religion & Health Policy	1	-.058	.123	-.185	.073	.045
Values & Health Beliefs	-.058	1	.333	.140	.011	.026
Sexual Orientation	.123	.333	1	.216	.552	.418
Language	-.185	.140	.216	1	.526	-.285
Disability	.073	.011	.552	.526	1	-.052
Religion & Health Service Delivery	.045	.026	.418	-.285	-.052	1

Table 5: Stage II Pilot Session 2: Cross Tabulation

Transgender

	After A	B	C	D	
Before A	4	10	3	0	17
B	3	24	5	0	32
C	1	13	11	1	26
D	1	1	0	1	3
	9	48	19	2	78

2 spoiled

Language

	After A	B	C	D	
Before A	57	2	3	11	73
B	0	0	0	0	0
C	2	1	1	0	4
D	2	0	0	0	2
	61	3	4	11	79

1 spoiled

Disability

	After A	B	C	D	
Before A	48	8	6	1	63
B	3	4	1	0	8
C	3	0	5	0	8
D	0	0	0	0	0
	54	4	12	1	79

1 spoiled

Religion

	After A	B	C	D	
Before A	1	5	1	0	7
B	3	25	6	1	35
C	0	24	8	1	33
D	0	2	2	0	4
	4	56	17	2	79

1 spoiled

Table 6: Stage II Pilot Session 2: post training responses to transgender and disability

		Disability post				
		A	B	C	D	
Transgender post	A	4	1	4	0	9
	B	35	6	6	1	48
	C	13	4	1	0	18
	D	1	1	0	0	2
		53	12	11	1	77

Table 7: Stage II Pilot Session 2 Post training Correlations

	Religion	Disability	Transgender	Language
Religion	1	-0.096	0.101	-0.030
Disability	-0.096	1	-0.184	0.137
Transgender	0.101	-0.184	1	0.073
Language	-0.030	0.137	0.073	1

Table 8: A response grid for ranking every item

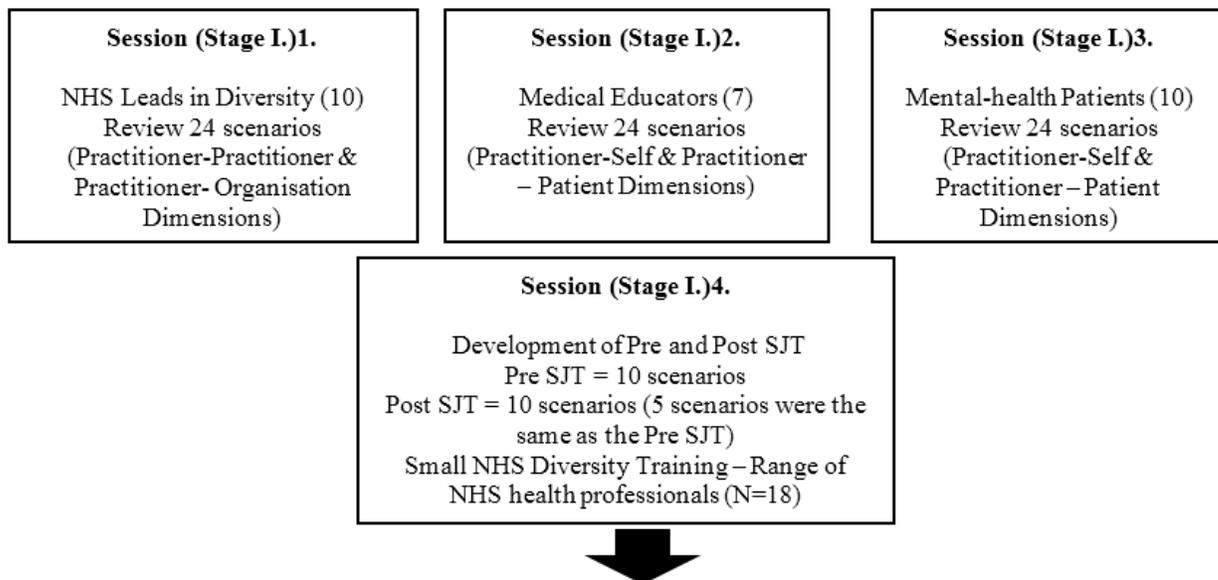
Please indicate how likely you would be to choose each option

	Very likely	Perhaps	Not very likely	Very unlikely
A Raise the issue in a team meeting in a way that everyone can safely and openly discuss the matter and resolve together.				
B Explore Rachel’s concerns and ask other staff for their experiences and feedback on this matter, but emphasise to all colleagues that bullying is not acceptable under any circumstances				
C Discuss with the other staff their concerns and if legitimate ask Rachel to use the unisex disabled toilets				
D Reassure Rachel that you would discuss the issue with the staff team leader to ensure she does have access to the female toilets				

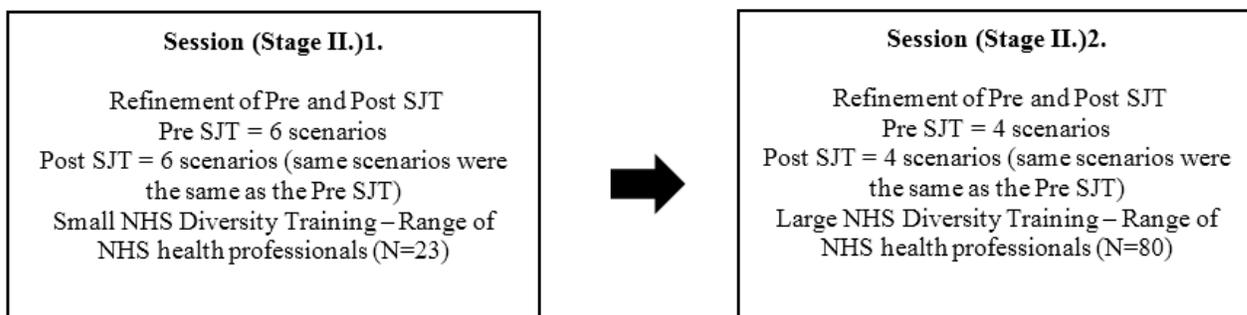
Scoring key

Scenario 1: Transgender	Scenario 2: Language	Scenario 3: Disability	Scenario 4: Religion
A=1	A=3	A=3	A=3
B=2	B=4	B=2	B=4
C=4	C=1	C=1	C=2
D=3	D=2	D=4	D=1

Figures



Stage II Pilot: Pre and Post Test Development and Construction



Stage III Pilot: Comparison with Control Groups (Non-NHS group)

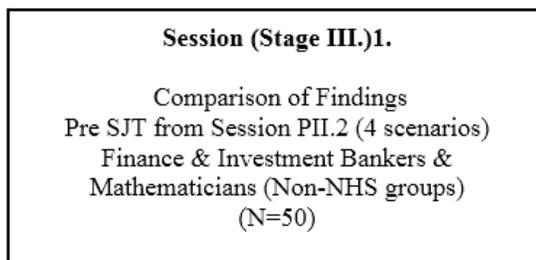


Figure 1

Overview of Pilot Stages Stage I Pilot: Key Stakeholder Input & Expertise Refinement of Scenario and Item Development, Response Format and Scoring and Test Construction

INSTRUCTIONS

This 'Situational Judgment Test' (SJT) is a collection of scenario-based questions that are **designed to evaluate your judgement abilities in healthcare situations involving diversity issues**. You will be presented with a set of six hypothetical work-based scenarios and will be asked to **choose the SINGLE most appropriate response** to the scenario. Please read each scenario based question carefully and try to answer as honestly as possible.

PLEASE CHOOSE THE MOST APPROPRIATE RESPONSE FOR EACH SCENARIO BY PLACING A TICK IN ONE OF THE BOXES AT THE SIDE.

SCENARIO 1

A Muslim female doctor claims that the organisational policy that all staff must be 'bare below the elbows' (hands and arms up to the elbow are exposed and free from clothing) is against her religious beliefs. Choose the one most appropriate action you would take.

- A Respect that this is part of her religious beliefs and accept that she will not be abiding by the 'bare below the elbows' rule
- B Discuss the issue with the rest of the team and ask for advice from a senior member of the team
- C Explain to your colleague that all health professionals have to abide by healthcare rules and policies and that no adjustments can be made
- D Allow your colleague to continue not abiding by the 'bare below the elbow' rule until any formal complaints are made

SCENARIO 2

One of your colleagues is very pro-life and objects to terminations of pregnancy based on her religious views. She has admitted to finding it difficult to bring herself to be compassionate towards any patient that was having a termination. She is otherwise a very caring and compassionate professional. How would you help your colleague?

- A Suggest she refuses to care for patients who are considering terminations
- B Allow your colleague to express openly her cultural beliefs and suggest she attends appropriate training to help her manage and deal with the situation she is struggling with
- C Discuss with your colleague that as health professionals we must treat all patients the same regardless of their beliefs and choices
- D Explain to your colleague that if she is treating those considering terminations differently that her behaviour is unacceptable and could result in disciplinary action

SCENARIO 3

Rachel formerly known as Richard is a transgender staff nurse. She has made you aware as a colleague that she is being bullied by several colleagues and that she is being prevented from using the female toilets. Choose the one most appropriate action you would take

- A Write to your service manager asking them to make the trust board aware of the need to provide separate toilets specifically for transgender members of staff
- B Explore Rachel's concerns and ask other staff for their experiences and feedback on this matter but highlight to all colleagues that bullying is not acceptable under any circumstances
- C Discuss with the other staff their concerns and if legitimate ask Rachel to use the unisex disabled toilets
- D Reassure Rachel that you would discuss the issue with the staff team leader to ensure she does have access to the female toilets

SCENARIO 4

Your colleague comes to you on their lunch break and is very frustrated about a patient that took up most of her morning as they were unable to understand what was written on their medication packets. She says, "they should really learn English if they are coming to this country, or at least try shouldn't they?" Choose the one most appropriate action you would take.

- A Sympathise with your colleague's frustration and point out that an interpreter should be called (if available) if there are difficulties
- B Explain to your colleague that their comment may be seen as racist and is therefore inappropriate
- C Explain to your colleague that as health professionals we are here to care for all patients with different needs and concerns and suggest an interpreter
- D Sympathise with your colleague's comment and take no further action

SCENARIO 5

You hear a patient who has dementia say, "Are you able to move around in that chair as much as the others?", to a colleague who is a wheelchair user. Your colleague responds in a rude and offensive manner. Choose the one most appropriate action you would take.

- A Intervene in the situation, asking your colleague to leave the patient immediately
- B Say to the patient that the Trust's Equality and Diversity stance makes their comment unacceptable
- C Discuss with your colleague that you witnessed unprofessional behaviour and that whilst it may be understandable, it is not acceptable
- D Ignore the situation as you feel it is none of your business and carry on with your work as usual

SCENARIO 6

A new member of staff, Miriam, reveals during her induction that she is Jewish and will need to finish work early on a Friday in the winter. She has not mentioned this earlier when specifically asked if there were any special circumstances. The manager appears apprehensive after her request and says she will have to discuss the request with a Senior Manager. The Senior Manager laughs and says, "Don't we all want to finish early on a Friday?" If you were the Senior Manager, what would you do in this case?

- A Acknowledge that Miriam has the right to request flexibility in her working hours and is entitled to finish early on Fridays
- B Refuse her request on the basis of increasing patient demands and staff pressures on Friday
- C Accept Miriam's request and discuss with the hospital team how this arrangement could potentially work for Miriam
- D Explain to Miriam that the contract of employment that she signed made the working hours clear, and that it would not be fair to her other colleagues to make special arrangements for her

THANK YOU FOR COMPLETING THE QUESTIONNAIRE

Figure 2

Phase II pilot session 1 SJT

INSTRUCTIONS

This 'Situational Judgment Test' (SJT) is a collection of scenario-based questions that are designed to evaluate your judgement abilities in healthcare situations involving diversity issues. You will be presented with a set of four hypothetical work-based scenarios and will be asked to choose the SINGLE most appropriate response to the scenario. Please read each scenario based question carefully and try to answer as honestly as possible.

PLEASE CHOOSE THE MOST APPROPRIATE RESPONSE FOR EACH SCENARIO BY PLACING A TICK IN ONE OF THE BOXES AT THE SIDE.

SCENARIO 1

Rachel formerly known as Richard is a transgender staff nurse. She has made you aware as a colleague that she is being bullied by several colleagues and that she is being prevented from using the female toilets. Choose the one most appropriate action you would take

- A Raise the issue in a team meeting in a way that everyone can safely and openly discuss the matter and resolve together.
- B Explore Rachel's concerns and ask other staff for their experiences and feedback on this matter but highlight to all colleagues that bullying is not acceptable under any circumstances
- C Discuss with the other staff their concerns and if legitimate ask Rachel to use the unisex disabled toilets
- D Reassure Rachel that you would discuss the issue with the staff team leader to ensure she does have access to the female toilets

SCENARIO 2

Your colleague comes to you on their lunch break and is very frustrated about a patient that took up most of her morning as they were unable to understand what was written on their medication packets. She says, "they should really learn English if they are coming to this country, or at least try shouldn't they?" Choose the one most appropriate action you would take.

- A Point out that an interpreter should be called (if available) if there are difficulties
- B Explain to your colleague that their comment may be seen as racist and is therefore inappropriate
- C Explain to your colleague that as health professionals we are here to care for all patients with different needs and concerns and suggest an interpreter
- D Sympathise with your colleague's comment and raise this issue in a team meeting

SCENARIO 3

You hear a patient who has dementia say, "Are you able to move around in that chair as much as the others?", to a colleague who is a wheelchair user. Your colleague responds in a rude and offensive manner. Choose the one most appropriate action you would take.

- A Intervene in the situation, asking your colleague to leave the patient immediately
- B Say to the patient that the Trust's Equality and Diversity stance makes their comment unacceptable
- C Discuss with your colleague that you witnessed unprofessional behaviour and that whilst it may be understandable, it is not acceptable
- D Ignore the situation as you feel it is none of your business and carry on with your work as usual

SCENARIO 4

A new member of staff, Miriam, reveals during her induction that she is Jewish and will need to finish work early on a Friday in the winter. She has not mentioned this earlier when specifically asked if there were any special circumstances. The manager appears apprehensive after her request and says she will have to discuss the request with a Senior Manager. The Senior Manager laughs and says, "Don't we all want to finish early on a Friday?" If you were the Senior Manager, what would you do in this case?

- A Acknowledge that Miriam has the right to request flexibility in her working hours and is entitled to finish early on Fridays
- B Grant her request as it is made on religious grounds
- C Discuss the request with the team around how this arrangement could potentially work for Miriam
- D Explain to Miriam that the contract of employment that she signed made the working hours clear, and that it would not be fair to her other colleagues to make special arrangements for her

(REMOVED SCENARIO 1)

A Muslim female doctor claims that the organisational policy that all staff must be 'bare below the elbows' (hands and arms up to the elbow are exposed and free from clothing) is against her religious beliefs. Choose the one most appropriate action you would take.

- A Respect that this is part of her religious beliefs and accept that she will not be abiding by the 'bare below the elbows' rule
- B Discuss the issue with the rest of the team and ask for advice from a senior member of the team
- C Explain to your colleague that all health professionals have to abide by healthcare rules and policies and that no adjustments can be made
- D Allow your colleague to continue not abiding by the 'bare below the elbow' rule until any formal complaints are made

(REMOVED SCENARIO 2)

One of your colleagues is very pro-life and objects to terminations of pregnancy based on her religious views. She has admitted to finding it difficult to bring herself to be compassionate towards any patient that was having a termination. She is otherwise a very caring and compassionate professional. How would you help your colleague?

- A Suggest she refuses to care for patients who are considering terminations
- B Allow your colleague to express openly her cultural beliefs and suggest she attends appropriate training to help her manage and deal with the situation she is struggling with
- C Discuss with your colleague that as health professionals we must treat all patients the same regardless of their beliefs and choices
- D Explain to your colleague that if she is treating those considering terminations differently that her behaviour is unacceptable and could result in disciplinary action

THANK YOU FOR COMPLETING THE QUESTIONNAIRE

Figure 3

Phase II pilot session 2 SJT and scoring key