PERSPECTIVES ON ADDRESSING RACIAL INEQUALITY IN THE NHS

by

MADELEINE ROBERTS

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF CLINICAL PSYCHOLOGY

Centre for Applied Psychology

School of Psychology

The University of Birmingham

December 2022

Thesis Overview

Literature Review

Racial discrimination is considered a contributor to the world health burden due to its association with an array of health difficulties. As such, continued research and reviews are vital for monitoring these associations. A meta-analysis was conducted to explore the association between perceived racial discrimination (PRD) and anxiety. A previous meta-analysis by Paradies et al (2015) was used as a comparative study. The search for the current review continued on from the end date of this 2015 review to provide an update. Fifty-five primary studies were included in the final analysis. The results revealed a small but significant association between PRD and anxiety. The quality of the evidence base and moderators of the association were explored. Recommendations for future research suggest the quality and diversity of the evidence base requires improvement.

Empirical Research Paper

Racial inequality in the NHS has been recognised by stakeholders and in research for years. However, strategies employed to enact change have been considered limited in their impact. Developing a shared understanding of stakeholders' beliefs, wants, and barriers to addressing racial inequality is considered an important but neglected area of research. Q-methodology was employed to explore stakeholder perspectives on how to address racial inequality most effectively in the NHS. Experts by Experience and NHS staff were recruited to provide their perspectives. Factor analysis revealed three distinct opinion groups. What defined and differentiated these opinion groups are explored. The strengths and limitations of the study are also considered with recommendations for future research made.

Acknowledgements

I would sincerely like to thank all the participants who took part in my study, as well as the participants and researchers from other studies drawn upon in this thesis. I have been privileged, inspired, and humbled to hear powerful global majority voices and happy to hear the voices of other White allies.

The patience and kindness of my parents, family, and amazing friends has been more important than they will ever know. I have moved away, missed events, missed talking, and missed spending time with you all. Knowing that you were behind me and would wait for me is something I will always be grateful for.

Devon, this journey would not have been the same without you. Your support academically, professionally, juggling a doctorate and COVID, and remembering to have fun was vital to me making it through those three years of training.

My research supervisors, Dr Theresa Powell and Dr Chris Jones, I thank you for your expertise and dedication in supporting me through my extended journey. You have taught me so much. I will be happy to give you your Friday mornings back but will very much miss learning from you.

Lastly, to Lil, my most incredible fiancée., to Oscar, our beautiful boy, and to Skye and Sprite, the best puppies ever. You are my rocks. In your own wonderful ways, you each gave me the strength and love to try my very best. Whether that was by cooking me dinner every night, constantly supplying me with supportive words, sharing your infectious sparkle, or just giving me a furry cuddle. Thank you.

Table of Contents

List of Figures

Literature Review

Empirical Paper

List of Tables

Literature Review

Empirical Paper

Preliminary Listings

Ethnicity is defined in this thesis as a group someone identifies with based on their ancestry, culture, language, or religion.

Ethnic minority person is used in this thesis to describe people who identify with an ethnic group that in the UK (or other country when applicable) is a minority compared to the majority ethnic group, namely White British in the UK.

Race is recognised in this thesis as a human-made categorisation system defined by a person's physical characteristics (e.g., skin colour). Due to the historical context of grouping people by race, and its limitation to only considering physical characteristics, ethnicity will instead be used to describe people in this thesis, although it is recognised that this categorisation also has its limitations. Please note where excerpts from documents or quotations drawn upon within this thesis have used the term 'race', the wording was not altered from its original phrasing.

Literature Review

Perceived racial discrimination (PRD) is used in this review, in keeping with the terminology of the research literature, to refer to subjective experiences of retrospective racial discrimination as perceived by participants.

Empirical Research Paper

Racial inequality was considered the most helpful term by people with lived experience interviewed in this study. Racial inequality was considered to best encompass the range of implicit, explicit, and systemic discrimination experienced by ethnic minority people in the UK.

1. LITERATURE REVIEW: THE RELATIONSHIP BETWEEN PERCEIVED RACIAL DISCRIMINATION AND ANXIETY: A COMPARATIVE META-ANALYSIS

Abstract

Aim

Racial discrimination has been associated with a variety of poor mental and physical health outcomes. Paradies et al (2015) provided a comprehensive review of the association between perceived racial discrimination (PRD) and several health outcomes. A significant association between PRD and anxiety was identified. However, an updated systemic review had not been completed since. The current meta-analysis aimed to provide an updated review of the association between PRD and anxiety.

Method

A systematic search of several electronic databases was completed. Primary studies published between October 2013 and June 2020 were systematically screened against a matrix of inclusion and exclusion criteria. Fifty-five primary studies, and 57 effects, were included in the final analysis. Pearson's correlation coefficients were extracted, or converted, from primary studies. Heterogeneity of variance and risk of bias were also investigated.

Results

A random effects model was used to generate an overall effect (r = .26). When compared to the overall effect found in Paradies et al (2015), there was no significant difference. Sub-group analyses revealed that studies with high levels of reporting bias and detection bias had significantly lower effect estimates. Further sub-group analyses and meta-regressions explored the impact of additional study and participant level moderators on the association between PRD and anxiety.

Conclusion

The findings show that PRD continues to have an association with anxiety. Future research was advised to focus on improving the quality and consistency of the evidence base. Focus on researching underrepresented groups was also recommended.

Introduction

Due to its association with poor health, racial discrimination has been described as a contributor to the global health burden (Conklin, 2011; Rehm & Shield, 2019). Research into the effects of racial discrimination is important for better understanding its impact at both individual and societal levels. Racial discrimination can be defined as the differential, unfair, or harmful treatment of someone because of their perceived ethnicity or race (de Mendoza et al., 2018). In Western societies, this is linked to a pervasive power and oppression relationship between the majority ethnicity, White British, and ethnic minority people (Kirkinis et al., 2021). Racial discrimination can mistakenly be limited to explicit acts directed from one individual to another (Joseph, 2015; McCoy, 2020). However, racial discrimination can be expressed through racial micro-aggressions (Ackerman-Barger et al., 2020), oppressive systemic structures (Evans et al., 2020) and the induction of internalised racist beliefs (Seet, 2021). Racial discrimination can therefore be experienced both directly (Atwal & Wang, 2019) and vicariously (Huynh et al., 2017; Szaflarski & Bauldry, 2019). The evidence base has previously been limited in capturing a breadth and depth of experiences. However, more recent research has started to include more diverse populations and demographics (Drakeford, 2019; Heard-Garris et al., 2018; Leath et al., 2019; Priest et al., 2013; Szaflarski & Bauldry, 2019). Despite this, there are still considered to be many limitations in the research literature in terms of the diversity of participants and study quality (Pizarro & Kohli, 2020; Schmitt et al., 2014). The continued expansion of this evidence base means systematic reviews will be crucial for summarising and tracking changes in trends.

When researching racial discrimination, perceived racial discrimination (PRD) is frequently used as the exposure variable. As racial discrimination cannot be experimentally manipulated, for ethical reasons and due to its subjective nature (Noh et al., 1999), PRD is

used to describe a participants' retrospective self-report of their experiences of racial discrimination (Paradies et al., 2015). Meta-analytic reviews have been used to summarise the evidence base exploring the association between PRD and health outcomes. This has allowed for an overview of the observed effects between PRD and health outcomes, an overview of the limitations across studies, and an overview of changes in the evidence base (Britt-Spells et al., 2018; Pearce et al., 2019). The meta-analysis completed by Paradies et al. (2015) provided an overview of the research literature exploring the impact of PRD across physical and mental health outcomes. Two-hundred and ninety-three primary studies, published between 1983 and 2013, were reviewed. Higher PRD was significantly associated with poorer mental health. A borderline significant effect was also found for worse physical health (r = .09; Brydges, 2019; Cohen, 1992). Paradies et al (2015) explored the influence of a range of moderators (e.g., publication year, age, and birth country) on any associations and compared the impact of these moderators to the findings of previous research.

Higher PRD was significantly associated with higher levels of anxiety and depression. Anxiety and depression are the two most prevalent mental health difficulties found cross-culturally (Antunes et al., 2018; Baxter et al., 2012; Lo et al., 2020), with similar prevalence rates and estimated contributions as worldwide health burdens (Baxter et al., 2012; Olatunji et al., 2007; Robinson et al., 2013; World Health Organization, 2017). However, Paradies et al (2015) reported a substantial discrepancy in the number of primary studies they found looking at anxiety versus those looking at depression. Whilst 109 primary studies were found to examine the relationship between PRD and depression, only 40 primary studies were found to focus on anxiety. Furthermore, subsequent meta-analyses reviewing the association between PRD and depression (Britt-Spells et al., 2018), as well as other mental health outcomes such as psychosis (Pearce et al., 2019), were since conducted. At the time of this review, an

updated meta-analysis exploring the relationship between PRD and anxiety had not been completed.

Completing a systematic review on the relationship between PRD and all the health outcomes examined by Paradies et al (2015) was beyond the scope of this review. Due to the contribution of anxiety to the global health burden, the disproportionately low number of studies identified by Paradies et al (2015), and the lack of subsequent meta-analyses, this review aims to provide an updated meta-analysis exploring the relationship between PRD and anxiety. Exploring the characteristics and quality of the evidence base will also be examined. The meta-analysis by Paradies et al. (2015) will be used as a comparative study, and thus the current search will be conducted from 2013 to date. To ensure this review can be best used as a continuation of Paradies et al. (2015), every effort will be made to replicate the search terms and search strategy from the 2015 review. The results from this review will then be compared to Paradies et al (2015). Any changes in the overall effect size, study quality, or the impact of moderators will be considered. Additional moderators will also be explored where possible.

Method

Identifying Primary Studies

Search of Electronic Databases

The aim of the current search was to obtain a comprehensive overview of the literature examining the relationship between PRD and anxiety since the review by Paradies et al (2015). An electronic systematic search of the online literature was carried out, in English, on the 14th June 2020. The search included both published and unpublished studies. ‘Primary studies’ will be used to refer to all papers used within this review. Whilst obtaining primary studies written in other languages would have been beneficial, reviewing articles not in English was beyond the scope of this review. The search completed by Paradies et al (2015)

identified primary studies up until the end of September 2013. Therefore, this search included primary studies from the beginning of October 2013 to 14th June 2020. PubMed, PsycInfo, Web of Science, Sociological Abstracts, Education Resources Information Centre (ERIC), Cumulative Index to Nursing and Allied Health Literature (CINAHL), and ProQuest (for dissertations and theses) were all searched. In the Paradies et al (2015) review, Social Work Abstracts, Academic Search Premier, and the authors' personal databases were also searched for additional references. Due to access restrictions, these databases could not be searched for this review. A full list of search terms and strategies can be found in Table 1. The search in the current review used the same search terms as Paradies et al (2015) with the addition of 'anxiety' being used as an overarching term to refine results for this more specific meta-analysis. Full inclusion and exclusion criteria used in the search refinement of primary studies are described in Table 2.

Table 1

Search Criteria

Construct	Free text search terms	Limits
1. Anxiety	anxi*/anxiety OR anxious OR anxieties (database dependent)	• October 2013 – June 2020 • English language • Journal articles, theses, dissertations, books and evaluation reports
	AND	
2. Health outcomes	birth* OR gestation* OR disease* OR BMI OR "body mass index" OR WHR OR "waist hip ratio" OR anthropometric* OR "blood pressure" OR hypertension OR cardiovascular OR overweight OR obes* OR diseases OR well-being OR wellbeing OR illness* OR depressi* OR anxi* OR distress OR stress OR suicid* OR sleep* OR ((social* OR behav* OR emotio* OR develop* OR psych*) AND (difficul* OR problem* OR delay* OR adjust*)) OR self-esteem OR "self esteem" OR "life satisfaction" OR "quality of life" OR resilien* OR alcohol OR tobacco OR smok* OR "substance use" OR drug* OR (health AND (care	

Construct	Free text search terms	Limits
	OR service* OR clinic*)) OR psychiatry OR psychology	
	AND	
3. Discrimination	discrim* OR bias OR prejud* OR hostil* OR harass* OR bully* OR "unfair treat*" OR oppress* OR prejudice	
	AND	
4. Racial/ethnic specific	rac* OR ethnic* OR cultur* OR religio* OR migra* OR immigra* OR refugee* OR "ethnic groups" OR "minority groups"	
	AND	
5. Type of study	longit* OR cohort* OR trial* OR "follow up" OR prospective OR retrospective OR "cross section*" OR "cross-section*" OR intervention* OR quantitative OR survey* OR "case-control" OR "case control" OR "randomised control* trial" OR "randomized control* trial" OR "before and after" OR "interrupted time series" OR questionnaire* OR registr* OR evaluat* OR audit* OR "longitudinal studies" OR "epidemiologic research design" OR "epidemiologic study characteristics" OR registries	

Note. Some of the limits were applied manually as not all databases allow automatic limitation.

Table 2

Inclusion and Exclusion Criteria.

Inclusion criteria	Justification
Exposure (PRD) *Experiences of PRD is the examined exposure. This includes self-reported PRD (i.e. being a direct target of racial discrimination, belonging to a group being targeted, or experiencing racial discrimination vicariously by witnessing someone else experiencing racial discrimination), proxy reports of PRD (i.e. a parent's report of their child's experiences) and internalised racial discrimination (i.e. individual acceptance and integration of racist beliefs). Studies where racial discrimination was evaluated by a*	Racial discrimination was considered a subjective experience unique to the individual. Therefore, participant PRD is the desired exposure as opposed to racial discrimination being evaluated by researchers who may perceive the experience differently and thus introduce confounding effects.

Inclusion criteria	Justification
researcher were excluded. Studies that used ecological and experimental exposure measures of racial discrimination rather than participant PRD were excluded.	
Exposure measurement tools *Measures of discrimination and related terms (see Table 1) where the reason for discrimination is due to someone's ethnicity and other related terms (see Table 1) were included. The exposure measure must specify the type of discrimination they are measuring. Measures of more general or other types of discrimination that do not allow for the isolation of PRD were excluded. In some cases, studies used a general measure of discrimination but either modified the measure to explicitly ask about PRD or participants were asked retrospectively to assign causation to their experiences of discrimination. These studies were included. If most of the questions in a measure asked about PRD specifically but the rest of the questions were about unspecified discrimination, these measures were also included. Studies measuring PRD as an outcome rather than an exposure were excluded. Studies solely measuring distress caused by PRD, rather than quantity of experiences, were also excluded.*	If an instrument did not specify it is measuring PRD, or if PRD cannot be isolated, this prevents the desired exposure from being accurately measured and risks the association between anxiety and other forms of discrimination being included. If a measure solely assesses the distress caused by PRD, the association between distress caused by PRD and anxiety could be explored but not the relationship between anxiety and PRD itself.
Outcome data *Studies were required to report quantitative data on the association between PRD and anxiety.*	To ensure that outcomes can be calculated into an effect size.
Type of article *The following article types were excluded: theoretical papers, commentaries, clinical guidance, conference papers, presentations, qualitative papers.*	These articles did not provide the data needed for analysis of quantitative data.
Participant characteristics *Participants must identify or be identified as belonging to a minority ethnic group.*	PRD was defined in this review as discrimination towards a marginalised or minority group.

Search Results

The results of the systematic search are presented in Figure 1. The search yielded a total of 2,384 primary studies which were imported into Zotero (Roy Rosenzweig Center for History and New Media, 2021). Duplicates were automatically and manually identified and removed. After duplicates were removed, 1,402 primary studies remained. The titles and abstracts of these primary studies were screened using the inclusion and exclusion criteria in Table 2. Following this, the full texts of the remaining 261 primary studies were screened against the inclusion and exclusion criteria. Nine full text primary studies could not be accessed due to access restrictions and were excluded. Seven primary studies that met the full inclusion and exclusion criteria but did not report sufficient or appropriate association data for the purposes of this review were excluded. Fifty-five primary studies met the full inclusion and exclusion criteria as well as access and data requirements. The most common reasons for excluding primary studies were anxiety not being a measured outcome, PRD not being a measured exposure (e.g. looking at other forms of discrimination exclusively or the type of discrimination was not specified such as when using the Everyday Discrimination Scale; Williams et al., 1997) or where the specific effects of PRD could not be isolated from other types of discrimination (e.g. discrimination based on ethnicity and sexual orientation).

Data Extraction

All data was extracted by the corresponding author. Pearson's correlation coefficient, *r*, was the desired effect for extraction. Pearson's correlation coefficient was used as it was the most reported effect and was the effect extracted in the 2015 review. The Pearson's correlation coefficient is commonly used in meta-analyses as it captures both the strength and direction of an association between two quantitative variables (Chee, 2015). To aid analysis, effect sizes that were not reported as Pearson's correlation coefficients were converted to *r*

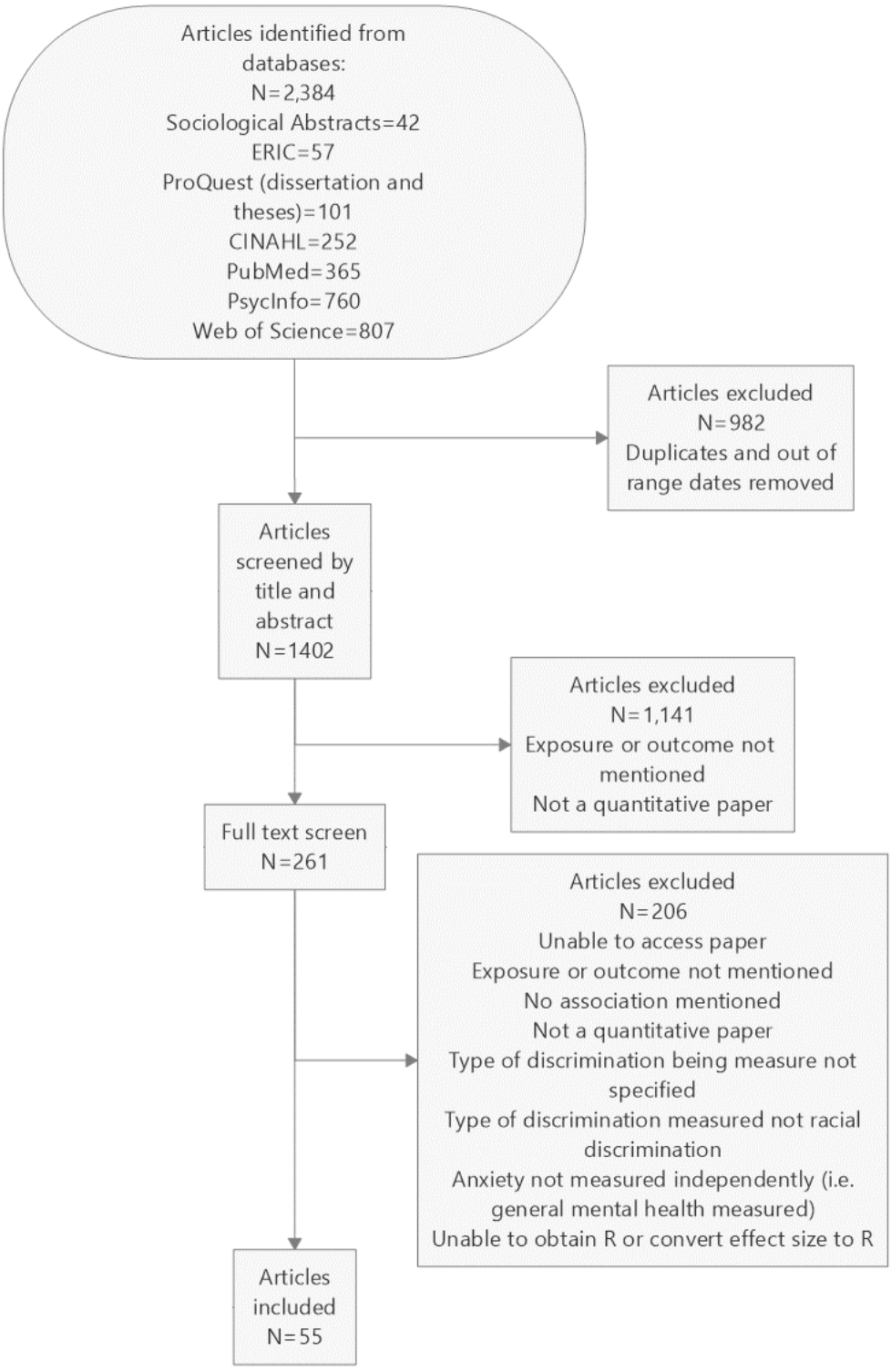
using an RStudio programme (Jones, 2020; RStudio Team, 2020). Other types of effects reported in the primary studies that required conversion included odds ratios (n=5), regression coefficients (n=3), and chi-square (n=2). If primary studies reported multiple measures of anxiety or reported the anxiety measure across multiple subgroups (i.e., effects reported separately for different genders), consideration was given to the appropriateness of including or excluding the additional effects. The inclusion of multiple reporting from the same primary articles can cause an overestimation of sample size (Lunny et al., 2021). To overcome this, multiple outcomes can be combined into a single outcome using the procedures described by Borenstein (2009). This procedure entails converting multiple correlation coefficients into *z*-scores, taking a weighted average, and then converting back into one correlation coefficient. This procedure was completed for thirteen primary studies. Cooke et al (2014) and Hope et al (2018) both reported multiple effects but were not subjected to the above procedure. It was not considered appropriate to combine the multiple effects due to size differences between effects. In Cooke et al (2014), the two subgroups 'African American' and 'Multiracial' were not combined as one subgroup had a small effect (r =.14) whilst the other had a large effect size (r =.50). In Hope et al (2018) 'Black' and 'Latinx' were also not combined as one subgroup reported no association (r=0) and the other had a small effect (r =.14). The reporting of multiple effects for these two primary studies meant that although 55 primary studies were identified, 57 effects were included in the analysis.

Heterogeneity of Variance

Heterogeneity describes the variation between study effects that is not due to chance (Ruppar, 2020). High heterogeneity of variance suggests that there are systematic differences between studies, rather than differences being due to chance alone. This reduces confidence

Figure 1

Results of The Systematic Search and the Application of the Inclusion/Exclusion Criteria



that the results are generalisable (Higgins et al., 2003; Sedgwick, 2015). Heterogeneity can result from the methodological variation in the primary studies (e.g., measurement error) or uncontrolled individual differences (e.g., age). After reviewing each primary study, considerable variation was suspected and thus it was felt to be important to assess heterogeneity as part of this review. Higgins I^2 is a commonly used measure of heterogeneity which represents the percentage of total variation caused by heterogeneity rather than by chance alone. Higgins et al (2003) suggests that an I^2 value of 25% indicates low heterogeneity, 50% moderate heterogeneity and 75% high heterogeneity. If high levels of heterogeneity are observed, the focus of subsequent analyses will be to identify potential sources of heterogeneity between primary studies.

Risk of Bias Assessment

Risk of methodological bias was assessed to consider its contribution to any observed heterogeneity. A set of quality criteria was developed to assess the risk of bias within the literature. The quality criteria used in this review was adapted from existing risk of bias frameworks, including The Cochrane Collaboration Risk of Bias Tool (Higgins et al., 2011) and the Risk of Bias Assessment Tool for Nonrandomised Studies (Kim et al., 2013), tailored to best assess bias across this literature. Risk of bias was assessed for each individual article across six domains: detection bias, generalisation, performance bias, reporting bias, selection bias, and statistical bias (Table 3). The criteria in Table 3 were used to rate each primary study against the six domains (Figure 2). A rating of ‘low risk’ was associated with a quality score of 2, ‘unclear risk’ with a quality score of 1 and ‘high risk’ with a quality score of 0. A quality index was calculated by summing the scores of all six domains and expressing this as a percentage. The highest potential index score was 12. The application of the risks of bias described in Table 2 to the individual studies is presented in Figure 2.

Table 3

Domains of Risk of Bias and the Criteria for Ratings of Low, Unclear, or High Risk

Domain	Details	Risk of bias
Detection bias	Were the measures administered in their original format? If measures were translated, was the process described? What were the measures' Cronbach's alpha levels? Were the measures clearly defined, valid, and reliably? Were the measures implemented consistently across participants?	High Risk – The measures were not administered in their original or agreed format (i.e. individual subscales of a measure were applied or analysed separately) or were implemented differently across participants. Cronbach's alpha <0.6. The outcome measures had poor validity and reliability. Outcome measures are translated but there is no detail on how this was conducted or there were problems in the translation. Unclear Risk – It is not clear whether the measures were administered in their original or agreed format. Cronbach's alpha >0.6 but <0.7. The outcome measures and their implementation are not clearly described. Outcome measures are translated with details given on how this was conducted. Low Risk – Measures were administered in their original or agreed format. Cronbach's alpha >0.7. The outcome measures are clearly defined, valid and reliable, and are implemented consistently across all participants.
Generalisation	Can the research findings be applied to settings other than that in which they were originally tested? Did the sample have idiosyncratic features? Was a power analysis provided and sample size justification and estimate given? Was the sample size large enough to detect an effect?	High Risk – Small sample with or without idiosyncratic features (<20 per group). The sample size is not adequate to detect an effect. Unclear Risk – Sufficient sample for generalisation but with some idiosyncratic features (>20 per group). Sample taken from only one population group (i.e. students) with attempts to generalise to entire population. A sample size justification, estimate and power analysis were not provided Low Risk – Sufficient sample for generalisation and representative of target population (>20 per group). A sample size justification, estimate and power analysis were provided. The sample size is adequate to detect an effect.

Performance bias	Were the levels of confidentiality and anonymity outlined clearly? Were participants rewarded for their participation? Was information given and procedures provided in a way that differentially motived participants?	High Risk – Responses are not confidential or anonymous. Participants were rewarded for their participation in the study. Participants were told which condition/what questionnaires they were completing and why and any proposed hypotheses. Systematic differences between groups in exposure to factors i.e. participants did not attend the same number of sessions Unclear Risk – The study does not report levels of confidentiality and anonymity. It is not clear if participants were rewarded for their participation (e.g. motivation to respond in a certain way). It is unclear how much information was provided to the participant prior to taking part in the study. Differences in exposure to factors that may affect performance i.e. different type of setting or different facilitators Low Risk – Study reports level of confidentiality and anonymity. Participants were not rewarded for their participation in the study. Information and procedures are provided in a way that does not differentially motivate participants.
Reporting bias	Is there evidence of selective outcome reporting? i.e. only significant results reported. Are there measures that have not been reported in the results that have been mentioned in the method? Are Cronbach's alphas reported?	High Risk – Not reported full outcome measures that are stated in the method section/reported only a subsample of results/only significant results. Cronbach's alpha not reported. Unclear Risk – Not all descriptive and/or summary statistics are presented. Low Risk – Reported all results of measures as outlined in the method. Cronbach's alpha reported for all measures.
Selection bias	Was the study population described and reported fully? What was the non-response rate? Was the sampling and recruitment method described? If so, was convenience or target sampling used?	High Risk – The source population is not described or reported. There is an unacceptable level of non-response rate, <30%. The recruitment method is not reported. Target sampling was used. Unclear Risk – The source population is not clearly described or fully reported. Non-response rate is not reported or <50% but >30%. The recruitment method is not clearly reported. Convenience sampling without additional bias. Low Risk – The source population is well described, and characteristics are clearly described. Non-response rate is reported and >50%. The recruitment method is clearly reported and well defined.
Statistical bias	Have appropriate statistical methods been used? Is there incomplete data due to attrition? Have Pearson's correlation coefficients, confidence intervals and *p*-values been reported?	High Risk – Statistics were not reported. Wrong statistical test was used and not appropriate for the study design. More than 50% data attrition loss. Unclear Risk – Unclear what statistical test was used. Attrition rate not reported. Confidence intervals or exact *p*-values not reported. Low Risk – Appropriate statistical methods used. An *r* value or another statistic which could be transformed into a statistical equivalent was reported. Less than 50% data loss attrition. Confidence intervals or exact *p*-values were given.

Figure 2

Ratings of Risk of Bias

Study Name	Statistical.Bias	Reporting.Bias	Detection.Bias	Selection.Bias	Performance.Bias	Generalisability	Quality.Index
Allen and Conklin 2017	Low risk	Low risk	Unclear risk	Unclear risk	High risk	Unclear risk	58%
Assari et al. 2017	Low risk	Unclear risk	Unclear risk	High risk	Low risk	Unclear risk	58%
Atwal and Wang 2018	Unclear risk	Low risk	High risk	Unclear risk	High risk	Unclear risk	42%
Bauermeister et al. 2018	Unclear risk	Low risk	Unclear risk	Unclear risk	High risk	Unclear risk	50%
Borders and Hennebry 2015	Unclear risk	Low risk	High risk	Unclear risk	High risk	Unclear risk	42%
Budhwani et al. 2015	Low risk	Unclear risk	High risk	High risk	Unclear risk	Unclear risk	42%
Cano et al. 2016	Unclear risk	Low risk	Unclear risk	Unclear risk	Low risk	Unclear risk	67%
Cariello et al. 2019	Low risk	Low risk	Unclear risk	Unclear risk	High risk	Low risk	67%
Carter et al. 2016	Unclear risk	Low risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	58%
Celebi et al. 2017	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	42%
Chae et al.2016	Low risk	Low risk	Unclear risk	High risk	High risk	Unclear risk	50%
Choi et al. 2013	Unclear risk	Unclear risk	Unclear risk	Unclear risk	High risk	Unclear risk	42%
Cooke et al. 2014 (African American)	Unclear risk	Low risk	Low risk	Unclear risk	Unclear risk	Unclear risk	67%
Cooke et al. 2014 (Multiracial)	Unclear risk	Low risk	Low risk	Unclear risk	Unclear risk	Unclear risk	67%
Corona et al. 2017	Unclear risk	Low risk	Low risk	Unclear risk	High risk	Unclear risk	58%
Doud et al. 2018	Low risk	Low risk	High risk	Low risk	Unclear risk	Unclear risk	67%
Doyle. 2019	Unclear risk	High risk	Unclear risk	Unclear risk	High risk	Unclear risk	33%
Drakeford. 2017	Low risk	Low risk	Unclear risk	Unclear risk	High risk	Low risk	67%
Durkee and Williams 2015	Unclear risk	Low risk	High risk	High risk	Unclear risk	Unclear risk	42%
Escovar et al. 2018	Low risk	Low risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	67%
Fang et al. 2016	Low risk	Low risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	67%
Franco and Carter 2019	Unclear risk	Low risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	58%
Graham et al. 2015a	Unclear risk	Low risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	58%
Graham et al. 2015b	Unclear risk	Low risk	High risk	Unclear risk	Unclear risk	Low risk	58%
Hope et al. 2018 (Black)	Unclear risk	Low risk	High risk	Unclear risk	High risk	Low risk	50%
Hope et al. 2018 (Lantinx)	Unclear risk	Low risk	High risk	Unclear risk	High risk	Low risk	50%
Hurd et al. 2014	Unclear risk	Low risk	Unclear risk	High risk	Unclear risk	Unclear risk	50%
Jackson 2015	Unclear risk	High risk	High risk	Unclear risk	Unclear risk	Low risk	42%
Jerald et al. 2017	Unclear risk	High risk	High risk	Unclear risk	High risk	Unclear risk	25%
Juang et al. 2016	Unclear risk	Low risk	High risk	High risk	Unclear risk	Unclear risk	42%
Kim 2014	Unclear risk	Low risk	High risk	Unclear risk	High risk	Low risk	50%
Kwon and Han 2019	Unclear risk	Low risk	Unclear risk	High risk	Unclear risk	Unclear risk	50%
Lee and Thai 2015	Unclear risk	Low risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	58%
Marin and Mancini 2017	Low risk	High risk	Unclear risk	High risk	Unclear risk	Unclear risk	42%
Martinez et al. 2020	Unclear risk	Low risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	58%
Mata-Greves and Torres 2018	Unclear risk	Low risk	High risk	Unclear risk	High risk	Unclear risk	42%
Mouzon et al. 2017	Low risk	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	50%
Neblett et al. 2016	Unclear risk	Low risk	High risk	Unclear risk	Unclear risk	Unclear risk	50%
Park et al. 2017	Unclear risk	Low risk	High risk	Unclear risk	High risk	Unclear risk	42%
Perreria et al. 2015	Low risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	58%
Preciado and D'Anna-Hernandez 2017	Unclear risk	Low risk	High risk	Unclear risk	High risk	Unclear risk	42%
Rodriguez-Seijas et al. 2015	Unclear risk	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	42%
Schwartz et al. 2018	Unclear risk	Low risk	High risk	Unclear risk	High risk	Low risk	50%
Serafini et al. 2017	Unclear risk	Unclear risk	High risk	Unclear risk	Unclear risk	Unclear risk	42%
Sosoo et al. 2019	Unclear risk	Low risk	High risk	Unclear risk	High risk	Unclear risk	42%
Sutter and Perrin 2016	Low risk	Unclear risk	High risk	Unclear risk	High risk	Unclear risk	42%
Swann et al. 2020	Unclear risk	Low risk	Unclear risk	Unclear risk	High risk	Unclear risk	50%
Takeda 2018	Unclear risk	Low risk	High risk	Low risk	High risk	Unclear risk	50%
Taylor and Ruiz 2018	Unclear risk	Low risk	High risk	Unclear risk	High risk	Unclear risk	42%
Thelemaque 2018	Unclear risk	Low risk	High risk	Unclear risk	High risk	Unclear risk	42%
Tuason et al. 2014	Unclear risk	Low risk	Unclear risk	Unclear risk	Unclear risk	Unclear risk	58%
Victoria 2014	Low risk	Low risk	Low risk	Low risk	Unclear risk	Low risk	92%
Walker et al. 2017	Low risk	Low risk	High risk	Unclear risk	High risk	Unclear risk	50%
Weeks and Sullivan 2019	Low risk	Unclear risk	Unclear risk	High risk	High risk	Unclear risk	42%
Wei et al. 2015	Unclear risk	Unclear risk	Low risk	High risk	Unclear risk	Unclear risk	50%
Whitfield 2019	Unclear risk	Low risk	Low risk	Low risk	Low risk	Unclear risk	83%
Williams et al. 2018	Unclear risk	Low risk	High risk	Unclear risk	Unclear risk	Unclear risk	50%

Statistical Bias. Statistical bias was most commonly rated as ‘unclear’. Sixteen papers were rated as low risk, forty-one were rated as unclear, and twelve were rated high risk. Primary studies rated as low risk generally reported exact test statistics, *p*-values, and attrition rates. All the unclear primary studies failed to report exact *p*-values and attrition rates.

Reporting Bias. Reporting bias across the primary studies was generally low risk. Forty-two primary studies were rated as low risk, eleven were rated as having unclear risk and four were rated as high risk. Low risk primary studies reported all measures described in the methods and reported all Cronbach’s alphas where appropriate. The studies rated ‘unclear’ did not report Cronbach’s alphas for some of the measures used. Primary studies that were rated as high risk did not report the Cronbach’s alphas for any of the measures used.

Detection Bias. Overall, detection bias was mixed in its level of risk, with most primary studies being rated as high or unclear risk. Six primary studies were rated as low risk, twenty-five primary studies were rated as unclear risk and twenty-six primary studies were rated as being high risk. The low-risk primary studies described the validity and/or reliability of the measures they used and two also clearly described the administration of the measures so that consistency across participants could be adhered to. The primary studies rated as having unclear risk met a variety of factors in this category. For example, the validity or reliability of their measures were unclearly defined or reported, measures were translated (although a description of how this was completed was included), it was unclear whether measures were reported in their original format, and it was unclear if there were systematic differences in the administration of the measures. High risk primary studies either only used some subscales of a measure, edited, or created measures, implemented measures differently across participants, or translated measures without providing sufficient information on how this was completed.

Selection Bias. Selection bias was mostly rated ‘unclear’ within the primary studies. Forty-one primary studies were rated as having an unclear risk of bias with four rated as low risk of bias and twelve rated as high risk of bias. The low-risk primary studies clearly reported their recruitment procedure and participant demographics. The unclear primary studies were considered vague in their reporting of sampling methods or demographics, or they used convenience sampling. The high-risk primary studies did not report sufficient demographics or recruitment procedures. In some primary studies, this was because they were using data originally part of another study and they signposted the reader to the original study instead of detailing it themselves. In addition, two primary studies had a response rate of less than 30% and one study used targeted sampling.

Performance Bias. Overall, performance bias was mostly rated ‘unclear’ and high risk across the primary studies. Three primary studies were rated as low risk, twenty-eight primary studies were rated as unclear risk and twenty-six primary studies were rated as high risk. The low-risk primary studies explicitly detailed confidentiality and anonymity and did not use rewards in their recruitment. The unclear risk primary studies did not clearly discuss anonymity, confidentiality, or whether participants were given rewards. The high-risk primary studies all rewarded participants for their participation.

Generalisability. Generalisability was most rated ‘unclear’ within the primary studies. Nine primary studies were rated as having low risk and forty-eight primary studies were rated as unclear risk. Low risk primary studies all discussed sample size and power or included a power analysis. Unclear risk primary studies did not discuss statistical power in the context of sample size. Whilst all primary studies had a sufficient sample size, some primary studies used idiosyncratic populations, for example, student or refugee samples.

Summary. Overall, the level of bias varied across the primary studies included in this meta-analysis, with quality index scores ranging between 21% and 79%. The study with the lowest rated quality index score had three out of six areas rated as high risk, with the other three rated as unclear (Jerald et al., 2017). The study with the highest rated quality index score had five out of six areas rated as low risk and one area rated as unclear (Victoria, 2014). High risk of bias was frequently seen across both detection and performance bias with 46% of the primary studies being rated as high risk for both areas. All primary studies had at least one domain rated as unclear risk and 79% of primary studies had at least one domain rated as high risk. As such a large proportion of primary studies were rated as having unclear or high risk across at least one of their domains, it was not feasible to remove primary studies with unclear or high risk of bias ratings. It is also argued that if such a large proportion of primary studies were rated as having unclear/high risk of bias, then this represents the current evidence base. Removing primary studies would therefore risk analysing an unrepresentative sample of the current literature base. As the decision was made to include primary studies with unclear and high-risk of bias, caution should be exercised during interpretation.

Results

After screening was completed, 55 primary studies reporting 57 effects across a total of 116,664 participants remained. Both study-level moderators (Table 4) and participant-level moderators (Table 5) were collated. A Fisher's r-to-z_r transformation was used to correct the bias within Pearson's correlation coefficients (Corey et al., 1998) by transforming the curved distribution of r into a distribution that was approximately linear (Fisher, 1922). Whereas the variance of the sampling distribution of r depends on the size of the correlation (i.e., as the correlation approaches the extreme values then variance increases or decreased) the variance of the transformed z_r distribution is independent of the correlation size.

Table 4

Study Moderators

Variable	Groups	Number of effects	% of studies
Total number of effects		57	100%
Sample size	18-100	8	14%
	100-200	14	25%
	201-300	11	19%
	300-1,000	17	30%
	1,000-60,664	7	12%
Year of publication	September 2013-2015	17	30%
	2016-2017	19	33%
	2018-2019	19	33%
	2020	2	4%
Type of publication	Academic Journal	45	79%
	Dissertation/Thesis	12	21%
Country	Israel	1	2%
	Italy	1	2%
	Turkey	1	2%
	USA	54	95%
Sampling procedure	Non-representative	47	82%
	Representative	5	9%
	Other/Not reported	5	9%
Data type	Cross-sectional	44	77%
	Longitudinal	6	11%
	Other	7	12%
Exposure instrument name	Schedule of Racist Events	5	9%
	Racism and Life Experiences Scales	3	5%
	Perceived Ethnic Discrimination Questionnaire	6	11%
	Subtle and Blatant Racism Scale	3	5%
	Multiple measures used	3	5%
	Other	28	49%
	Not recorded	1	2%
	Created	8	14%
Exposure number of items	8 or less	14	25%
	9 or more	39	68%
	Not recorded	4	7%

Variable	Groups	Number of effects	% of studies
Exposure Cronbach's alpha	0.79 or lower	8	14%
	0.80 or higher	37	65%
	Other/Not reported	12	21%
Exposure type	Direct	42	74%
	Indirect	1	2%
	Internalised	2	4%
	Mixed	6	11%
	Not recorded	6	11%
Timeframe of PRD exposure	Last six months	2	4%
	Last 12 months	15	26%
	Last > 12 months and < 5 years	0	0%
	More than 5 years – lifetime	14	25%
	Not specified	26	46%
Anxiety type	Bodily Symptoms	1	2%
	GAD	9	16%
	Mixed	2	4%
	Non-specific anxiety	43	75%
	Social	1	2%
	Other	1	2%

Table 5

Participant Characteristics

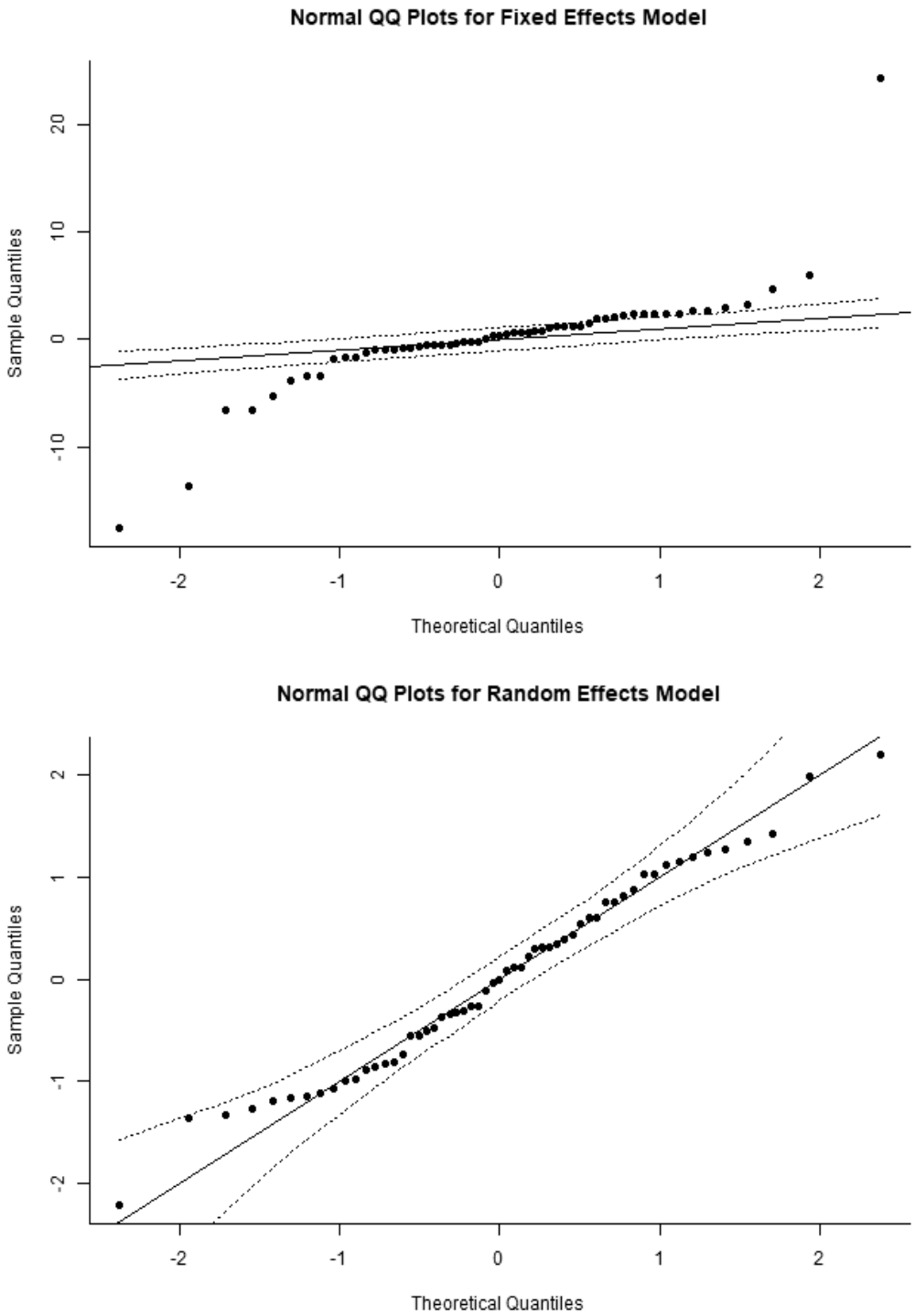
Characteristic	No. of studies	Characteristics	No. of participants	%
Age	46	Under 18 only	61,952	57%
		18 and over only	46,479	43%
Sex	45	Female	47,096	52%
		Male	43,509	48%
		Other	127	0.1%
US racial/ethnic groups	53	African American	3,859	7%
		American Indian	43	0.08%
		Arab	405	1%
		Asian American	1,291	2%
		Asian Pacific Islander	442	1%
		Black American	7,182	13%
		Latinx	20,390	37%
		Multiple ethnicities	19,677	36%
		Multiracial	30	0.05%
		Polynesian	628	1%
		Romanian	115	0.21%
		Syrian	361	1%
		International Students (in America)	805	1%
Birth country	22	Foreign Born	20,267	53%
		Locally Born	17,946	47%
Current education (under 18 samples)	3	Primary Education	40,102	66%
		Secondary Education	20,804	34%
Education completed (18 and over samples)	22	Completed less than high school	4,049	4%
		Completed high school	6,255	6%
		Completed more than high school	95,645	90%

Reported associations between PRD and anxiety ranged between $r = 0$ and .50 ($z_r = 0$ and .55). When considering how to model the association between PRD and anxiety, two models were considered. The fixed effects model assumes that the effect can be described as a single value and that all the primary studies employ a methodology of equal power to detect that effect. The weighting of the effects is directly proportional to the sample size of the study. Therefore, the estimation of the effect is dependent on the specific studies from which the average is derived. This makes it difficult to generalise and would allow only 'conditional inference' (Hedges & Vevea, 1998). The random effects model assumes that effect size reported by a particular study is one of a distribution of possible true effects, and the purpose of the meta-analysis is to model this distribution of these true effects. The random effects model also assumes that the distribution of the effects may be influenced by factors other than sample size (e.g., different methodologies, different measurement tools, systematic individual differences in response to PRD). Thus, the random effects model is better suited to cases where the primary studies might be influenced by multiple factors, where the effects reported in the primary studies are derived from a mixture of methodologies, and where the purpose of the meta-analysis is to generalise beyond the observed studies.

The distribution of primary study effects is shown in Figure 3. The Quantile-Quantile (Q-Q) plots show the observed distribution compared to the normal distribution indicated by the continuous straight line. Q-Q plots are presented for both the fixed effects model and the random effects model. The DerSimonian and Laird (1986) method was used to calculate 'between study' variation. Figure 3 shows the fixed effects model as having marked non-normality. As such, the data was considered to best fit the random effects model. The random effects model was considered the most appropriate method for the calculation of the variation of the true effect.

Figure 3

Q-Q Plots of the Distribution of Correlations within the Primary Studies



Note. The top QQ plot shows the fixed effects model and the bottom QQ plot shows the random effects model.

The omnibus test

A forest plot in Figure 4 shows the amount of variation between each individual study and estimates the averaged overall effect of the primary studies combined (Lewis & Clarke, 2001). The random effects model suggested an overall weighted average correlation of *r* =.2564 ($z = 15.21$, $p < 0.0001$) and a 95% confidence interval between 0.2246 to 0.2877. Overall, this indicates a small but significant association between PRD and anxiety as measured and reported in the current primary studies. A high level of heterogeneity between the primary studies was observed (Higgin's $I^2 = 94\%$, $\text{tau}^2 = 0.0120$, $Q = 1017.29$, $p < 0.0001$). This suggests that the estimates of the correlation between PRD and anxiety may be biased by the presence of uncontrolled confounding factors. The focus of subsequent analyses will be to try and identify the sources of heterogeneity between the correlations reported in the primary studies.

The Impact of Influential Primary Studies

The impact of excessively influential primary studies was assessed using a "leave-one-out" analysis which calculates the random effects model with each of the primary studies removed in turn. The change in the weighted average effect size (i.e., influence) and the change in heterogeneity (i.e., discrepancy) was recorded. The results of this analysis are presented on the Baujat plot (Baujat et al., 2002; Figure 5). Mouzon et al (2017), Schwartz et al (2018), and Takeda (2018) were identified as both influential and discrepant from the rest of the primary studies. These studies were re-reviewed to consider their removal if concerns over their appropriateness were identified. No concerns were found. The random effects model was also recalculated with these primary studies removed. The adjusted random effects model showed a very small change, a reduction of 1.8% (r=0.2517, 95% CI 0.2197 – 0.2831). Accordingly, it was decided that no primary studies would be removed based on this analysis.

Figure 4

Forest Plot of Correlations Between Perceived Racial Discrimination and Anxiety

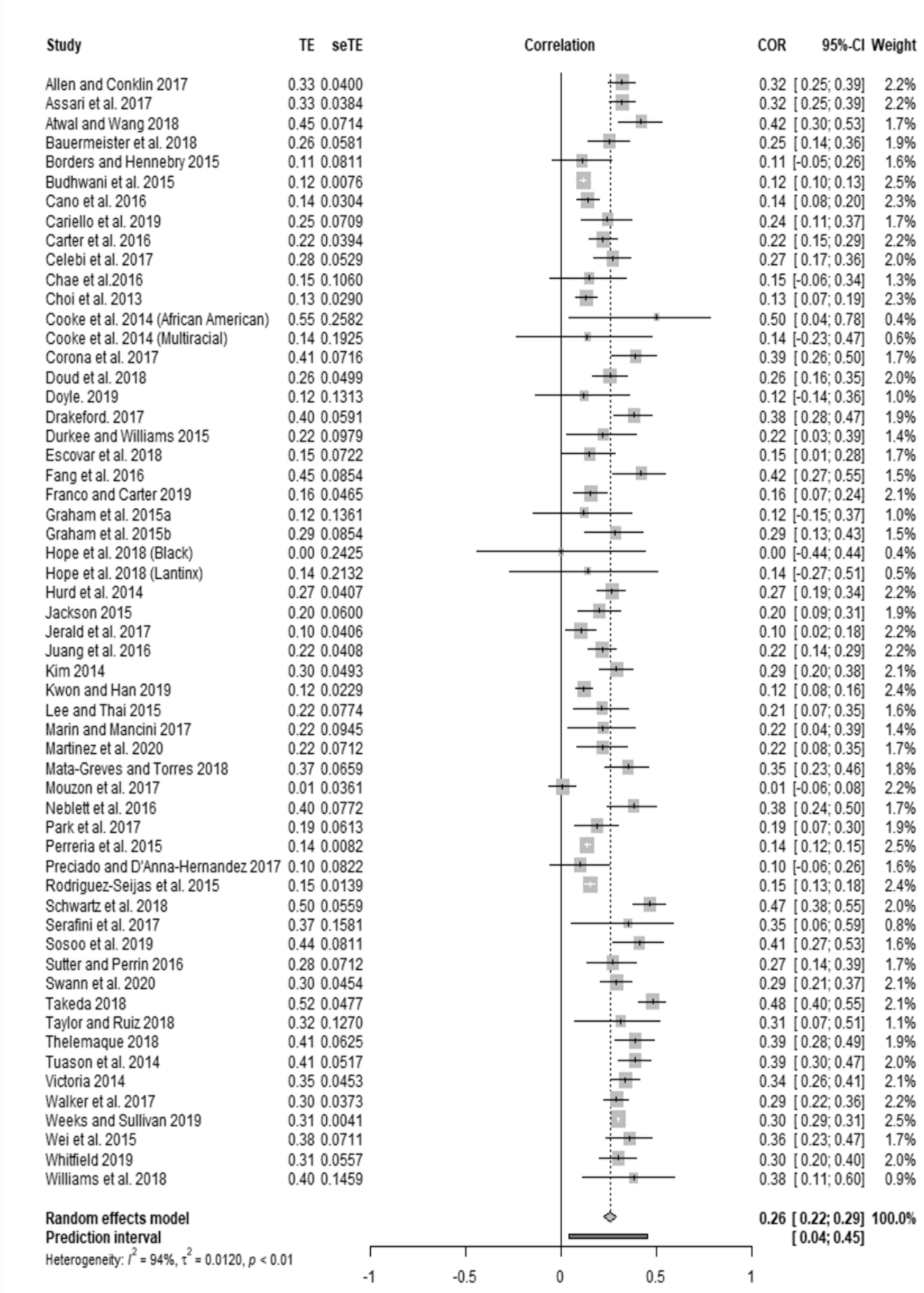
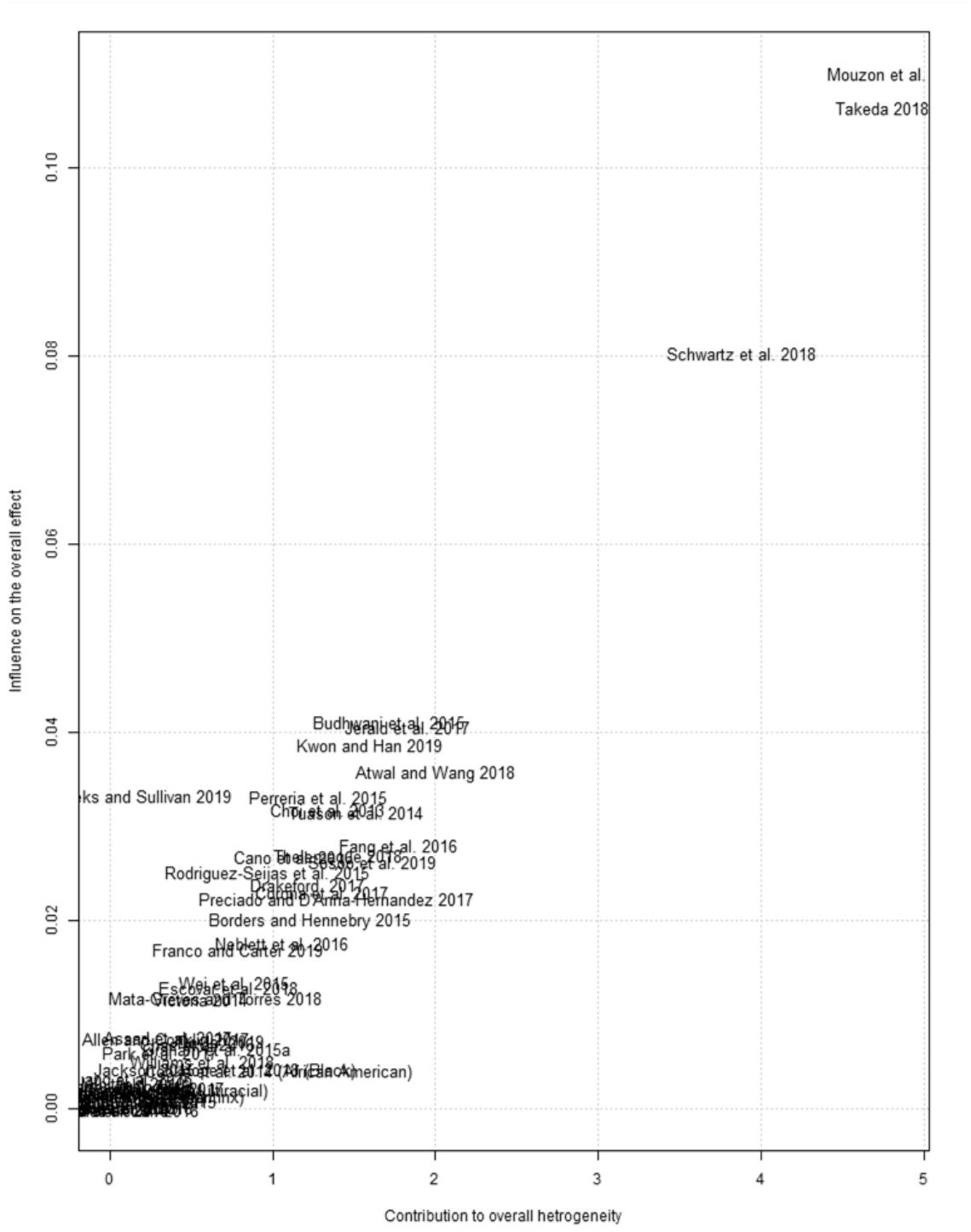


Figure 5

Baujat Diagnostic Plot of Sources of Heterogeneity



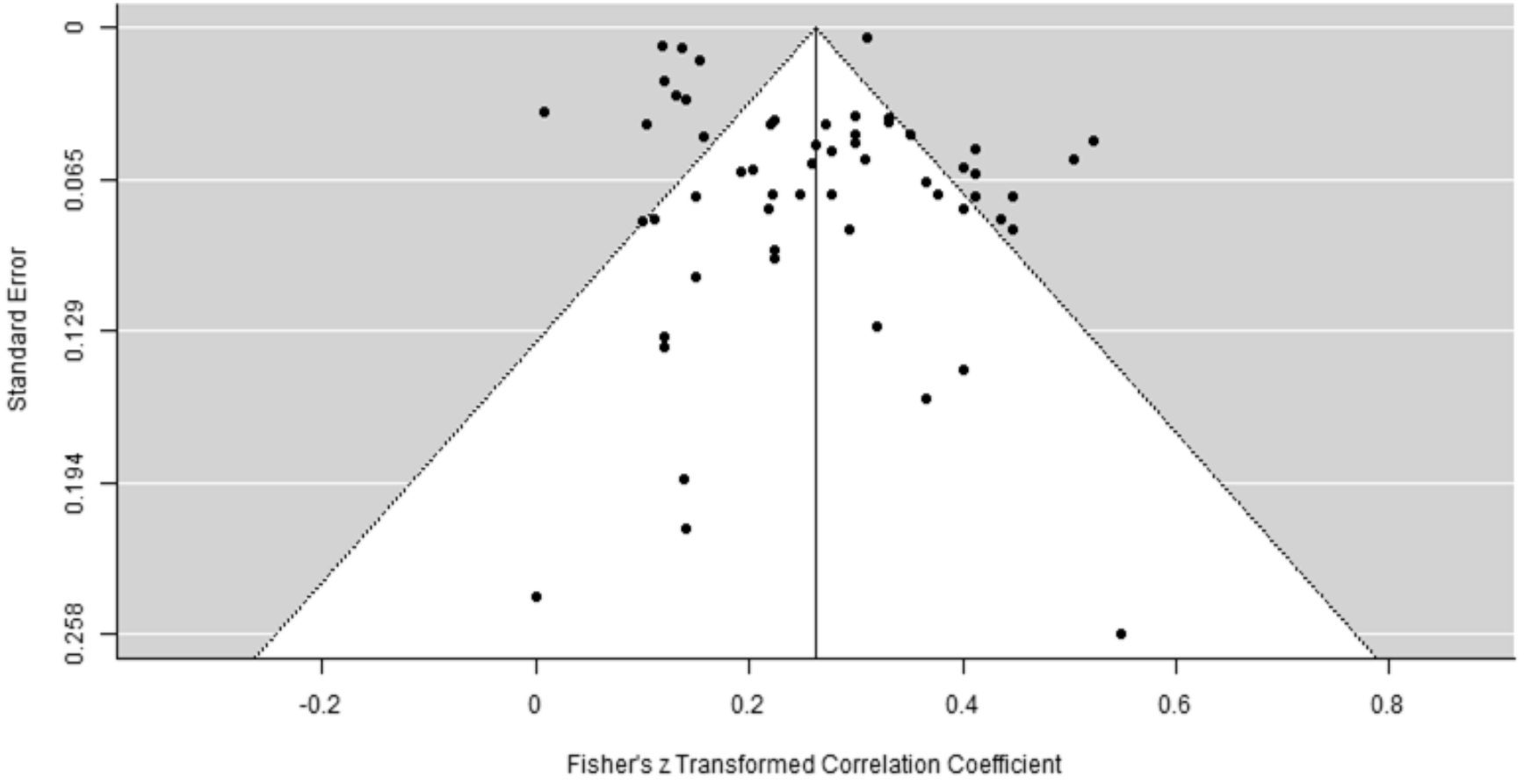
Note. The vertical axis reports the influence of the article on the overall effect and the horizontal axis reports the discrepancy of the article within the rest of the literature.

The Impact of Publication and Small Study Biases

Publication bias is caused by the tendency for papers with statistically significant results to be published whilst papers with non-significant results are less likely to be published (Murad et al., 2018). Small study bias is the tendency for primary studies with smaller sample sizes to show greater variability in their measurement of correlations. These biases can be identified using a funnel plot (Begg & Mazumdar, 1994). A funnel plot is a scatterplot of the magnitudes of each study's correlation coefficient against standard error. Studies with smaller sample sizes typically have larger standard errors as they are less precise at estimating effects. As study sample size increases, therefore, standard error typically decreases. The vertical axis of a funnel plot is inverted. As such, it is the expected that studies with smaller sample sizes, and larger standard errors, will widely scatter at the bottom whilst studies with larger sample sizes, and smaller standard errors, will more narrowly cluster at the top, creating a funnel shape. An asymmetrical funnel plot suggests the presence of publication bias (Sedgwick et al., 2013). Furthermore, an absence of plots in the area associated with small sample sizes and non-significant results (the lower left-hand corner) also suggests there is some publication bias leading to an overestimation of the true effect. The funnel plot for the current study is presented in Figure 6. The vertical line in the middle of the plot represents the overall effect (r = .26). Visual inspection does not strongly indicate the presence of publication bias as primary studies are plotted symmetrically on either side of the overall effect including towards the lower left corner. However, visual interpretation alone can be considered too subjective (Terrin et al., 2005). Therefore Eggers's test of asymmetry (Egger et al., 1997), a statistical test of publication bias, was also used. The Egger's test was non-significant (t=-0.03443, p=0.9727).As such, no adjustment for publication bias and small study effects was undertaken.

Figure 6

Funnel Plot of the Correlation.



Note. The 95% confidence interval of the expected distribution of correlations is shown as an inverted “funnel”.

The Impact of Study-Level Risk Of Bias

To assess the contribution of study level risk of bias upon heterogeneity, a series of subgroup analyses were conducted. The risk of bias for ratings of unclear risk and high risk of bias were combined and formed a category called 'any risk'. For each of the six areas of bias, studies with any risk were compared to those with low risk. Table 6 summarises the results of this subgroup analysis.

For reporting bias, primary studies that were rated as having low risk of bias had significantly higher correlation estimates (r = .28) than primary studies rated with any risk (r = .20; Figure 7). Higgins I^2 values suggest that primary studies rated as low risk (I^2=78%) were less heterogeneous that primary studies rated with 'any risk' (I^2 = 98%).

Table 6

Subgroup Analyses Results Comparing the Effects Across Low Risk and Any Risk Studies

Area of bias	Low risk			Any risk			X^2	p
	Correlation	95% CI	k	Correlation	95% CI	k		
Statistical bias	0.25	0.18 to 0.30	16	0.26	0.22 to 0.30	41	0.20	0.66
Reporting bias	0.28	0.25 to 0.32	42	0.20	0.13 to 0.26	15	5.26	0.02
Detection bias	0.34	0.29 to 0.39	6	0.25	0.22 to 0.28	51	8.44	<0.01
Selection bias	0.35	0.24 to 0.44	4	0.25	0.22 to 0.28	53	3.27	0.07
Performance bias	0.25	0.12 to 0.37	3	0.26	0.22 to 0.29	54	0.01	0.94
Generalisability bias	0.31	0.24 to 0.37	9	0.25	0.21 to 0.28	48	2.49	0.11

Note. k denotes the number of studies.

Similarly, for detection bias, primary studies rated as having low risk of bias had significantly higher correlation estimates (r = .34) than primary studies rated with any risk (r = .25; Figure 8). The Higgins I^2 values suggest that primary studies rated as low risk (I^2=0%) were far less heterogeneous than primary studies rated with 'any risk' (I^2 = 95%).

The other four areas of bias did not reveal any significant differences. As primary studies at risk of detection and reporting bias show significantly smaller effect sizes, it is likely that the true overall effect is underrepresented to some degree.

Figure 7

Reporting Bias Forest Plot

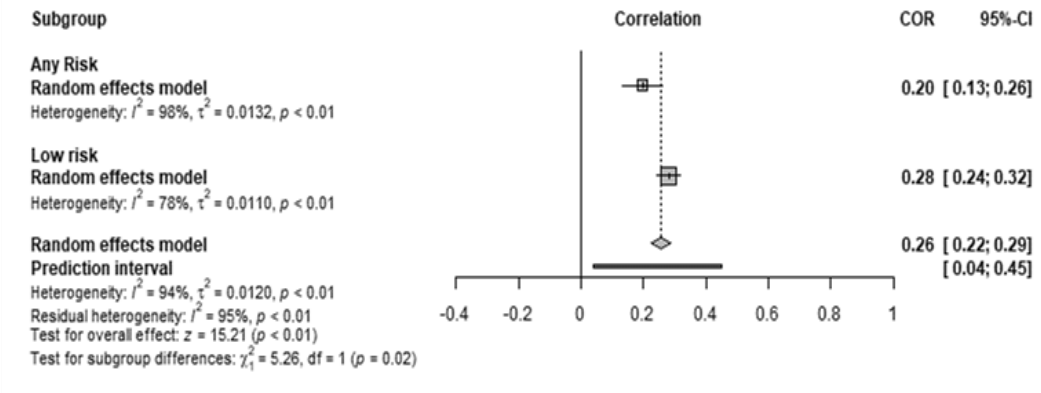
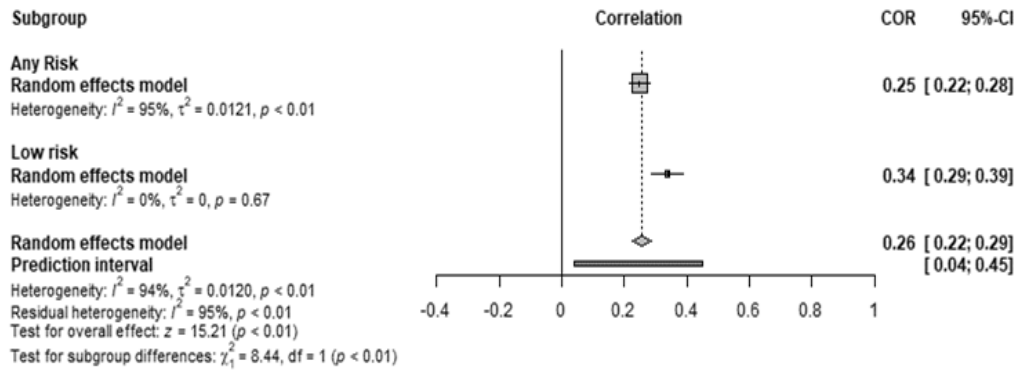


Figure 8

Detection Bias Forest Plot



Analysis of Moderators

Both study and participant level moderators were explored using sub-group analysis and meta-regressions. Study-level moderators included, publication year, publication status, country, study design, sampling procedure, exposure measures, exposure timeframe, exposure type, exposure number of items and exposure instrument internal reliability, anxiety type, anxiety measures, anxiety measure Cronbach's alpha and original reporting type of the effect. Participant-level moderators included participant age, sex, ethnicity, level of education and birthplace.

Analysis of Study Level Moderators

No significant effects for the publication year (p = .16), publication status (p = .12), study country (p = .99), study design (p = .81), sampling procedure (p = .40), or exposure timeframe (p = .17). Imbalances in the numbers of primary studies across groups meant some sub-groups could not be analysed (e.g., experimental (n = 1) and non-experimental designs (n = 56). This is likely to have increased the risk of type II error for detecting any truly significant differences in effect sizes.

Primary studies that originally reported r had significantly higher correlations (r = .27) and lower heterogeneity (I^2 = 66%) than studies originally reporting a different type of effect that was converted to r (r = .17, I^2 = 99%, p = .036). Full details are reported in Table 7 and Figure 9.

Table 7

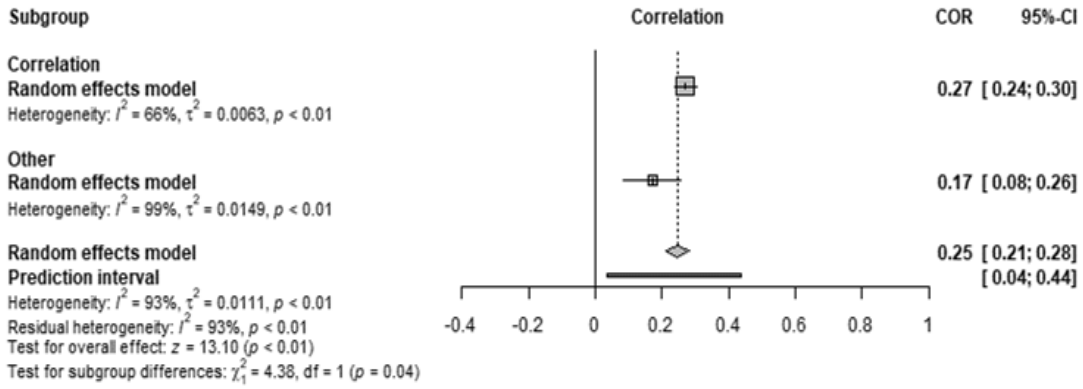
Subgroup Analyses Results Comparing the Effects Across Original Reporting of Effect

Original reporting	Correlation	95% CI	*k*	X^2	*p*
Correlation	0.2699	0.2368; 0.3024	35	4.38	0.0363
Other	0.1717	0.0832; 0.2576	8		

Note. *k* denotes the number of studies.

Figure 9

Original Reporting Forest Plot



Exposure Variable. A meta-regression was completed to explore the effect of the number of items in the exposure measures as a moderator. A significant effect was found. With every increase in exposure measure item number, there was a 0.0057 (SE = 0.022, z = 2.53, p = .011) increase in r. As Paradies et al (2015) compared exposure measures with 8 or less items to exposure measures with 9 and over items, a sub-group analysis was also completed. No significant difference was found (p = .51). A meta-regression was completed to explore the effect of the Cronbach's alpha of the exposure measures; this was not significant (p = .092). When looking at the exposure measures used in each article, subgroup analysis revealed no significant difference between the top four most used measures of PRD (schedule of racist events, racism, and life experiences scales, perceived ethnic discrimination questionnaire, subtle and blatant racism scale). However, primary studies using a validated measure had significantly higher correlation estimates (r = .27) and lower heterogeneity (I^2 = 79%) than primary studies using an unvalidated measure (r = .17, I^2 = 99%, p = .016). Full details are reported in Table 8 and Figure 10.

Table 8

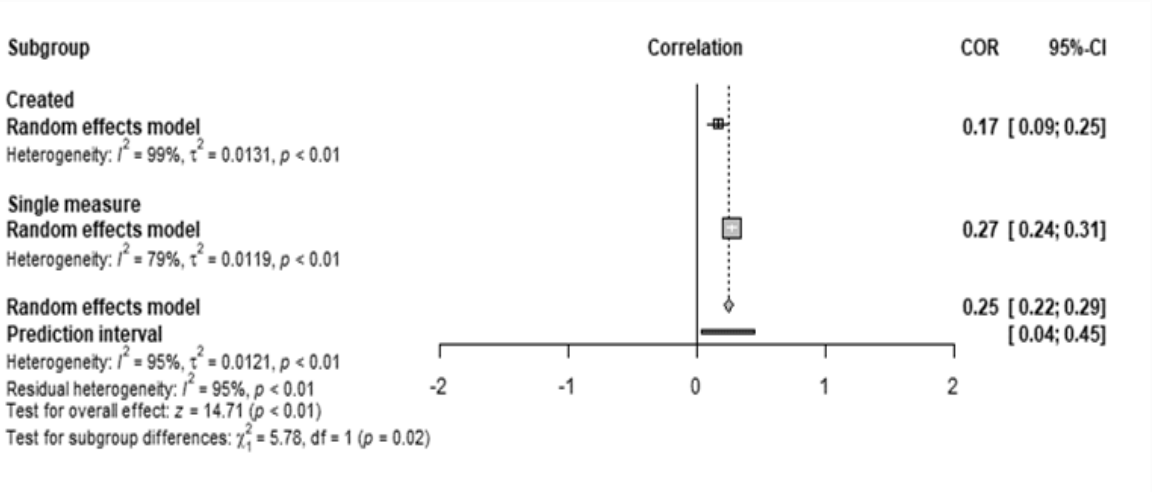
Subgroup Analyses Results Comparing the Effects Across Exposure Measures

Exposure Measure	Correlation	95% CI	*k*	X^2	*p*
Created measure	.17	0.0851; 0.2455	8	5.78	0.0162
Single validated measure	.27	0.2369; 0.3074	46		

Note. *k* denotes the number of studies.

Figure 10

Exposure Measure Forest Plot



For exposure type, only primary studies that looked at direct and mixed PRD could be compared due to the low frequencies of other exposure types (e.g., indirect (n = 1) and internalised (n = 2)). Primary studies measuring direct exposure had significantly higher correlation estimates (r = .24) and higher heterogeneity (I^2 = 89%) than primary studies measuring mixed exposure types (r = .12, I^2 = 59%, p < .001). Full details are reported in Table 9 and Figure 11.

Table 9

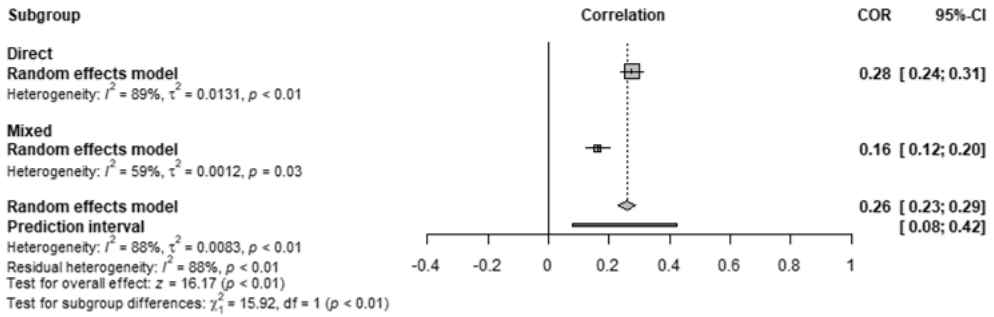
Subgroup Analyses Results Comparing the Effects Across Exposure Type

Exposure Type	Correlation	95% CI	*k*	X^2	*p*
Direct	0.2365	0.2365; 0.3130	42		
Mixed	0.1232	0.1232; 0.2022	6	15.92	<0.001

Note. *k* denotes the number of studies.

Figure 11

Exposure Type Forest Plot



Outcome Variable. For anxiety type, only primary studies that measured GAD and unspecified anxiety could be compared in a sub- group analysis due to the low frequency of other anxiety types being reported (e.g., bodily symptoms (n = 1) and social anxiety (n = 1)). No significant difference was found (p = 0.49). A meta-regression revealed no significant effects for the Cronbach's alpha of the anxiety measures (p = .092). Due to the large variety of anxiety measures, the differences between each measure were not explored. The most frequent measure of anxiety (n = 11), the Depression, Anxiety and Stress Scale (DASS), was compared to the rest of the anxiety measures categorised as 'Other'. Primary studies using the DASS had significantly higher correlation estimates (r = .33) and lower heterogeneity (I^2 = 26%) than primary studies using one of the 'Other' anxiety measures (r = .25, I^2 = 95%, p = .0046). Full details are reported in Table 10 and Figure 12.

Table 10

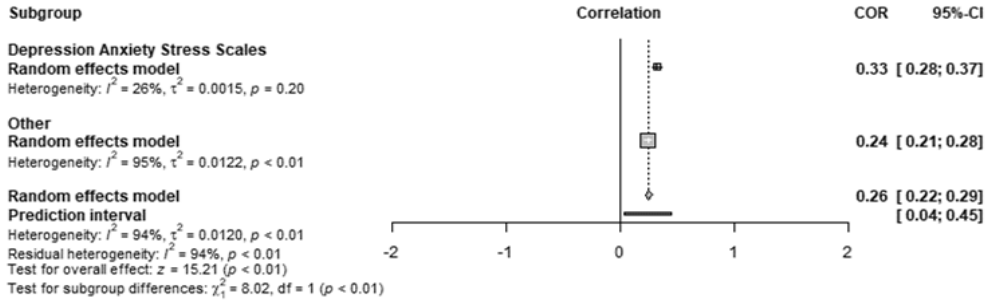
Subgroup Analyses Results Comparing the Effects Across Anxiety Measures

Original reporting	Correlation	95% CI	*k*	X^2	*p*
DASS	0.3254	0.2819; 0.3677	10	8.02	0.0046
Other	0.2448	0.2095; 0.2795	47		

Note. *k* denotes the number of studies.

Figure 12

Anxiety Measure Forest Plot



Analysis of Participant Level Moderators

Age. A sub-group analysis was completed to explore the effect of age as a moderator. Despite the discrepancy in group sizes (18 years and older (n = 39), under 18 years (n = 6)), primary studies with participants under 18 observed significantly higher effects (r = .30) and lower heterogeneity (I^2 = 34%) than studies that included participants who were 18 and over (r = .24, I^2 = 88%, p = .029). Full details are reported in Table 11 and Figure 13.

Biological Sex. A sub-group analysis explored the effect of sex as a moderator. The majority of primary studies included more than one sex (n = 31), these studies were labelled ‘mixed’ and were compared with studies that only recruited female (n = 5) or male (n = 5) participants. No significant difference was found (p = .35).

Table 11

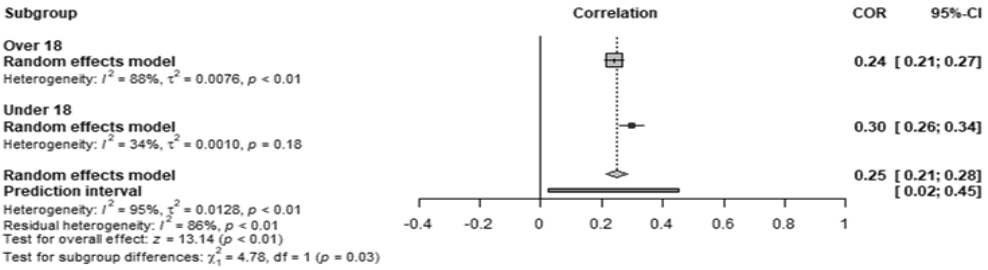
Subgroup Analyses Results Comparing the Effects Across Age Category

Age Category	Correlation	95% CI	*k*	X^2	*p*
18 and Over	0.2397	0.2080; 0.2708	39	4.78	.0288
Under 18	0.2982	0.2560; 0.3393	6		

Note. *k* denotes the number of studies.

Figure 13

Age Category Forest Plot



Birthplace. A sub-group analysis was completed to explore the effect of birthplace. Most studies (*n* = 38) did not record birthplace or did not enable it to be extracted. No studies explicitly stated that they only included locally born participants. Studies that included both foreign and local born participants were labelled as 'mixed' (n = 15) and were compared to studies that included only foreign-born participants (n = 4). Studies that included 'foreign born' only participants had significantly higher correlations (*r* = .35) than studies which included mixed participants (*r* = .21, *p* = .02). Full details are reported in Table 12 and Figure 14.

Ethnicity. Sub-group analysis was used to explore the effect of ethnicity as a moderator. No significant effect was found (*p* = .25)

Education. Sub-group analysis was used to assess the effect of level of education completed as a moderator. No significant difference was found (*p* = .072).

Table 12

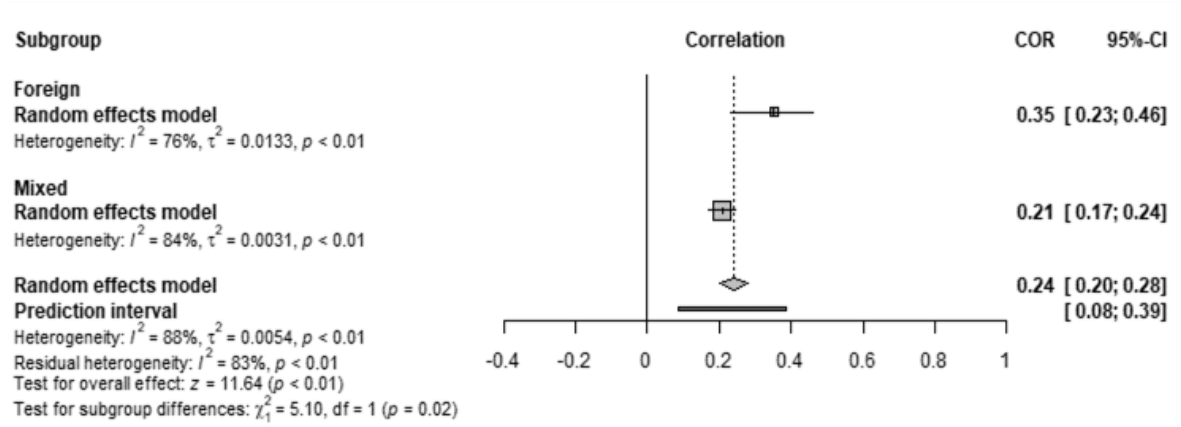
Subgroup Analyses Results Comparing the Effects Across Birthplace

Birthplace	Correlation	95% CI	*k*	X^2	*p*
Foreign born	0.3522	0.2332; 0.4608	4	5.10	0.0239
Mixed	0.2083	0.1716; 0.2444	15		

Note. *k* denotes the number of studies.

Figure 14

Birthplace Forest Plot



Comparison to Previous Meta-Analysis

Table 13 summarises the results of the study-level moderator analysis in this review to that in Paradies et al (2015). The overall effect from this current meta-analysis was also compared to the overall effect found by Paradies et al (2015). Figure 15 shows that the difference between the overall effect found in this review (r = .26) and Paradies et al (2015) review (r =.24) was not statistically significant (p = .59).

Table 13

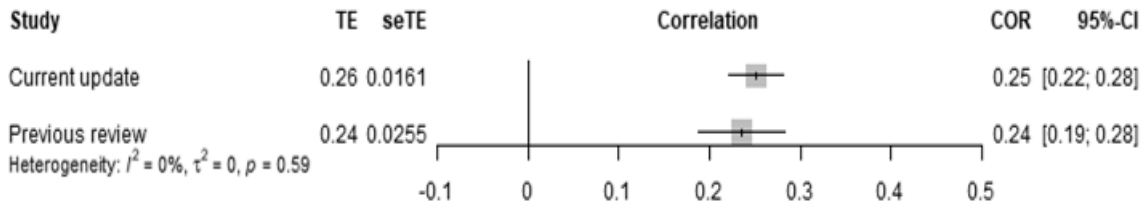
Study Level Moderator Comparisons Between Paradies et al. (2015) and the Current Review

Moderator	Paradies et al. (2015)			Current Review		
	Significant	Moderator Level	Correlation (*n*)	Significant	Moderator	Correlation (*n*)
Publication status	No	Published	0.24 (32)	No	Published	0.23 (45)
		Unpublished	0.25 (8)		Unpublished	0.30 (12)
Publication year	Yes	2005 or earlier	0.16 (16)	No	Publication year	Regression
		2006 or later	0.30 (22)			
Country of study	No	United States	0.23 (31)	No	United States	0.26 (54)
		Non-United States	0.28 (9)		Non-United States	0.26 (3)
Longitudinal vs cross sectional	Not analysed	Cross-sectional	N/A	No	Cross-sectional	0.27 (44)
		Longitudinal	N/A		Longitudinal	0.28 (6)
Sampling procedure	Yes	Representative	0.14 (7)	No	Representative	0.24 (5)
		Non-representative	0.27 (31)		Non-representative	0.27 (47)
Exposure Type	No	Direct	0.27 (26)	Yes	Direct	.24 (42)
		Group or vicarious	0.17 (9)		Mixed (direct and vicarious)	.12 (6)
Exposure timeframe	No	3 years or less	0.21 (7)	No	1 year or less	0.29 (15)
		More than 3 years	0.34 (9)		More than 5 years	0.23 (14)
		Not specified	0.21 (19)		Not specified	0.25 (26)
Exposure number of items	No	8 or less	0.26 (14)	No	8 or less	0.24 (14)
		9 or more	0.23 (19)		9 or more	0.26 (39)
	Not analysed	Regression	N/A	Yes	Item number	Regression
Exposure Cronbach's Alpha	No	Lower than 0.8	0.26 (5)	No	Lower than 0.8	0.16 (10)
		0.8 or higher	0.25 (20)		0.8 or higher	0.26 (35)
	Not analysed	Regression	N/A	No	Cronbach's Alpha	Regression
Exposure instrument		Not available.		No	Subtle and blatant racism scale	0.31 (3)

			Schedule of racist events	0.27 (5)
			Racism and life experience scales	0.30 (3)
			Perceived ethnic discrimination questionnaire	0.36 (6)
Validated vs unvalidated	Not analysed	Yes	Validated	.27 (46)
			Unvalidated	.17 (8)
Experimental vs not	Not analysed	Not available due to insufficient number of studies		
Anxiety type	Not analysed	No	GAD	0.14 (9)
			Unspecified	0.26 (43)
Anxiety instrument	Not analysed	Yes	DASS	0.33 (10)
			Other	0.24 (47)
Cronbach's alpha (anxiety)	Not analysed	No	Lower than 0.8	0.28 (7)
			0.8 or higher	0.19 (37)
		No	Cronbach's Alpha	Regression
Original reporting of effect	Not analysed	Yes	Correlation	0.27 (47)
			Other	0.17 (10)

Figure 15

Forest Plot Comparing the Overall Effects of the Previous and Current Reviews



Discussion

According to Cohen's categorisation of small (r = .10), medium (r = .30) and large (r = .50) effects, this review found a statistically small but significant overall effect (r = .25; Cohen, 1992). The overall effect found in this review was not significantly different to the effect found by Paradies et al. (2015). The current review does show that the research literature continues to find evidence that experiences of PRD and anxiety are positively associated. Due to the high level of heterogeneity present, further analysis focused on identifying which factors may have contributed to the variation of effect sizes between primary studies.

Risk of Bias

All areas of bias revealed varying study quality. The only exception to this was generalisability, in which no primary studies were rated as high risk. The variation in quality across studies may have contributed to the high heterogeneity variance of the overall effect. Studies rated as low risk for reporting bias and detection bias had significantly larger effects than those rated as higher risk. This demonstrates that higher quality primary studies reported higher effect sizes. Primary studies with high risk for reporting and detection bias may not have accurately or consistently measured the association between PRD and anxiety. This suggests that lower quality studies in the evidence base may contribute to an underestimation of the association between PRD and anxiety. When looking at what caused studies to be rated unclear or high risk, it raised further concerns about the effect of confounds on the association between PRD and anxiety (Pearce et al. 2019). For example, many studies did not indicate whether they had reviewed the power of their analysis or did not report sufficient demographics.

Study level-moderators

Paradies et al (2015) and the current review did not find a significant difference in association across published and unpublished primary studies. This review additionally drew upon funnel plots and Egger's test of publication bias which did not provide evidence for publication bias.

Whilst Paradies et al. (2015) found that older studies had significantly smaller effect sizes than newer studies, a significantly larger effect size in this updated review was not found. This may reflect the smaller timeframe across which this meta-analysis searched (five years) compared to Paradies et al (2015; all studies up until 2013) but could also suggest a plateau in effect sizes. Further reviews spanning a greater length of time will be needed to track this trajectory before more firm conclusions can be made.

In the current review, a significant difference was not found between non-representative and representative samples. This is not in line with the findings of Paradies et al. (2015) where non-representative samples had higher effect sizes. However, in the current study, there was a large imbalance in the number of studies within each group, with only 5 studies using a representative sampling procedure compared to 47 studies using a non-representative sampling procedure. Therefore, it is possible that the sub-group analysis was low in power to detect a significant difference if there was one, risking a Type II error (Brookes et al., 2001).

Paradies et al (2015) did not find a significant difference between exposure type, although this may have been due to a type II error as previously described. Whilst a direct comparison could not be made in the current review due to the way studies reported their data, a significantly larger effect was found for direct experiences of PRD compared to direct and vicarious trauma combined. This could suggest that experiencing racial discrimination

directly has a stronger relationship with anxiety compared to indirect exposure. However, as a direct comparison could not be made, only a tentative conclusion can be drawn.

Despite both Paradies et al. (2015) and the current study finding no significant difference between US and non-US studies, there were many more US studies than non-US studies in both reviews. This again reduces confidence in the power of the sub-group analysis. Whilst not significant, both Paradies et al. (2015) and the current review found that non-US studies had higher effect sizes than US studies. Further research based outside the US is encouraged to allow for a less biased comparison in future analyses.

It is also important to highlight, as found in both Paradies et al (2015) and the current review, that across primary studies many different questionnaires were used to measure PRD. This brings into question whether studies are defining and measuring PRD in the same way. Paradies et al (2015) was unable to run sub-group analyses across PRD measures due to how infrequently each measure was encountered. In the current review, the top four most used measures were used frequently enough to enable a sub-group analysis. No significant differences were found between the top four most frequently used measures. Whilst this can increase confidence in the construct validity of the top four most frequently used measures, a comprehensive understanding of whether there is significant variation across the range of measures used could not be completed. Although not included in Paradies et al (2015), the current review was able to compare the effects of validated measures used against unvalidated measures. Unvalidated measures reported significantly lower effect sizes and higher heterogeneity than those using validated measures. With 14% of studies using unvalidated measures, this may have also contributed to an underestimation of the overall association between PRD and anxiety in the current review.

The Cronbach's alpha of PRD measures did not reveal any significant findings. However, Cronbach's alpha only measures internal consistency, and this may not be a sufficient measure of reliability on its own. A recommendation from Paradies et al. 2015 was for primary studies to outline the reliability (e.g., interrater, test-retest), validity, and psychometric properties of the measures they used more explicitly. Given the risk of bias assessment in this current review, this does not seem to have improved for the studies included in this current review.

A significant difference was not found between PRD measures that had 8 or less items and those that had 9 or more items in Paradies et al (2015) or the current review. However, when a meta-regression was used to consider the impact of item number on effect sizes, a significant effect was found. This suggests that using 8 and 9 items as a cut off may not have been effective at identifying a difference. Future reviews may want to consider using a different cut off in their sub-group analyses. This difference in findings between a sub-group analysis and meta-regression also highlights the importance of considering which statistical analyses are most appropriate and helpful for the research.

Novel Moderators

The current review analysed five study level moderators that were not included in Paradies et al (2015) as it was felt that they would further add to the findings of this review. One moderator, study design, could not be analysed due to an insufficient number of studies using an experimental study design. This makes sense considering the ethical implications of experimentally manipulating racial discrimination. This also highlights a limitation that is likely to be a consistent part of this evidence base. No significant findings occurred for two moderators: anxiety type and the anxiety measure Cronbach's alpha. However, there was a significant effect found for anxiety measure and reporting of effect size. The DASS was

compared to all other anxiety measures combined. Studies using the DASS reported significantly higher effect sizes. This could suggest that DASS measures anxiety differently to the other anxiety measures. Another novel finding was that studies which originally reported effects as correlations reported significantly higher effect estimates than those which were converted. This again provides evidence that the overall effect identified is an underestimation. Future meta-analyses should take this into consideration when thinking about whether to include studies where effects must be converted.

Participant level-moderators

In the current review, significant differences were not found across sex, education, or ethnicity. However, this review did find that PRD was significantly moderated by participant age. Despite the small group size for participants under 18, effects were significantly larger than those for participants 18 years and over. PRD was also significantly moderated by birth country status. Birthplace could not be compared directly due to the way the included studies reported this factor. However, foreign born participants did have significantly higher effect sizes when compared to local and foreign-born participants combined. It is hard to draw conclusions from these findings due to the imbalanced group sizes and inability to make direct comparisons between groups. Despite larger effects for participants under 18 and foreign-born participants, this study found that, in line with Paradies et al (2015), most participants were over 18 and born locally. This highlights the importance of continuing to diversify study samples as the evidence suggests those who may be the most impacted on are the least researched.

Limitations

The current review was not able to match Paradies et al (2015) consistently, which limited the ability to accurately compare between the two reviews. Discrepancies included

access restrictions to databases, the number of years over which the reviews occurred, and the way in which data was analysed. Furthermore, the researchers screening and extracting data were different and this could have contributed to subjective differences.

Need For Increased Study Diversity

The inability to include primary studies not written in English was a significant limitation. The importance of future reviews rectifying this was additionally highlighted by larger effect sizes found in non-US studies compared to US studies. A recommendation for future reviews would be to make all attempts to overcome this barrier and to include studies from as many different countries, in as many different languages as possible.

A major limitation that hindered analysis was the large discrepancy between group sizes across several factors. For some imbalanced groups, this meant analysis was not completed. For imbalanced groups where sub-group analyses were still completed, this may have introduced an increased risk of type-II errors. . A lack of or inefficient reporting in primary studies was another limitation that may have contributed to the imbalance between groups. Not reporting study and participant level factors (i.e. participant birthplace) was commonly encountered and impacted on both moderator analyses and quality appraisals. Occasionally, studies provided demographics (e.g., foreign or local born status) but not in a way that enabled effect sizes for the different groups to be isolated and extracted. For studies where this occurred, they could not be effectively included in the sub-group or moderator analyses. Form the information that could be extracted, this study found 90% of participants had completed education beyond high school and 86% of participants were from just three ethnic groups (Latinx, Multiple Ethnicities, Black American) highlighting that several ethnic, demographic, and socioeconomic groups were underrepresented. Where sub-group analyses did not find significant differences across these groups, there was a risk of a type-II error. It

may be that if the groups sizes were more balanced, a significant effect could have been found. For some comparisons where the risk of type-II error was high, the much smaller group showed a higher effect estimate (e.g., under 18s and foreign-born participants). This could mean that if a true effect was missed, it was the underrepresented participant group which had a higher association with anxiety. This adds to the pertinence of increasing the diversity of studies and better representing marginalised groups. It has been argued that sub-group analyses can be completed inappropriately where there's a lack of power and that this may result in misleading conclusions that fail to identify a difference when there was one (Brookes et al., 2001). Readers should be cautious of accepting non-significant findings where this may be the case. It may be important for future meta-analyses to take this into account when presenting their results. The imbalance between groups also highlights potential gaps for future research. It will be important for future research to explore the relationship between PRD and anxiety across a more diverse array of ethnic groups increase representation and facilitate analysis across .

High Heterogeneity of Variance

The high heterogeneity of variance found in this review meant it was likely that there were variables other than chance contributing to the variation in effects between primary studies and thus to the overall effect. Factors were identified as contributing to systematic variation and increased heterogeneity in this review (e.g., high risk of reporting or detection bas, studies using unvalidated measures, studies reporting alternative effect types). Future research may be able to reduce the influence of these confounding variables by increasing methodological quality (e.g., ensuring all measures are validated). Other factors that resulted in significant differences in effect sizes across studies were not due to methodological limitations but rather due to some groups moderating the association between PRD and

anxiety differently. For example, under 18 and over 18s. As effect sizes significantly differed across groups and contributed to the high heterogeneity found in this review, it may be more helpful to separate these groups as to gain a more accurate effect size for individual groups and so between group variation is not lost within the overall effect (Harrer et al., 2021).

Furthermore, in the current review, a decision was made to include all primary studies that met the inclusion criteria, regardless of their quality or the study/participant characteristics. This decision was based upon the desire to replicate Paradies et al.'s study, which did not screen out studies based on their quality, and to gain a holistic overview of the current evidence base rather than an overview of just high quality studies. This decision was further supported by the limited number of studies that were rated low risk, meaning most studies would have been excluded based on low quality. However, this decision was balanced against the finding that low quality ratings in some areas were associated with significantly lower effect estimates and higher heterogeneity. As such, there was a trade-off between being exhaustive and being rigorous, with the overall effect representing the findings of the current evidence base as a whole, whilst likely being an underrepresentation of the true effect between PRD and anxiety. As the evidence base expands, improves, and diversifies, it may be possible and important, for future meta-analyses to screen out studies based on their quality, or to complete separate analyses for different study and participant level moderators. The hope is that this would reduce heterogeneity and ascertain a more representative overall effect size.

Measures

It is particularly pertinent for this area of research to consider that many psychometrics are validated on WEIRD (white, educated, industrialised, rich, and democratic) participants (Laajaj et al., 2019). If validation and research groups differ, it should also be stated whether

validation has been completed for the demographic to which the measure is being applied (Britt-Spells et al., 2018). This was a factor that very few of the primary studies took into consideration.

Due to the personal nature of PRD and anxiety, self-report questionnaires were observed to be used across most primary studies. It has been suggested that individuals are likely to under report their experiences due to forgetting or wanting to respond in a socially desirable way (Pearce et al., 2019) and this may be particularly applicable when participants are self-reporting about experiences of PRD (Holmes, 2009). However, as experimentally manipulating racial discrimination is not possible for ethical reasons, this limitation requires future researchers to think carefully about how they can minimise the bias introduced by self-report measures (van Berkel et al., 2020). Studies exploring the effects of PRD on other mental health difficulties have also encountered this lack of consistency and standardisation regarding psychometrics (Britt-Spells et al., 2018). Some reviews identifying this problem have chosen to provide a narrative summary of the research instead of employing meta-analytic methods (Pearce et al, 2019).

Taking into consideration the large range of measures used to assess PRD and anxiety, as well as the concerns regarding reliability and validity of these measures, a recommendation for future research is to support the identification and consistent use of appropriate and high-quality measures in this research area.

Conclusion

The current review provided a comprehensive update on the evidence base exploring PRD and anxiety following the review by Paradies et al (2015). Whilst a significant difference between the two reviews was not found, it seems likely that the current review could represent an underestimated effect size. The number of low-quality studies and a lack of studies

including underrepresented demographic groups were shown to significantly contribute towards this. It is hoped that by improving study quality, including underrepresented groups in future research, and refining the measures applied to this research, a more accurate overall effect size can be established. Identifying and exploring the high heterogeneity across the studies was vital in understanding the true landscape of the evidence base. Excluding studies to reduce heterogeneity would have resulted in a limited understanding of the differences between sub-groups and the impact of study quality. However, future studies should focus on improving the underrepresentation of certain participant and study level factors as well as improving consistency and quality across the evidence base. Systematic reviews and meta-analyses should continue to be updated regularly to track changes.

Whilst this review did not find a significant change in the association between PRD and anxiety compared to Paradies et al (2015), it still found a significant effect and continues to provide evidence for PRD being associated with anxiety, a global health burden. This review adds to the call for Clinical Psychologists to draw upon the Social Graces (Burnham, 2018) and the Power Threat Meaning Framework (Johnstone et al., 2018) to improve their practice. These models place emphasis on the impact of individual and systemic relationships on a person's mental health, well-being, and ability to reach their potential and thrive. Given that a significant effect was again found between PRD and anxiety, recommendations can be given to Clinical Psychologists to adapt their practice to incorporate learning from these findings. Experiences of racial discrimination and its association with anxiety, should be routinely considered in service planning and enquired about in mental health and psychological assessments. This would be key for developing accessible services, accurate and helpful client formulations, and thus appropriate, person-centred treatment plans.

References

Ackerman-Barger, K., Boatright, D., Gonzalez-Colaso, R., Orozco, R., & Latimore, D. (2020). Seeking inclusion excellence: Understanding racial microaggressions as experienced by underrepresented medical and nursing students. *Academic Medicine, 95*(5), 758–763. https://doi.org/10.1097/ACM.0000000000003077

Allen, G. E. K., Conklin, H., & Kane, D. K. (2017). Racial discrimination and psychological health among Polynesians in the U.S. *Cultural Diversity & Ethnic Minority Psychology, 23*(3), 416-424. https://doi.org/10.1037/cdp0000133

Antunes, A., Frasquilho, D., Azeredo-Lopes, S., Neto, D., Silva, M., Cardoso, G., & Caldas-de-Almeida, J. M. (2018). Disability and common mental disorders: results from the world mental health survey initiative Portugal. *European Psychiatry*, *49*, 56–61. https://doi.org/10.1016/j.eurpsy.2017.12.004

Assari, S. (2017). Psychiatric disorders differently correlate with physical self-rated health across ethnic groups. *Journal of Personalized Medicine*, *7*(4). https://doi.org/10.3390/jpm7040006

Atwal, K., & Wang, C. (2018). Religious head covering, being perceived as foreigners, victimization, and adjustment among Sikh American adolescents. *School Psychology, 34*(2), 233-243. https://doi.org/10.1037/spq0000301

Bauermeister, J. A., Muessig, K. E., Flores, D. D., LeGrand, S., Choi, S., Dong, W., Harper, G. W., & Hightow-Weidman, L. B. (2018). Stigma diminishes the protective effect of social support on psychological distress among young black men who have sex with men. *AIDS Education and Prevention, 30*(5), 406-418. https://doi.org/10.1521/aeap.2018.30.5.406

Baujat, B., Mahé, C., Pignon, J.-P., & Hill, C. (2002). A graphical method for exploring heterogeneity in meta-analyses: Application to a meta-analysis of 65 trials. *Statistics in Medicine*, *21*(18), 2641–2652. https://doi.org/10.1002/sim.1221

Baxter, Scott, K., Vos, T., & Whiteford, H. (2012). Global prevalence of anxiety disorders: A systematic review and meta-regression. *Psychological Medicine*, *43*(5), 897-910. https://doi.org/10.1017/S003329171200147X

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*(4), 1088–1101. https://doi.org/10.2307/2533446

Borders, A., & Hennebry, K. A. (2015). Angry rumination moderates the association between perceived ethnic discrimination and risky behaviors. *Personality and Individual Differences*, *79*, 81–86. https://doi.org/10.1016/j.paid.2015.01.047

Borenstein. (2009). *Introduction to meta-analysis*. Wiley.

Brookes, S. T., Whitley, E., Peters, T. J., Mulheran, P. A., Egger, M., & Davey Smith, G. (2001). Subgroup analysis in randomised controlled trials: quantifying the risks of false-positives and false-negatives. *Health Technology Assessment*, *5*(33). https://doi.org/10.3310/hta5330

Britt-Spells, A. M., Slebodnik, M., Sands, L. P., & Rollock, D. (2018). Effects of perceived discrimination on depressive symptoms among black men residing in the united states: A meta-analysis. *American Journal of Men's Health*, *12*(1), 52–63. https://doi.org/10.1177/1557988315624509

Brydges, C. R. (2019). Effect size guidelines, sample size calculations, and statistical power in gerontology. *Innovation in Aging*, *3*(4). https://doi.org/10.1093/geroni/igz036

Budhwani, H., Hearld, K. R., & Chavez-Yenter, D. (2015). Generalized anxiety disorder in racial and ethnic minorities: A case of nativity and contextual factors. *Journal of Affective Disorders*, *175*, 275–280. https://doi.org/10.1016/j.jad.2015.01.035

Burnham, J. (2018). Developments in Social GRRRAAACCEEESSS: visible–invisible and voiced–unvoiced. *Culture and Reflexivity in Systemic Psychotherapy*, 139-160.

Cano, M. A., Castro, Y., de Dios, M. A., Schwartz, S. J., Lorenzo-Blanco, E. I., Roncancio, A. M., Martinez, M. J., Sheehan, D. M., Auf, R., Pina-Watson, B., Huynh, Q.-L., & Zamboanga, B. L. (2016). Associations of ethnic discrimination with symptoms of anxiety and depression among Hispanic emerging adults: A moderated mediation model. *Anxiety, Stress, and Coping*, *29*(6), 699-707. https://doi.org/10.1080/10615806.2016.1157170

Cariello, A. N., Perrin, P. B., Williams, C. D., Espinoza, G. A., Morlett-Paredes, A., Moreno, O. A., & Trujillo, M. A. (2019). Moderating influence of enculturation on the relations between minority stressors and physical health via anxiety in Latinx immigrants. *Cultural Diversity and Ethnic Minority Psychology*, *26*, 356–366. https://doi.org/10.1037/cdp0000308

Carter, S. E., Walker, R. L., Cutrona, C. E., Simons, R. L., & Beach, S. R. H. (2016). Anxiety mediates perceived discrimination and health in African-American women. *American Journal of Health Behavior*, *40*(6), 697–704. https://doi.org/10.5993/AJHB.40.6.2

Celebi, E., Verkuyten, V., & Bagci, S. (2017). Ethnic identification, discrimination, and mental and physical health among Syrian refugees: The moderating role of identity needs. *European Journal of Social Psychology*. *47*(7), 832-843. https://onlinelibrary.wiley.com/doi/abs/10.1002/ejsp.2299

Chae, D. H., Epel, E. S., Nuru-Jeter, A. M., Lincoln, K. D., Taylor, R. J., Lin, J., Blackburn, E. H., & Thomas, S. B. (2016). Discrimination, mental health, and leukocyte telomere length among African American men. *Psychoneuroendocrinology*, *63*, 10–16. https://doi.org/10.1016/j.psyneuen.2015.09.001

Chee, J. (2015). *Pearson's product-moment correlation: sample analysis* [University of Hawaii at Mānoa School of Nursing]. https://doi.org/10.13140/RG.2.1.1856.2726

Choi, K.-H., Paul, J., Ayala, G., Boylan, R., & Gregorich, S. E. (2013). Experiences of discrimination and their impact on the mental health among African American, Asian and pacific islander, and Latino men who have sex with men. In *American Journal of Public Health*. *103* (5), 868–874. https://doi.org/10.2105/AJPH.2012.301052

Cohen. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037//0033-2909.112.1.155

Conklin, H. D. (2011). *Perceived racism and mental health: A meta-analytic review* [Doctoral dissertation, Brigham Young University].

Cooke, C. L., Bowie, B. H., & Carrere, S. (2014). Perceived discrimination and children's mental health symptoms. *Advances in Nursing Science*, *37*(4), 299-314. https://doi.org/10.1097/ANS.0000000000000047

Corey, D. M., Dunlap, W. P., & Burke, M. J. (1998). Averaging correlations: expected values and bias in combined Pearson *r* s and fisher's *z* transformations. *The Journal of General Psychology*, *125*(3), 245–261. https://doi.org/10.1080/00221309809595548

Corona, R., Rodríguez, V., McDonald, S., Velazquez, E., Rodríguez, A., & Fuentes, V. (2017). Associations between cultural stressors, cultural values, and Latina/o college students' mental health. *Journal of Youth & Adolescence*, *46*(1), 63–77. https://doi.org/10.1007/s10964-016-0600-5

Daoud, N., Gao, M., Osman, A., & Muntaner, C. (2018). Interpersonal and institutional ethnic discrimination, and mental health in a random sample of Palestinian minority men smokers in Israel. *Social Psychiatry and Psychiatric Epidemiology, 53*(10), 1111-1122. https://doi.org/10.1007/s00127-018-1531-0

de Mendoza, V. B., Huang, Y., Crusto, C. A., Sun, Y. V., & Taylor, J. Y. (2018). Perceived racial discrimination and DNA methylation among African American women in the interGEN Study. *Biological Research for Nursing*, *20*(2), 145–152. https://doi.org/10.1177/1099800417748759

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*(3), 177–188. https://doi.org/10.1016/0197-2456(86)90046-2

Doyle, J. A. (2019). '*I'm black and I'm strong:' the health effects of resilience in the face of discrimination among black American men in metropolitan Seattle* [Doctoral dissertation University of Washington].

Drakeford, N. M. (2019). *Strong Black woman: an exploration of coping, suppression, and psychological distress.* [Doctoral dissertation, University of Akron].

Durkee, M. I., & Williams, J. L. (2015). Accusations of acting white: links to black students' racial identity and mental health. *Journal of Black Psychology*, *41*(1), 26–48. https://doi.org/10.1177/0095798413505323

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Escovar, E. L., Craske, M., Roy-Byrne, P., Stein, M. B., Sullivan, G., Sherbourne, C. D., Bystritsky, A., & Chavira, D. A. (2018). Cultural influences on mental health

symptoms in a primary care sample of Latinx patients. *Journal of Anxiety Disorders, 55*, 39–47. https://doi.org/10.1016/j.janxdis.2018.03.005

Evans, M. K., Rosenbaum, L., Malina, D., Morrissey, S., & Rubin, E. J. (2020). Diagnosing and treating systemic racism. *New England Journal of Medicine*, *383*(3), 274–276. https://doi.org/10.1056/NEJMe2021693

Fang, K., Friedlander, M., & Pieterse, A. L. (2016). Contributions of acculturation, enculturation, discrimination, and personality traits to social anxiety among Chinese immigrants: a context-specific assessment. *Cultural Diversity & Ethnic Minority Psychology*, *22*(1), 58–68. https://doi.org/10.1037/cdp0000030

Fisher, R. A., & Russell, E. J. (1922). On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A-Mathematical Physical and Engineering Sciences*, *222*(594–604), 309–368. https://doi.org/10.1098/rsta.1922.0009

Franco, M., & Carter, S. (2019). Discrimination from family and substance use for multiracial individuals. *Addictive Behaviors*, *92*, 203–207. https://doi.org/10.1016/j.addbeh.2019.01.008

Graham, J. R., Calloway, A., & Roemer, L. (2015). The buffering effects of emotion regulation in the relationship between experiences of racism and anxiety in a black American sample. *Cognitive Therapy And Research*, *39*(5), 553–563. https://doi.org/10.1007/s10608-015-9682-8

Graham, J. R., West, L., & Roemer, L. (2015). A preliminary exploration of the moderating role of valued living in the relationships between racist experiences and anxious and depressive symptoms. *Journal of Contextual Behavioral Science, 4*(1), 48–55. https://doi.org/10.1016/j.jcbs.2014.11.001

Harrer, M., Cuijpers, P., Furukawa, T., & Ebert, D. (2021). *Doing meta-analysis with R: A Hands-On Guide*. Chapman and Hall. https://doi.org/10.1201/9781003107347

Heard-Garris, N. J., Cale, M., Camaj, L., Hamati, M. C., & Dominguez, T. P. (2018). Transmitting trauma: A systematic review of vicarious racism and child health. *Social Science & Medicine, 199*, 230–240. https://doi.org/10.1016/j.socscimed.2017.04.018

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*, 486–504. https://doi.org/10.1037/1082-989X.3.4.486

Higgins, Altman, D. G., Gotzsche, P. C., Juni, P., Moher, D., Oxman, A. D., Savovic, J., Schulz, K. F., Weeks, L., Sterne, J. A. C. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal, 343*, 5928–5928. https://doi.org/10.1136/bmj.d5928

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal*, *327*(7414), 557–560.

Holmes, J. D. (2009). Transparency of self-report racial attitude scales. *Basic and Applied Social Psychology*, *31*(2), 95–101. https://doi.org/10.1080/01973530902876884

Hope, E. C., Velez, G., Offidani-Bertrand, C., Keels, M., & Durkee, M. I. (2018). Political activism and mental health among Black and Latinx college students. *Cultural Diversity & Ethnic Minority Psychology*, *24*(1). https://doi.org/10.1037/cdp0000144

Hurd, N. M., Varner, F. A., Caldwell, C. H., & Zimmerman, M. A. (2014). Does perceived racial discrimination predict changes in psychological distress and substance use over time? An examination among black emerging adults. *Developmental Psychology, 50*(7), 1910–1918. https://doi.org/10.1037/a0036438

Huynh, V. W., Huynh, Q.-L., & Stein, M.-P. (2017). Not just sticks and stones: indirect ethnic discrimination leads to greater physiological reactivity. *Cultural Diversity & Ethnic Minority Psychology*, *23*(3), 425–434. https://doi.org/10.1037/cdp0000138

Jackson, S. C. (2015). *Contemporary racism and the Asian American experience: the impact of lifetime racial microaggression stress on psychological functioning and risk behavior in Asian American young adults* [Doctoral dissertation].

Jerald, M. C., Cole, E. R., Ward, L. M., & Avery, L. R. (2017). Controlling images: How awareness of group stereotypes affects Black women's well-being. *Journal of Counseling Psychology*, *64*(5). https://doi.org/10.1037/cou0000233

Jones, C. (2020). *Meta analysis of summary effects using generic inverse variance* (4.1) [English]. University of Birmingham.

Johnstone, L. & Boyle, M. Cromby, J., Dillon, J., Harper, D., Kinderman, P., Longden, E., Pilgrim, D. & Read, J. (2018). *The power threat meaning framework: Towards the identification of patterns in emotional distress, unusual experiences and troubled or troubling behaviour, as an alternative to functional psychiatric diagnosis*. Leicester: British Psychological Society

Joseph, F. (2015). White fatigue: naming the challenge in moving from an individual to a systemic understanding of racism. *Multicultural Perspectives*, *17*(3), 115–124. https://doi.org/10.1080/15210960.2015.1048341

Juang, L. P., Moffitt, U., Kim, S. Y., Lee, R. M., Soto, J. A., Hurley, E., Weisskirch, R. S., Blozis, S. A., Castillo, L. G., Huynh, Q.-L., & Whitborne, S. K. (2016). Cognitive reappraisal and expressive suppression: Links to racial-ethnic discrimination and adjustment among Latino/a and Asian-heritage college students. *Journal of Adolescence*, *53*, 21–33. https://doi.org/10.1016/j.adolescence.2016.08.012

Kim, M. (2014). *The influence of cultural context on risk, resilience, and mental health on emerging Latino adults*. [Doctoral dissertation, University of Notre Dame].

Kim, S. Y., Park, J. E., Lee, Y. J., Seo, H.-J., Sheen, S.-S., Hahn, S., Jang, B.-H., & Son, H.-J. (2013). Testing a tool for assessing the risk of bias for nonrandomized studies showed moderate reliability and promising validity. *Journal of Clinical Epidemiology*, *66*(4), 408–414. https://doi.org/10.1016/j.jclinepi.2012.09.016

Kirkinis, K., Pieterse, A. L., Martin, C., Agiliga, A., & Brownell, A. (2021). Racism, racial discrimination, and trauma: A systematic review of the social science literature. *Ethnicity & Health*, *26*(3), 392–412. https://doi.org/10.1080/13557858.2018.1514453

Kwon, S., & Han, D. (2019). Discrimination, mental disorders, and suicidal ideation in Latino adults: decomposing the effects of discrimination. *Journal of Immigrant & Minority Health, 21*(1), 143–150. https://doi.org/10.1007/s10903-018-0726-5

Laajaj, R., Macours, K., Pinzon Hernandez, D. A., Arias, O., Gosling, S. D., Potter, J., ... & Vakis, R. (2019). Challenges to capture the big five personality traits in non-WEIRD populations. *Science Advances, 5*(7). https://doi.org/10.1126/sciadv.aaw5226

Leath, S., Mathews, C., Harrison, A., & Chavous, T. (2019). Racial identity, racial discrimination, and classroom engagement outcomes among black girls and boys in predominantly black and predominantly white school districts. *American Educational Research Journal*, *56*(4), 1318–1352. https://doi.org/10.3102/0002831218816955

Lee, M. R., & Thai, C. J. (2015). Asian American phenotypicality and experiences of psychological distress: More than meets the eyes. *Asian American Journal of Psychology, 6*, 242–251. https://doi.org/10.1037/aap0000015

Lewis, S., & Clarke, M. (2001). Forest plots: Trying to see the wood and the trees. *BMJ : British Medical Journal*, *322*(7300), 1479–1480.

Lo, S. M., Wong, H. C., Lam, C. Y., & Shek, D. T. L. (2020). Common mental health challenges in a university context in Hong Kong: a study based on a review of medical records. *Applied Research in Quality of Life*, *15*(1), 207–218. https://doi.org/10.1007/s11482-018-9673-5

Lunny, C., Pieper, D., Thabet, P., & Kanji, S. (2021). Managing overlap of primary study results across systematic reviews: Practical considerations for authors of overviews of reviews. *Biomed Central*, *21*(1), 140. https://doi.org/10.1186/s12874-021-01269-y

McCoy, H. (2020). Black Lives Matter, and Yes, You are Racist: the parallelism of the twentieth and twenty-first centuries. *Child and Adolescent Social Work Journal*, *37*(5), 463–475. https://doi.org/10.1007/s10560-020-00690-4

Marin, I. A., & Mancini, F. (2017). Romanian immigrants in Italy: mental health and potential protective and vulnerability factors. *Revista de Cercetare si Interventie Sociala, 59*, 21–33.

Martinez, J. H., Eustis, E. H., Arbid, N., Graham-LoPresti, J. R., & Roemer, L. (2020). The role of experiential avoidance in the relation between racial discrimination and negative mental health outcomes. *Journal of American College Health, 70*(2), 461-468. https://doi.org/10.1080/07448481.2020.1754221

Mata-Greves, F., & Torres, L. (2018). Rejection and Latina/o mental health: Intragroup marginalization and intragroup separation. *The American Journal of Orthopsychiatry, 89*(6). https://doi.org/10.1037/ort0000368

Mouzon, D. M., Taylor, R. J., Keith, V. M., Nicklett, E. J., & Chatters, L. M. (2017). Discrimination and psychiatric disorders among older African Americans. *International Journal of Geriatric Psychiatry, 32*(2). https://doi.org/10.1002/gps.4454

Murad, M. H., Chu, H., Lin, L., & Wang, Z. (2018). The effect of publication bias magnitude and direction on the certainty in evidence. *British Medical Journal*, *23*(3), 84–86. https://doi.org/10.1136/bmjebm-2018-110891

Neblett, E. W., Jr., Bernard, D. L., & Banks, K. H. (2016). The moderating roles of gender and socioeconomic status in the association between racial discrimination and psychological adjustment. *Cognitive and Behavioral Practice, 23*(3), 385–397. https://doi.org/10.1016/j.cbpra.2016.05.002

Noh, S., Beiser, M., Kaspar, V., Hou, F., & Rummens, J. (1999). Perceived racial discrimination, depression, and coping: A study of Southeast Asian refugees in Canada. *Journal of Health and Social Behavior, 40*(3), 193. https://doi.org/10.2307/2676348

Olatunji, B. O., Cisler, J. M., & Tolin, D. F. (2007). Quality of life in the anxiety disorders: A meta-analytic review. *Clinical Psychology Review*, *27*(5), 572–581. https://doi.org/10.1016/j.cpr.2007.01.015

Paradies, Y., Ben, J., Denson, N., Elias, A., Priest, N., Pieterse, A., Gupta, A., Kelaher, M., & Gee, G. (2015). Racism as a determinant of health: a systematic review and meta-analysis. *PLOS ONE* , 9(10). https://doi.org/10.1371/journal.pone.0138511

Park, I. J. K., Wang, L., Williams, D. R., & Alegria, M. (2017). Does anger regulation mediate the discrimination-mental health link among Mexican-origin adolescents? A longitudinal mediation analysis using multilevel modelling. *Developmental Psychology, 53*(2), https://doi.org/10.1037/dev0000235

Pearce, J., Rafiq, S., Simpson, J., & Varese, F. (2019). Perceived discrimination and psychosis: A systematic review of the literature. *Social Psychiatry and Psychiatric Epidemiology*, *54*(9), 1023–1044. https://doi.org/10.1007/s00127-019-01729-3

Perreira, K. M., Gotman, N., Isasi, C. R., Arguelles, W., Castañeda, S. F., Daviglus, M. L., Giachello, A. L., Gonzalez, P., Penedo, F. J., Salgado, H., & Wassertheil-Smoller, S. (2015). Mental health and exposure to the united states: Key correlates from the Hispanic community health study of Latinos. *The Journal of Nervous and Mental Disease, 203*(9), 670–678. https://doi.org/10.1097/NMD.0000000000000350

Pizarro, M., & Kohli, R. (2020). "I Stopped Sleeping": Teachers of color and the impact of racial battle fatigue. *Urban Education*, *55*(7), 967–991. https://doi.org/10.1177/0042085918805788

Preciado, A., & D'Anna-Hernandez, K. (2017). Acculturative stress is associated with trajectory of anxiety symptoms during pregnancy in Mexican-American women. *Journal of Anxiety Disorders, 48*, 28–35. https://doi.org/10.1016/j.janxdis.2016.10.005

Priest, N., Paradies, Y., Trenerry, B., Truong, M., Karlsen, S., & Kelly, Y. (2013). A systematic review of studies examining the relationship between reported racism and health and wellbeing for children and young people. *Social Science and Medicine, 95*, 115-127. https://doi.org/10.1016/j.socscimed.2012.11.031

Rehm, J., & Shield, K. D. (2019). Global burden of disease and the impact of mental and addictive disorders. *Current Psychiatry Reports*, *21*(2), 1-7. https://doi.org/10.1007/s11920-019-0997-0

Robinson, O. J., Vytal, K., Cornwell, B. R., & Grillon, C. (2013). The impact of anxiety upon cognition: Perspectives from human threat of shock studies. *Frontiers in Human Neuroscience, 7*(203). https://doi.org/10.3389/fnhum.2013.00203

Rodriguez-Seijas, C., Stohl, M., Hasin, D. S., & Eaton, N. R. (2015). Transdiagnostic factors and mediation of the relationship between perceived racial discrimination and mental

disorders. *Jama Psychiatry*, *72*(7), 706–713. https://doi.org/10.1001/jamapsychiatry.2015.0148

Roy Rosenzweig Center for History and New Media. (2021). *Zotero* (5.0.96) [English]. Roy Rosenzweig Center for History and New Media. www.zotero.org/download

RStudio Team. (2020). *RStudio: Integrated Development for R. RStudio* (1.3.1073) [English]. RStudio, PBC. https://www.rstudio.com/

Ruppar, T. (2020). Meta-analysis: How to quantify and explain heterogeneity? *European Journal of Cardiovascular Nursing*, *19*(7), 646–652. https://doi.org/10.1177/1474515120944014

Schmitt, M. T., Branscombe, N. R., Postmes, T., & Garcia, A. (2014). The consequences of perceived discrimination for psychological well-being: A meta-analytic review. *Psychological Bulletin*, *140*(4). https://doi.org/10.1037/a0035754

Schwartz, S. J., Salas-Wright, C. P., Perez-Gomez, A., Mejia-Trujillo, J., Brown, E. C., Montero-Zamora, P., Meca, A., Scaramutti, C., Soares, M. H., Vos, S. R., Javakhishvili, N., & Dickson-Gomez, J. (2018). Cultural stress and psychological symptoms in recent Venezuelan immigrants to the United States and Colombia. *International Journal of Intercultural Relations, 67,* 25-34. https://doi.org/10.1016/j.ijintrel.2018.09.001

Sedgwick, P. (2013). Meta-analyses: How to read a funnel plot. *British Medical Journal, 346*. https://doi.org/10.1136/bmj.f1342

Sedgwick, P. (2015). Meta-analyses: What is heterogeneity? *British Medical Journal*, *350*. https://doi.org/10.1136/bmj.h1435

Seet, A. Z. (2021). Serving the White nation: Bringing internalised racism within a sociological understanding. *Journal of Sociology*, *57*(2), 213–230. https://doi.org/10.1177/1440783319882087

Serafini, K., Donovan, D. M., Wendt, D. C., Matsumiya, B., & McCarty, C. A. (2017). A comparison of early adolescent behavioral health risks among urban American Indians/Alaska natives and their peers. In American Indian and Alaska native mental health research. *American Indian and Alaska Native Mental Health Research, 24*(2), 1–17. https://doi.org/10.5820%2Faian.2402.2017.1

Sosoo, E. (2019). *The influence of internalized racism on the relationship between racial discrimination and depressive and anxiety symptom distress.* [Doctoral dissertation, The University of North Carolina at Chapel Hill]

Sutter, M., & Perrin, P. B. (2016). Discrimination, mental health, and suicidal ideation among LGBTQ people of color. *Journal of Counseling Psychology*, *63*(1), 98–105.

Swann, G., Stephens, J., Newcomb, M. E., & Whitton, S. W. (2020). Effects of sexual/gender minority- and race-based enacted stigma on mental health and substance use in female assigned at birth sexual minority youth. *Cultural Diversity and Ethnic Minority Psychology*, *26*, 239–249. https://doi.org/10.1037/cdp0000292

Szaflarski, M., & Bauldry, S. (2019). The effects of perceived discrimination on immigrant and refugee physical and mental health. *Advances in Medical Sociology*, *19*, 173–204. https://doi.org/10.1108/S1057-629020190000019009

Takeda, K. (2018). *Discrimination, social support, and internalizing symptoms among Asian-Pacific Islander lesbian gay, bisexual, transgender, queer, and questioning individuals: exploring intersectionality and ecological systems of support*. [Doctoral dissertation, Pace University].

Taylor, Z. E., & Ruiz, Y. (2018). Contextual stressors and the mental health outcomes of Latino children in rural migrant-farmworker families in the Midwest. *Journal of Rural Mental Health, 41*, 284–298. https://doi.org/10.1037/rmh0000082

Terrin, N., Schmid, C. H., & Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. *Journal of Clinical Epidemiology, 58*(9), 894–901. https://doi.org/10.1016/j.jclinepi.2005.01.006

Thelemaque, T. (2018). *An examination of the moderation effects of potentially traumatic events on the relationship between racial and ethnic discrimination and psychological distress symptoms in black Americans*. [Doctoral dissertation, Saint Louis University].

Tuason, M. T. G., Ancheta, I. B., & Battie, C. A. (2014). What impacts the psychological health of Filipino American women? *Asian American Journal of Psychology, 5*(4), 307–315. https://doi.org/10.1037/a0037781

van Berkel, N., Goncalves, J., Hosio, S., Sarsenbayeva, Z., Velloso, E., & Kostakos, V. (2020). Overcoming compliance bias in self-report studies: A cross-study analysis. *International Journal of Human-Computer Studies, 134*, 1–12. https://doi.org/10.1016/j.ijhcs.2019.10.003

Victoria, R. (2014). *Exploring how skin color and racial identity modify the relationship between perceptions of racism and psychological distress among latinas/os.* [Doctoral dissertation, Columbia University].

Walker, R., Francis, D., Brody, G., Simons, R., Cutrona, C., & Gibbons, F. (2017). A longitudinal study of racial discrimination and risk for death ideation in African American youth. *Suicide and Life-Threatening Behavior, 47*(1), 86–102). https://doi.org/10.1111/sltb.12251

Weeks, M. R., & Sullivan, A. L. (2019). Discrimination matters: relations of perceived discrimination to student mental health. *School Mental Health*, *11*(3), 425–437. https://doi.org/10.1007/s12310-019-09309-1

Wei, M., Liang, Y.-S., Du, Y., Botello, R., & Li, C.-I. (2015). Moderating effects of perceived language discrimination on mental health outcomes among Chinese international students. *Asian American Journal of Psychology, 6*(3), 213–222. https://doi.org/10.1037/aap0000021

Whitfield, I. R. (2019). *Racial microaggressions, psychological well-being, and ethnic identification among people of color in the U.S.* [Dissertation, Pace University].

Williams, D. R., Yu, Y., Jackson, J. S., & Andersofn, N. B. (1997). Racial differences in physical and mental health: socio-economic status, stress and discrimination. *Journal of Health Psychology*, *2*(3), 335–351. https://doi.org/10.1177/135910539700200305

Williams, M. T., Kanter, J. W., & Ching, T. H. W. (2018). Anxiety, stress, and trauma symptoms in African Americans: negative affectivity does not explain the relationship between microaggressions and psychopathology. *Journal Of Racial and Ethnic Health Disparities*, *5*(5), 919–927. https://doi.org/10.1007/s40615-017-0440-3

World Health Organization. (2017). *Depression and other common mental disorders: depression and other common mental disorders*. https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf

Wyatt, S. B., Williams, D. R., Calvin, R., Henderson, F. C., Walker, E. R., & Winters, K. (2003). Racism and cardiovascular disease in African Americans. *The American Journal of the Medical Sciences*, *325*(6), 315–331. https://doi.org/10.1097/00000441-200306000-00003

2. EMPIRICAL RESEARCH PAPER: PERSPECTIVES ON ADRRESSING RACIAL INEQUALITY IN THE NHS

Abstract

Aims

The implications of racial discrimination in the NHS include disparities across career outcomes for staff and more negative healthcare outcomes for service users. Strategies employed to tackle racial discrimination have been criticised for their limited impact. Research into individual and organisational change is reviewed and the value of considering an array of stakeholders' perspectives is evidenced. This study aimed to explore stakeholders' perspectives on how to address racial discrimination most effectively in the NHS.

Method

Q-methodology provided an appropriate methodology with which to explore the research question. A Q-set of strategies on how to address racial inequality in the NHS was developed from both a literature search and interviews with stakeholders. Both NHS staff and experts by experience were recruited to complete a forced choice distribution grid Q-sort and data was analysed through factor analysis.

Results

The arrays from 28 Q-sorts were analysed and a three factor solution explaining 51.11% of the total study variance was considered appropriate. The Q-sorts most highly associated with each factor were identified and used to create a representative Q-sort for each factor. The three factors are defined with consensus and distinguishing Q-statements also reviewed.

Conclusion

Each identified factor represents an opinion group. These opinion groups are summarised and compared. It is recognised that the opinions groups could be used as targets for interventions. However, the differences between opinion groups are considered significant barriers to implementing solutions. The strengths and weaknesses of the study are discussed with recommendations for the future suggested.

Introduction

The National Health Service (NHS) is the largest healthcare service and employer in the UK. Despite this, racial inequalities[1] faced by staff and service users have been acknowledged within the research literature for many years (Esmail & Carnall, 1997; Kapadia et al., 2022; Kline, 2014; Wight, 2022). This has included the NHS being described as 'systemically racist', which is defined as a society or organisation set up in a way that ethnic minority groups are disadvantaged or mistreated (Adebowale & Rao, 2020; Beagan et al., 2022; Iacobucci, 2020; Kline, 2014; Ross et al., 2020). Attempts to address racial inequality have also spanned many years (Esmail & Carnall, 1997; NHS England, 2022a) with equality targets and the core values of 'respect' and 'inclusion' strived for. Nonetheless, racial inequality remains an experience for ethnic minority staff and service users (Bamrah & Chakravorty, 2022).

It has been shown that racial discrimination can impact staff well-being and productivity, with research finding that experiences of racial discrimination are associated with lower job satisfaction and increased stress levels (Deitch et al., 2003; Kaltiso et al., 2021). Racial disparities in the NHS workforce have been found across several major areas including recruitment, training, career progression, and unfair treatment (Adebowale & Rao, 2020; Iacobucci, 2020). The Workforce Race Equality Standard (WRES) was initiated in 2015 with the aim of ensuring ethnic minority staff would receive "*equal access to career opportunities and receive fair treatment in the work place*" (NHS England, 2022a). Data is collected against nine key indicators of racial equality. The percentage of ethnic minority staff is then compared to White staff across areas of representation, recruitment, training, and

[1] Racial inequality was considered the most appropriate term by people with lived experience interviewed in this study to encompass the range of implicit, explicit, and systemic racial and ethnic discrimination experienced by ethnic minority people in the UK.

disciplinary action. The latest data from the WRES (NHS, 2022) concluded that, whilst improvements had been made, there were still clear inequalities across all areas of the workforce, including the highest ever self-reported level of discrimination since the first WRES report in 2015.

The impact of racial inequality extends to NHS service users. Not only is exposure to racial discrimination a risk for poorer physical and mental health outcomes (Cave et al., 2020; Paradies et al., 2015), but ethnic minority people are also more likely to have negative healthcare related experiences (Ben et al., 2017). For example, compared to White British service users, ethnic minority users, in particular men, are more likely to be detained under the mental health act (Crown, 2018; Mann, 2014; Williams & Bunn, 2022), and women are at higher risk of maternal morbidity (Nair et al., 2014) . The serious case inquiry into the death of David Bennett, a man from an Afro-Caribbean background who was killed whilst being restrained on an NHS inpatient ward, revealed how complex, prevalent, and devastating the impacts of systemic racism could be (Norfolk, Suffolk and Cambridgeshire & Strategic Health Authority, 2003).

Looking Back and Moving Forward

The pervasive nature of racial inequality in the NHS brings into question the helpfulness of the strategies employed to address it (Hassen et al., 2021) which have been considered by many to be limited in their impact (Gay & Bamford, 2007; Kar, 2020) and this has been discussed by a variety of stakeholders across a range of forums including policies and reports (NHS, 2022; Ross et al., 2020), blogposts (Adeyemi, 2019; Chand, 2018; Dyer, 2019), videos (NHS England, 2016) and leadership initiatives (Jolliff, 2019). Key documents offer guidance, recommendations, and targets for addressing racial inequality in the NHS on an individual-level (e.g., training to increase staff's knowledge of different cultures or of their

own implicit biases), organisational-level (e.g., collecting data to capture racial disparities like the number of ethnic minority staff who accessed additional training compared to White staff), and policy-level (e.g., drives to achieve equality targets) (Adeyemi, 2019; Hassen et al., 2021; NHS, 2022; NHS England, 2022b; The NHS Staff Council, 2021). Despite this, Adeyemi (2019) summarises the sense that whilst many strategies have been implemented to address racial inequality, they have not been sufficient.

> *"...what hashtag, conference, workshop, diversity training, evidence-gathering process have we not engaged with to bring about change? What is that internalised barrier that keeps the NHS from addressing the problem once and for all?"* (Adeyemi, 2019).

Research into anti-racism interventions within healthcare settings have concluded that a heavy focus on individual-level strategies is largely insufficient within wider organisational contexts (Guschke & Christensen, 2021; Hassen et al., 2021). The implicit association test (Greenwald et al., 1998) has long revealed that even those who explicitly deny being consciously biased can hold unconscious biases, and these have the potential to influence behaviour (Tobon et al., 2021). Targeting individual-level factors (e.g., implicit biases) through strategies such as training may understandably seem like an appropriate solution. However, attempts to increase knowledge and awareness in isolation does not mean individual behaviour change will follow (Noon, 2018). Psychological studies have produced a plethora of theories on behavioural change, including change specifically within organisations, revealing that individual, relational, and broader level factors are important to consider (Vakola et al., 2004). Social Norms theory (Johnson, 2012) considers the importance of interpersonal influences on individual behaviours whilst Cognitive Dissonance Theory

(Festinger, 1957) and the Transtheoretical/Stages of Change Model (Prochaska & DiClemente, 1982) highlight the importance of individual motivation and readiness. When trying to initiate behaviour change Cognitive Dissonance Theory would argue that if new information goes against someone's prior opinions, it is likely that they will be driven to maintain the balance between their original attitudes and behaviours to reduce the discomfort that accompanies such dissonance. This may mean new information is dismissed and behaviour remains the same (Harmon-Jones, 2019). The Transtheoretical/Stages of Change Model further highlights that successful behaviour change varies depending on people's cognitive and affective states of readiness. Movement through the stages of change are likely to be more or less effective depending on the stage someone aligns with (Glanz et al., 2008). Furthermore, identifying and attempting to address individual bias can be considered limited if the organisational environment is not conducive to change, for example individual-level strategies applied to address racial inequality within a context of systemic racism (Noon, 2018). Indeed, focusing on individual-level factors without considering wider societal and systemic factors has been considered a "*weakness of contemporary approaches*" (Tate & Page, 2018).

The solution to this does not necessarily lie with combining individual-level strategies with policy- and organisational-level approaches as they too have their limitations. Firstly, without real life implementation, policy and organisational-level approaches can become tokenistic and inconsequential (Hassen et al., 2021). In addition, higher level approaches (e.g., aiming to increase the number of ethnic minority staff in senior roles) do not necessarily acknowledge the drivers of racial inequality (e.g., hierarchies of power and systemic racism). As such, there is a risk that instead they contribute to an avoidance of acknowledging the bottom line – that racism causes racism (Hassen et al., 2021; Tate & Page, 2018).

Furthermore, policy and organisational level strategies are designed, measured, and reported, in a way that inherently represents the priorities, ideas, views, and resource availability of the NHS as an organisation. As such, attempts to address racial inequality could be considered innately influenced by organisational biases and limited in their representation of service user and staff views. As described by models of organisational change within the NHS, a lack of shared understanding, shared beliefs, and motivation can act as substantial barriers to change within organisations (NHS England, 2018).

'Design think' is an approach increasingly applied to health care innovation research and contrasts the top-down approach often employed by health care organisations (Chan, 2018; McLaughlin et al., 2019; Scott et al., 2003). It suggests that strategies applied by organisations are likely to be ineffective if they do not incorporate the needs and feedback of employees and users (Altman et al., 2018). Instead, building a comprehensive understanding of these stakeholders' wants, needs, and barriers is considered key to developing effective solutions (Roberts et al., 2016). Further research into how racial inequality should be addressed in the NHS is considered a priority, with the current body of research exploring racial inequalities in the UK's health system inadequate (Adebowale & Rao, 2020; Salway et al., 2020; Surash, 2020). This is especially true of research that provides a platform for a breadth of perspectives to be heard (Anekwe, 2020). At the time of this project, research exploring the perspectives of NHS staff and service users on how racial inequality should be most effectively addressed was a gap in the evidence base This study aimed to examine these perspectives in a novel way by using Q-methodology.

Overview of Q-methodology

Q-methodology is a mixed-methods approach (Ladan et al., 2018) developed with the aim of studying subjectivity (e.g. people's views, attitudes, values, etc) flexibly, rigorously,

and within a scientific framework (Combes et al., 2004; Herrington & Coogan, 2011; QMethod Software, 2022; Stephenson, 1935). Q-methodology is inherently explorative and aims to identify perspectives on a given topic at a given time. This does not allow for the direct generalisation of results to a wider population (i.e., to infer that a certain percentage of the population share a certain perspective). However, generalisation is not considered necessary for the results of a Q-study to be meaningful or applicable across different contexts (Watts & Stenner, 2012). Simply witnessing the existence of particular perspectives and how they interact, can offer key insights (Skorpen et al., 2012). Compared to other methods used for studying subjectivity (e.g., interviews and questionnaires), Q-methodology's strength lies in its ability to concisely present data whilst facilitating exploration of both collective and individual differences (Herrington & Coogan, 2011; Ladan et al., 2018). Additionally, when studying subjectivity around sensitive topics, research instruments such as Likert scales are particularly prone to socially desirable response styles from participants (Kowalski et al., 2018; Schuetzler et al., 2018; Thielmann et al., 2016; Willburne et al., 2021). However, unlike Likert scales, Q-methodology prevents participants from rating an unlimited number of statements to the extreme positive or negative end of a scale. Indeed, it has been shown that in psychological research, Q-methodology is less affected by socially desirable responding (Fluckinger, 2014). This was particularly important to consider as it was hypothesised that socially desirable responding could indeed be a threat to validity for this study. Furthermore, questionnaires are often pre-determined before participants provide their perspectives. The content of the statements used in Q-methodology on the other hand are derived from stakeholders themselves.

Q-methodology has already been used to explore a range of related topics including staff and service user perspectives on health conditions (Broderick et al., 2017; Forrest-Bank

& Jenson, 2015; Skorpen et al., 2012), healthcare treatments (Combes et al., 2004; Gough et al., 2014), healthcare services (Shabila et al., 2014), health inequalities (McHugh et al., 2019; Sylvester, 2000), and inequality in the workplace (DeCourville & Hafer, 2001). Overall, Q-methodology was considered a particularly beneficial approach for this study.

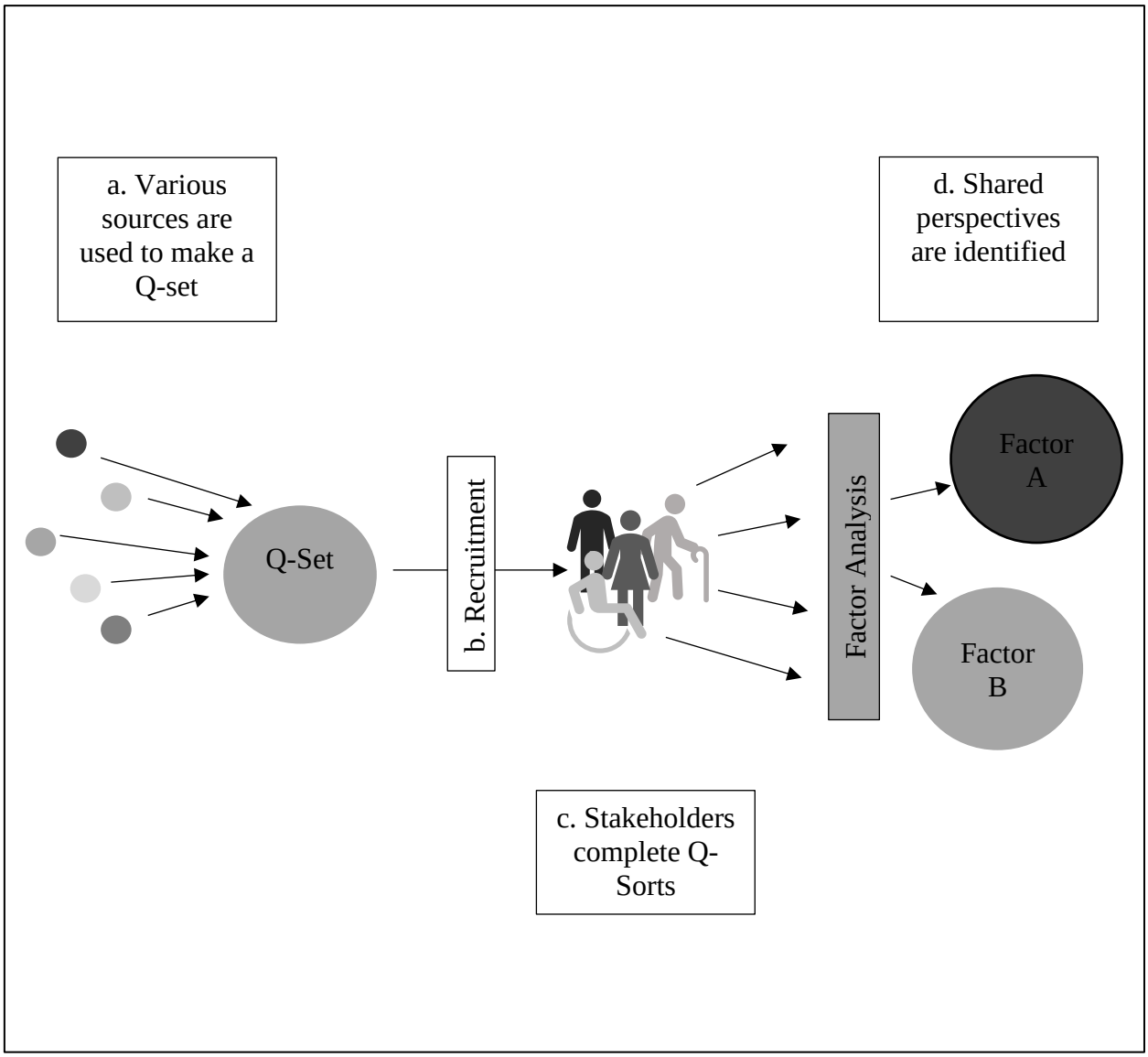
Method

This study was approved by the Health Research Authority (Appendix A). Q-methodology is comprised of four main stages: (a) developing a Q-set, (b) recruiting participants, (c) completing a Q-sort, and (d) data analysis. This process is detailed below and depicted in Figure 1.

a. A Q-set is a bank of statements or perspectives intended to reflect the spectrum of views on a given topic. It is important that the Q set is representative, as participants need to be able to authentically express their subjectivity through the Q-set (Herrington & Coogan, 2011). Perspectives that make up a Q-set can be obtained from a range of sources (e.g., policy documents (Gough et al., 2014), social media comments (Leonard et al., 2021), photographs (Combes et al., 2004), and stakeholders through interviews or questionnaires for example (McHugh et al., 2019)). The most important factor to consider when developing a Q-set is that sources are useful and relevant for addressing the research question (Watts & Stenner, 2012).In a final Q-set, 40 – 80 statements are considered enough to sufficient range of perspectives on a topic whilst remaining manageable for participants (Watts & Stenner, 2012).

Figure 1.

An Image to Depict the Q-Methodology Process



b. Once a Q-set has been developed, participants are recruited for a card sorting task called a Q-sort. The most important recruitment consideration is that participants have meaningful perspectives pertaining to the research question. Whilst this may often mean capturing perspectives from people with the strongest viewpoints, it may also mean capturing perspectives from people who have little enthusiasm or expertise on a topic (Watts & Stenner, 2012). Purposive sampling is routinely used in Q-

methodology, meaning participants are deliberately sought and recruited based on the belief that they offer an important perspective (Shabila et al., 2014).

c. Recruited participants are then required to complete a Q-sort which is a card sorting activity using the statements in the Q-set. The statements of the Q-set are sorted by each participant based on their perspectives. This gives the Q-set meaning and is a vessel for the expression of participants' subjectivity (Herrington & Coogan, 2011).

Finally, data is analysed using factor analysis. Factor analysis is used to explain as much study variance as possible using the fewest number of factors (Tinsley & Tinsley, 1987). In Q-methodology, each factor is considered an opinion group. An overview of the procedure used in this study can be found in Figure 2 and is also described in detail below.

Q-Set Development

To maximise the breadth and depth of the Q-set for this study, statements were gathered from a range of sources including written materials and interviews with stakeholders. Written materials (e.g., media articles, policy documents, and research papers) were reviewed with key perspectives and themes extracted as quotes and logged in an encrypted Excel spreadsheet. Interviews were completed with five stakeholders. In Q-methodology, stakeholders consulted at this stage are not considered study participants, but instead are used to develop the study itself. . Stakeholder interviewees were considered to have lived experience or offered another important perspective on addressing racial inequality within the NHS to try and obtain a wide range of perspectives. The demographics for recruited interviewees are shown in Table 1. Interviewees were aged between 26 and 65 with the majority of interviewees identifying as a man. The 5 interviewees self-identified across Black British – African, Black British – Caribbean, and White British. Interviewees held a range of positions within the NHS that spanned across clinical, expert by experience, and managerial

roles with the majority of them having been in their roles for over 10 years. Interviewees were recruited through dissemination of a participant information sheet via local collaborators. Those who were interested contacted the principal researcher directly. Interviewees were provided with a consent form and demographic questionnaire prior to the interview, as well as a debrief form afterwards. Interviews were semi-structured (Appendix B) and completed online via Zoom. Interview lengths ranged from 45 minutes to 1 hour 15 minutes. Interviews were recorded to allow for the statements to be extracted after the interview. Recordings were relistened to in their entirety, and identified statements were logged into the encrypted Excel spreadsheet.

In total, 295 statements were collated. One hundred and fifty-seven statements originated from media sources and 138 statements originated from interviews with stakeholders. To refine the initial 295 statements, the principal researcher and a second rater independently combined or removed duplicate statements until a consensus was reached. Each rater then grouped statements into categories. The categories attributed to each statement were then compared and consolidated by the principal researcher. The categorisation of statements helped to assess the balance of the final Q-set, and whether it was representative of the initial bank of statements. The principal researcher and two other researchers also reworded Q-statements to support their clarity. The final Q-set was made up of 65 Q-statements.

Figure 2

Q-Methodology Procedure Flow Chart

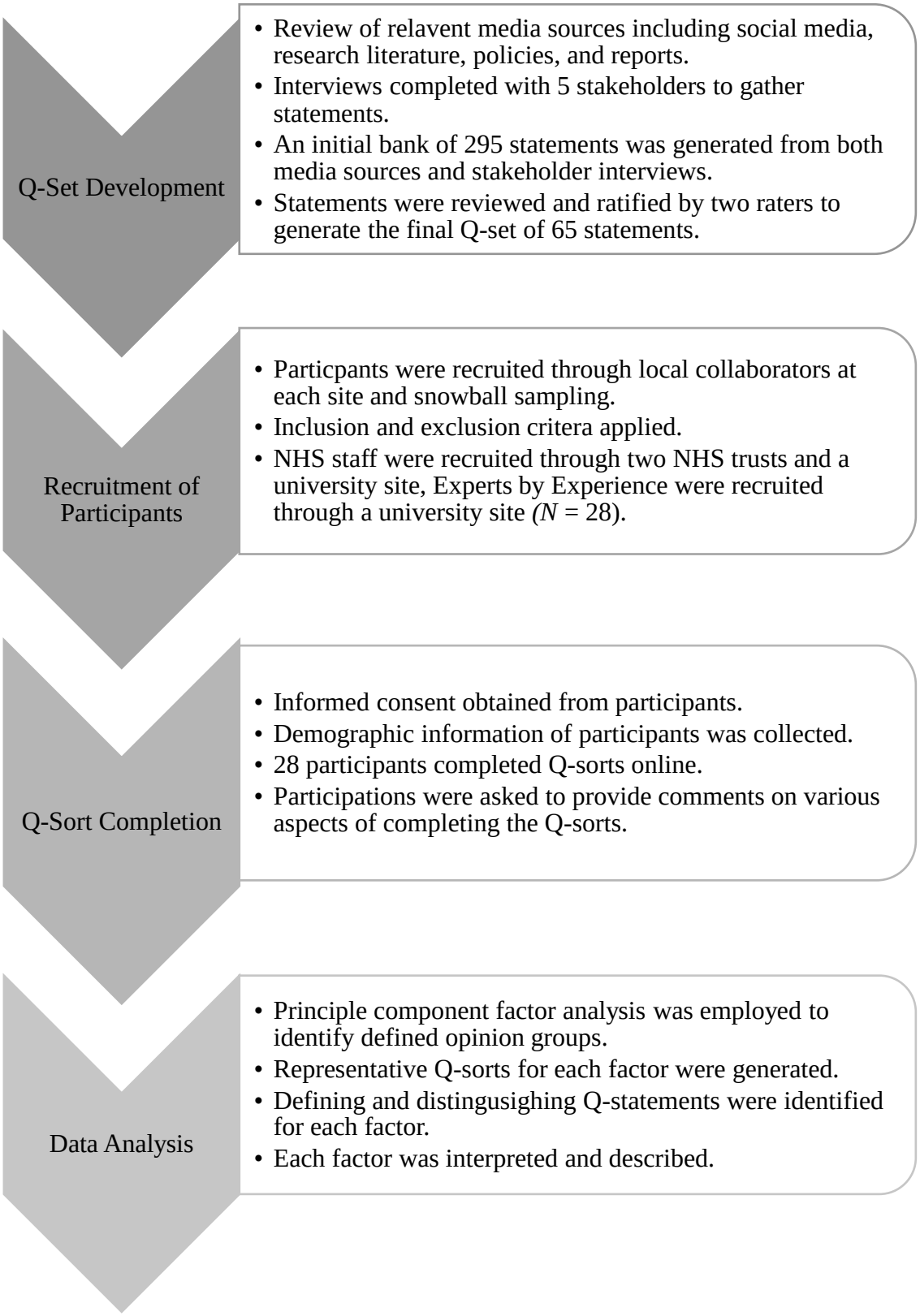


Table 1

Summary of Interviewee Demographics (N = 5)

		n
Age	26 – 35	1
	36 – 45	1
	46 – 55	1
	56 – 65	2
Gender	Man	4
	Woman	1
Ethnicity	Black British – African	1
	Black British - Caribbean	2
	White – English/Welsh/Scottish/Northern Irish/British	2
Job Role	Psychologist	1
	Corporate Services	1
	Management	1
	Expert by Experience	1
	Consultant Neuropsychologist	1
Length of Job Role	3 – 4 years	1
	5 – 10 years	1
	>10 years	3

Participants

Participants recruited to complete a Q-sort were NHS staff (recruited through two NHS Trusts and a university) and NHS Experts by Experience (EbE; recruited through the same university). Inclusion and exclusion criteria outlined that staff must have been in their current NHS roles for at least 3 months whilst EbEs must have had experience of using an NHS service or of caring for someone who had used an NHS service. All participants needed access to a computer and the internet. Recruiting participants who did not speak or understand written English, or who would need assistance to participate, was unfortunately beyond the scope of this research. Participants were recruited through local collaborators (e.g., research and innovation departments, expert by experience group coordinators, lead clinicians, etc)

who forwarded a template email, PIS, and web link to access the Q-sort online. As interested participants could independently and anonymously participate, it was not possible to calculate a response rate.

Q-Sort Completion

Whilst traditionally Q-sorts are completed in person with a researcher present, the COVID-19 pandemic and social distancing requirements at the time of this study meant that an online Q-sort was considered safer and more accessible. An online Q-sort was created using Q-Sortware software (Pruneddu, 2011). Online Q-sorts using this software have been shown to be reliable (Pruneddu, 2013) and good alternatives to in person testing (Watts & Stenner, 2012). The Q-Sortware software facilitated the creation, distribution, and collection of Q-Sort data online. The online Q-sort was accessed by all recruited participants through one anonymous web link. Once participants accessed the link, they were presented with a participant information sheet and were required to complete a consent form and demographic questionnaire. Participants could then view instructions on how to complete the Q-sort activity before being directed to the task.

Card Sorting Activity

Participants completed the card sorting activity in two stages. In stage 1, participants were presented with each of the final 65 Q-statements that made up the Q-set. Each individual Q-statement was shown to participants on a virtual card. Participants were required to sort each card based on how effectively they felt it answered the prefix ‘To effectively address racial inequality in the NHS…’ The category options were ‘most effective’, ‘least effective’ and ‘unsure/indifferent/mixed feelings’. There were no limits for how many cards could be attributed to each category at this stage. The purpose of this initial card sort was to support participants with organising their cards for the final card sort.

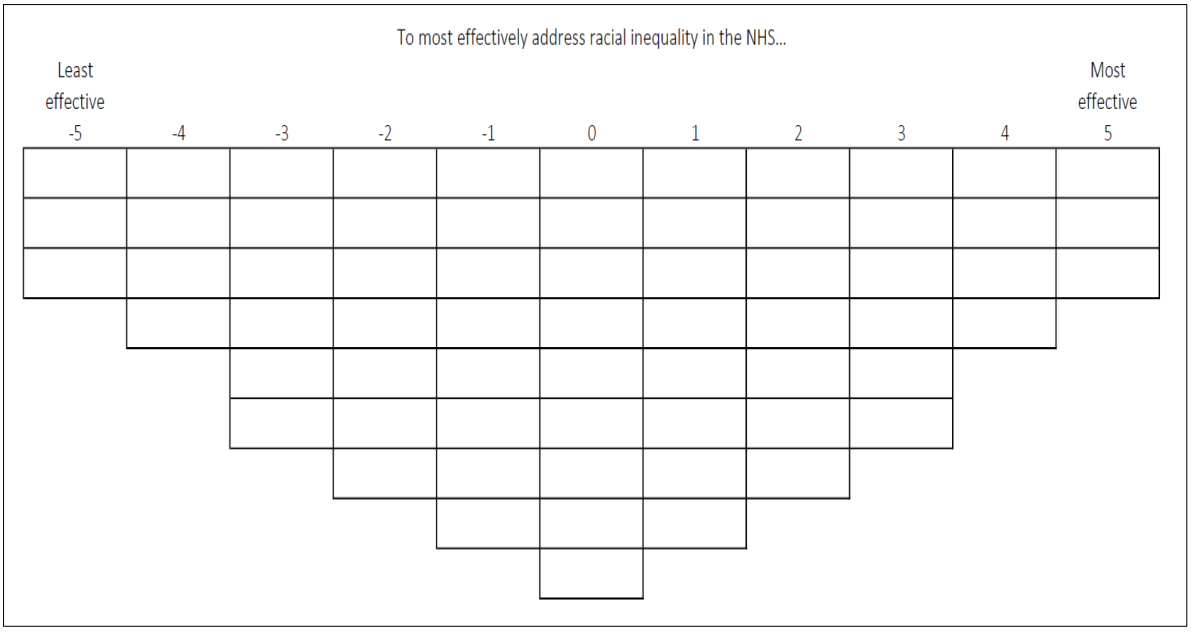
In stage 2, the cards were displayed to participants in the three categories as they had sorted them in step 1. Participants were asked to further sort the 65 Q-statements cards into a sorting grid depending on how effective they thought that solution was. The distribution grid used in this study is represented in Figure 2. The grid contained 65 spaces, one for each Q-statement in the Q-set, arranged across an 11-point scale from 'least effective' (-5) to 'most effective' (+5). This scale was chosen as, although Q-sets containing 60 items and above are recommended to have 13 ranks, a shortened scale is advised for complex topics to reduce the burden on participants (Brown, 1980). A quasi-normal forced-choice distribution was chosen for this study. Forced-choice distributions are standardised across participants, facilitating clearer comparisons, and reducing participant burden (Watts & Stenner, 2012). This meant participants had to allocate a certain number of Q-statements to each rank. For example, participants were required to attribute 6 Q-statements to the ranks ±3. Participants were able to move the Q-statement cards around the grid but had to ensure that the correct number of cards were allocated to each rank before they could finalise their Q-sort. The mean time taken to complete the Q-Sortware procedure was 31 minutes with a range from 14 to 80 minutes.

Post-Sort Questionnaire

When completed face to face, post-sorting interviews are often administered following a Q-sort. The insight gained from these interviews is used to aid interpretation of the results, increases study validity, and reduces researcher bias when interpreting the results (Gallagher & Porock, 2010). To ensure that this valuable information was still captured, a post-sort questionnaire was presented after participants had finalised their Q-sorts. Participants were then asked a mixture of open and closed questions about the Q-statement

Figure 3

Forced-Choice Quasi-Normal Frequency Distribution Used in this Study



cards (e.g. ‘were there any cards that surprised you?’) and about the way they completed the card sort (e.g. ‘what made you rank cards as ‘least effective’’).

Data Analysis Strategy

Each participant’s Q-sort created a unique array detailing which Q-statements were assigned to which rank. In this way, each Q-statement was assigned a number between -5 and +5 for each participant. The arrays for all 28 Q-sorts were collated. The “QMethod” package (Zabala, 2014) of the R programming language (RStudio Team, 2020) was used to analyse the data and extract opinion groups using principal component factor analysis. In Q-methodology, factors represent opinion groups on how to address racial inequality most effectively in the NHS.

Results

Characteristics of the Participants

The demographics of participants who completed the Q-sorts can be found in Table 2. Chi- square tests were used to explore these demographic characteristics. Chi-square tests revealed that significantly more respondents identified with the White British ethnic category

than any other ethnic category. Although, when compared to the general working age population, there were significantly more participants who identified with an ethnic minority group than would be expected by chance ($X^2 = 4.938$, $p = .026$; NHS Digital, 2019). There were significantly fewer experts by experience compared to NHS staff. There was an almost equal proportion of those identifying as women and men. Around two thirds of respondents had been in their job roles for less than 6 years.

Identifying the Number of Factors (Opinion Groups)

To identify the number of factors in the final solution, Q-sorts were factor-analysed using principal components analysis in the "QMethod" package (RStudio Team, 2020; Zabala, 2014). A factor's statistical strength and explanatory power is indicated by its eigenvalue. The higher the eigenvalue, the more variance that factor explains (Watkins, 2018). The Kaiser-Guttman Criteria states that factors with eigenvalues of less than 1 account for less variability than a single variable, in this case than a single Q-sort, and are not considered useful in a final solution (Silva et al., 2020; Watts & Stenner, 2012). Eight of the 27 factors had eigenvalues greater than 1. Table 3 details the eigenvalues, percentage variability, and cumulative variability of Factors 1 through 9 (subsequent factors are omitted as the eigenvalues were less than 1).

Table 2

Summary of Q-sort Respondent Demographics (N = 28)

		n	X^2	*p*
Age	18 - 25	2		
	26 – 35	9		
	36 – 45	8		
	46 – 55	3		

		n	X^2	*p*
	56 – 65	6	6.643	.16
Gender	Man	13		
	Woman	15	0.167	.68
Ethnicity	Asian or Asian British - Indian	1		
	Asian or Asian British - Pakistani	1		
	Asian or Asian British – Any other Asian	1		
	Black or Black British - Caribbean	1		
	Mixed or Multiple Ethnic backgrounds –	1		
	Mixed or Multiple Ethnic backgrounds –	1		
	White – any other white background	2		
	White – English/Welsh/Scottish/Northern	20	89.14	<.001
Job Role	Clinical Psychologist	9		
	Expert by Experience	4		
	Psychological Practitioner	4		
	Allied Health Professional	3		
	Administration	1		
	Assistant Psychologist	1		
	Doctor	1		
	Healthcare Scientist	1		
	Management	1		
	Nurse	1		
	Psychotherapist	1		
	Researcher	1	27.71	.003
Length of Job Role	6-11 months	8		
	1 – 5 years	10		
	6 – 10 years	4		
	11 – 20 years	3		
	21 – 30 years	2		
	31 – 40 years	1	13.57	.018

Using the Kaiser-Guttman Criteria in isolation can result in the over extraction of factors, impacting the helpfulness of the solution (Morton & Altschul, 2019). To avoid over extraction and to maximise interpretability, more than two flagged participants per factor was required (Brown, 1980). Table 3 and Table 4, show that the first 3 factors have eigenvalues greater than 1 and more than two participants are identified as exemplifying each factor. As such, a 3-

factor solution was retained. This 3-factor solution explained 51.11% of the study variance. This was considered appropriate given as solutions explaining 35-40% of the study variance are considered ‘good’ (Kline, 1994). The 3-factor principal components solution was rotated to aid interpretability (Watts & Stenner, 2012). There are a variety of methods available for factor rotation which can be broadly summarised as either oblique or orthogonal rotations. When using oblique rotations factors are allowed to correlate. Orthogonal methods on the other hand produce maximal separation between factors (Goretzko et al., 2021). An orthogonal rotation was considered most appropriate for this analysis where differentiated opinion groups were being sought (Osborne, 2019). The orthogonal “varimax” method was chosen for this study (Dilbeck, 2017). A summary of the varimax rotated factors is given in Table 5.

Table 3

Eigenvalues and Percent Variability for Factors 1 Through 9

Factor	Eigenvalues	Percent Variability	Cumulative Variability
1	9.9315	35.47	35.47
2	2.3784	8.4941	43.964
3	2.0012	7.1472	51.111
4	1.679	5.9966	57.108
5	1.1799	4.214	61.322
6	1.1103	3.9654	65.287
7	1.0581	3.7791	69.066
8	1.0013	3.576	72.642
9	0.857	3.0607	75.703

Table 4

Table Of Factor Loadings for The First Eight Factors for The Unrotated Principal Components

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8
ID1	0.060112	**0.720163**	0.203019	-0.046454	0.250709	0.057642	0.026889	0.064526
ID2	-0.05024	0.194784	0.378155	0.278488	0.225024	0.595082	-0.080109	0.292799
ID3	-0.10927	**0.839776**	0.088186	0.067757	-0.064666	0.148498	-0.105289	0.166009
ID4	0.484244	0.430983	0.100257	0.156178	0.257884	0.131857	0.318014	-0.17235
ID5	0.229033	0.183404	0.010399	0.360187	**0.633523**	0.125687	0.143619	-0.090177
ID6	0.126107	0.390855	0.16749	0.562699	0.375027	0.064306	0.258924	0.127013
ID7	0.49414	0.304006	0.485142	0.124523	-0.149682	-0.104716	0.254285	0.150604
ID8	-0.061821	**0.634563**	0.10789	0.278845	0.163132	0.178568	0.42512	-0.111737
ID9	0.122478	0.324808	0.178191	0.469754	0.634247	0.017469	-0.137038	0.152819
ID10	0.200418	0.079751	0.089204	**0.793836**	0.098974	0.161225	0.061977	-0.096672
ID11	0.259315	0.573079	0.129442	0.370108	0.22785	0.203075	0.144149	-0.002097
ID12	0.212609	0.050692	**0.725083**	0.336902	-0.058871	0.123644	-0.017687	-0.23406
ID13	0.445377	0.112963	0.184084	0.463365	0.295962	0.318132	0.32631	0.122057
ID14	0.052531	0.102282	0.41742	0.035137	0.258759	0.519489	0.466364	0.164349
ID15	0.259392	0.234153	0.410511	0.174395	0.528986	0.118283	0.385466	0.064253
ID16	0.152241	0.299755	-0.329313	-0.056154	0.30174	0.526027	0.178056	-0.279581
ID17	**0.607249**	-0.148244	-0.033763	0.360303	0.219504	0.043078	0.209824	0.153824
ID18	0.566683	0.056045	0.270044	0.416471	0.17485	0.01425	-0.021476	0.311738
ID19	0.270978	0.237304	0.079533	0.247287	0.119219	**0.706336**	0.021872	0.059125
ID20	0.117984	0.265343	**0.740916**	-0.094736	0.023746	0.138531	0.105099	-0.121091
ID21	0.345471	0.380618	-0.041865	0.367399	-0.026495	0.32532	0.245576	0.438863
ID22	0.247958	0.058034	0.06112	-0.004638	**0.798194**	0.282398	-0.116034	0.05354
ID23	0.305967	0.143952	-0.097383	-0.045926	0.065092	0.09581	0.201475	**0.719285**
ID24	0.165736	0.057447	0.117727	0.110279	-0.086096	0.009094	**0.823817**	0.167486
ID25	**0.805065**	0.060777	0.062674	0.117095	0.116418	0.297525	-0.105155	-0.001646
ID26	**0.790252**	0.013742	0.042005	0.039738	0.191362	0.012276	0.164483	0.152244
ID27	-0.104191	0.050641	**0.743198**	0.093908	0.334422	-0.033479	0.118573	0.161031
ID28	0.265011	0.480914	0.050484	0.433026	0.16421	0.240043	0.182226	0.177956

Note. Flagged respondents for each factor are shown in bold.

Describing the Factors in Terms of the Q-sorts and Q-statements

The Q-sorts most associated with each of the final three factors were identified to support the distinguishing of each opinion group. Factor loadings, the correlation coefficient between each Q-sort and each factor, were used to identify the Q-sorts most associated with each factor. A distinguishing Q-sort had to be both correlated with its corresponding factor score and the square of its factor loading had to be higher than the sum of the square factor loadings for the other factors (i.e., $fx^2 > (fy^2 + fz^2)$; Zabala, 2014). Q-sorts that satisfied these criteria were 'flagged' against the Factor that they were highly associated with. If a Q-sort did not load highly onto any of the factors, or loaded highly onto more than one factor, it was not flagged. Twenty-four of the 28 Q-sorts loaded significantly and exclusively onto one of the three factors. The remaining four participants were excluded from the interpretation. Factor loadings and the flagged Q-sorts are shown in Table 6. To consider the relationship between each Q-statement and each of the three factors, *z*-scores were calculated. *Z*-scores are the weighted average of the scores for each Q-statement and indicate the direction and strength of such relationship. To aid interpretability, *z*-scores were converted to factor scores. Factor scores are *z*-scores rounded towards the units of measurement used in the original Q-sorts. Both *z*-scores and factor scores are shown in Table 7.

Table 5

Eigenvalues and Cumulative Variability Associated with the Varimax Rotated Factors

Factor	eigenvalues	Percent Variability	Cumulative Variability	Std Error
1	5.688854	20.31734	20.31734	0.1428571
2	5.219746	18.64195	38.95929	0.1740777
3	3.402498	12.15178	51.11106	0.2425356

Table 6

Three Factor Solution Factor Loadings per Q-sort

Q-Sort	Factor 1	Factor 2	Factor 3
Q-Sort 1	**0.6463**	-0.0375	0.2363
Q-Sort 2	**0.541**	0.1548	0.3302
Q-Sort 3	**0.6925**	-0.2142	0.1251
Q-Sort 4	0.4883	0.4604	0.2082
Q-Sort 5	**0.5181**	0.4426	0.0593
Q-Sort 6	**0.6103**	0.3773	0.3021
Q-Sort 7	0.097	0.3846	**0.6058**
Q-Sort 8	**0.7291**	0.0297	0.2527
Q-Sort 9	**0.5895**	0.3358	0.1936
Q-Sort 10	0.3426	**0.4431**	0.1843
Q-Sort 11	**0.6788**	0.3234	0.2248
Q-Sort 12	0.0955	0.1569	**0.7215**
Q-Sort 13	0.4165	**0.6842**	0.2925
Q-Sort 14	0.4212	0.2477	0.464
Q-Sort 15	0.4634	0.4152	0.4933
Q-Sort 16	**0.5681**	0.186	-0.3284
Q-Sort 17	0.0256	**0.7866**	0.0634
Q-Sort 18	0.138	**0.6814**	0.3314
Q-Sort 19	**0.5536**	0.3869	0.0677
Q-Sort 20	0.2106	-0.0391	**0.7252**
Q-Sort 21	0.4837	**0.5235**	0.0942
Q-Sort 22	**0.4388**	0.3923	-0.025
Q-Sort 23	0.1423	**0.4487**	-0.0103
Q-Sort 24	0.0787	0.3332	0.3415
Q-Sort 25	0.1396	**0.7303**	0.0509
Q-Sort 26	0.0163	**0.7724**	0.1109
Q-Sort 27	0.1709	-0.0027	**0.7384**
Q-Sort 28	**0.6188**	0.4126	0.1665

Note. Flagged Q-Sorts for each factor are shown in bold.

Table 7

Z-scores and Factor Scores per Q-Statement for Each Factor

Q-statement	Factor 1		Factor 2		Factor 3	
	z-score	f-score	*z*-score	f-score	*z*-score	f-score
Q1	0.2968	0	-0.189	-1	1.2432	3
Q2	0.528	1	-0.867	-2	1.9389	5
Q3	0.6122	2	1.0754	3	1.4361	4
Q4	0.2438	0	0.7442	2	-0.397	-1
Q5	0.4832	1	-0.082	0	-0.763	-2
Q6	-0.6059	-2	-0.394	-1	-0.063	0
Q7	-1.8525	-5	-0.286	-1	-1.143	-3
Q8	1.4749	5	1.5731	5	0.6459	2
Q9	-1.6211	-4	-1.492	-4	-1.833	-5
Q10	1.232	4	1.1611	3	0.4846	1
Q11	0.5462	1	1.5225	4	0.5933	2
Q12	0.5249	1	1.2004	3	0.3948	1
Q13	-1.2002	-3	0.1731	0	-0.403	-1
Q14	-0.8828	-2	-0.158	0	-0.748	-2
Q15	-0.8494	-2	-0.246	-1	-0.959	-3
Q16	-0.9819	-3	0.4987	1	1.5254	4
Q17	0.8494	3	1.9147	5	-0.023	0
Q18	0.2152	0	1.4888	4	0.157	0
Q19	0.5154	1	-1.485	-4	0.592	2
Q20	-2.5527	-5	-1.862	-5	-1.509	-4
Q21	0.6343	2	0.8711	2	-0.019	0
Q22	-0.6199	-2	-1.241	-3	0.451	1
Q23	0.2941	0	0.3341	1	0.063	0
Q24	0.0961	-1	1.987	5	0.4457	1
Q25	0.876	3	1.1394	3	1.2601	3
Q26	1.3707	4	1.2783	4	0.2257	0
Q27	-0.2867	-1	-0.109	0	0.911	3
Q28	0.2952	0	0.6807	2	-0.37	-1
Q29	-0.6535	-2	-1.393	-3	-2.372	-5
Q30	0.4558	1	-0.707	-2	1.5112	4
Q31	-1.7495	-4	-1.189	-3	-1.814	-5
Q32	-1.194	-3	-1.617	-4	-0.643	-2
Q33	-1.1551	-3	-0.324	-1	-0.221	-1
Q34	-1.2296	-3	-1.627	-4	-1.208	-3
Q35	1.1302	3	0.5662	1	0.8945	3
Q36	0.3497	0	-0.996	-3	0.6681	2
Q37	0.6718	2	0.9688	3	-0.618	-2
Q38	0.309	0	1.3484	4	-0.514	-2
Q39	0.9737	3	0.9518	3	-1.061	-3
Q40	0.6612	2	-0.435	-2	0.8873	2
Q41	0.4032	1	0.9372	2	-0.918	-2
Q42	1.0582	3	-0.856	-2	0.4096	1
Q43	1.3898	5	-0.379	-1	0.9217	3

Q-statement	Factor 1		Factor 2		Factor 3	
	z-score	f-score	*z*-score	f-score	*z*-score	f-score
Q44	0.0412	-1	-0.43	-2	-0.378	-1
Q45	-0.1441	-1	0.0945	0	-1.34	-4
Q46	-0.2836	-1	0.3356	1	-0.14	0
Q47	0.6344	2	0.0225	0	1.053	3
Q48	0.08	-1	-1.327	-3	0.391	1
Q49	0.7713	2	-0.081	0	-0.298	-1
Q50	-0.9788	-2	-0.748	-2	-0.959	-3
Q51	0.373	1	0.2791	1	0.2093	0
Q52	0.869	3	0.7674	2	0.4096	1
Q53	0.0354	-1	-0.263	-1	-0.665	-2
Q54	-1.404	-3	-1.914	-5	-1.774	-4
Q55	-0.0192	-1	0.0488	0	-0.982	-3
Q56	0.1324	0	-0.145	0	0.1689	0
Q57	1.5338	5	0.3505	1	0.6788	2
Q58	0.6231	2	0.4993	1	1.9787	5
Q59	-0.3763	-2	-1.189	-3	1.2745	4
Q60	1.3111	4	0.3713	1	-0.447	-1
Q61	1.2224	4	0.6809	2	0.4027	1
Q62	-1.4616	-4	-1.638	-5	-1.777	-4
Q63	0.2507	0	-0.769	-2	-0.148	-1
Q64	-2.4939	-5	0.8955	2	1.6303	5
Q65	-1.7728	-4	-0.325	-1	0.6496	2

Note. f-score = factor score

Interpreting the factors

The final stage of the analysis requires interpretation of each of the three opinion groups. The factor scores of the 'flagged' Q-sorts for each factor were averaged to create one representative Q-sort, encapsulating the shared viewpoint of that factor (Donner, 2001). The factor scores of the representative Q-sorts for each factor are shown in Table 8. Each Factor was also explored in detail and given a title to summarise the way in which it was defined by the Q-statements. Q-statements that were considered 'consensus' (i.e., a Q-statement whereby there were no statistically significant differences between the *z*-scores across each factor) and 'distinguishing' (i.e., Q-statements whereby the *z*-scores were statistically significantly different across each factor) were also explored.

Table 8

Q-statements and Their Factor Scores for Each Factor

Q-Statement		F1	F2	F3
Q1	…a director of equality should be appointed in all NHS Trusts	0	-1	3
Q2	…a member of staff with an ethnic minority background should always be included on NHS interview panels	1	-2	5
Q3	…each NHS Trust should have a central equality and diversity team for staff to access resources when needed	2	3	4
Q4	…data on the ethnicity of all NHS service users and staff must be collected	0	2	-1
Q5	…independent organisations should be the ones assessing levels of racial inequality in the NHS	1	0	-2
Q6	…ethnic minority staff and service users should be considered ‘experts’ in equality and diversity	-2	-1	0
Q7	…NHS staff not born in Britain should have the opportunity to be educated on what different social mannerisms mean in British culture	-5	-1	-3
Q8	…ethnic minority staff should be better supported when racism is directed at them from service users	5	5	2
Q9	…broad brush solutions need to be applied as standard across the NHS	-4	-4	-5
Q10	…each member of staff must personally think addressing racial inequality is important	4	3	1
Q11	…ethnic minority staff and service users should be continually involved in decision making	1	4	2
Q12	…NHS leaders should be demanding more evidence about what is effective for addressing racial inequality	1	3	1
Q13	…all NHS staff should be trained in the routine collection of service user ethnicity	-3	0	-1
Q14	…the percentage of ethnic minority and White NHS staff should be the same as that of the general population	-2	0	-2
Q15	…each ethnic minority service user should have an assessment to see how accessible NHS services are for them	-2	-1	-3
Q16	…every member of staff should have generic equality and diversity training	-3	1	4
Q17	…greater involvement of ethnic minority people in healthcare research is needed	3	5	0
Q18	…interventions applied to address racial inequality will need to have measurable outcomes	0	4	0
Q19	…it must be accepted that it’s not just ‘a few bad apples’ who are racist	1	-4	2
Q20	…addressing racial inequality should be solely left to NHS leaders to deal with	-5	-5	-4
Q21	…it will be important to create a safe space for ethnic minority staff before expecting them to talk about racism	2	2	0
Q22	…all NHS staff should be required to attend regular team meetings where racial inequality is discussed	-2	-3	1
Q23	…more ethnic minority mentors should be available	0	1	0
Q24	…more research is needed on the health effects of racial inequality	-1	5	1
Q25	…multicultural perspectives must be incorporated into the culture of the NHS	3	3	3
Q26	…healthcare interventions must be appropriately adapted for the needs of ethnic minority cultures and communities	4	4	0
Q27	…NHS professionals should have specific training on racism that incorporates real life scenarios	-1	0	3
Q28	…services should have allocated workers to engage people from hard-to-reach ethnic minority communities	0	2	-1
Q29	…ethnic minority staff / service users should be provided with extra resources and opportunities to compensate for their disadvantage	-2	-3	-5
Q30	…staff need to accept they may not realise they’re racist	1	-2	4
Q31	…treatment should be withheld from service users who are racist to staff	-4	-3	-5

Q32	…Trusts that are not meeting equality targets should be penalised	-3	-4	-2
Q33	…personal views on racial inequality need to be put aside	-3	-1	-1
Q34	…all NHS staff should complete questionnaires about how comfortable they feel talking about race	-3	-4	-3
Q35	…racial inequality must be kept on the agenda of the NHS continually	3	1	3
Q36	…recruitment processes should place more importance on an applicant's underlying competencies over their qualifications	0	-3	2
Q37	…serious efforts are needed to ensure access to high quality interpreting services	2	3	-2
Q38	…research is required to understand where ethnic minority service users experience the most racial inequality whilst receiving NHS treatments	0	4	-2
Q39	…the efficacy of NHS treatments for ethnic minority groups must be reviewed	3	3	-3
Q40	…senior levels of the NHS should include more ethnic minority people	2	-2	2
Q41	…the outcomes of ethnic minority service users should be tracked across their entire healthcare journey	1	2	-2
Q42	…staff in the NHS must be antiracist (actively oppose racism)	3	-2	1
Q43	…staff must be willing to recognise their biases	5	-1	3
Q44	…stronger allyship from White staff is needed	-1	-2	-1
Q45	…ethnic minority service users' attitudes towards using NHS services should be surveyed	-1	0	-4
Q46	…interventions that have worked to address racial inequality in other large institutions should be used in the NHS	-1	1	0
Q47	…strategies must be completely focused on addressing racism that is deeply embedded in the NHS as an organisation	2	0	3
Q48	…all NHS teams should be assessed on how they respond to issues relating to race	-1	-3	1
Q49	…the existence of racial inequality in the NHS must first be accepted before progress can be made	2	0	-1
Q50	…we should celebrate the achievements of ethnic minority staff more	-2	-2	-3
Q51	…it should be acknowledged that medical diagnoses can be more stigmatising for certain ethnic minority communities	1	1	0
Q52	…the negative impact of poor socioeconomic factors on ethnic minority people's health should be acknowledged	3	2	1
Q53	…the NHS should provide preventative interventions for ethnic minority people, for example, support to access school, education and career development opportunities	-1	-1	-2
Q54	…the career development of ethnic minority staff must take priority over that of White staff	-3	-5	-4
Q55	…NHS services should be able to allocate funding individually to each service user to meet their needs in the way that feels most culturally appropriate	-1	0	-3
Q56	…the White workforce needs to be upskilled to understand the unspoken rules of ethnic minority cultures	0	0	0
Q57	…there must be measurable action not just conversations, research and policies	5	1	2
Q58	…there should be a zero-tolerance policy on racial inequality	2	1	5
Q59	…there should be compulsory questions about equality and diversity in job interviews	-2	-3	4
Q60	…healthcare treatments need to be tailored to effectively meet the needs of ethnic minority service users	4	1	-1
Q61	…the lived experience of ethnic minority staff and service users should be learnt from	4	2	1
Q62	…ethnic minority staff must be kept in their roles for the long term	-4	-5	-4
Q63	…the long history of racism in the NHS should not be forgotten	0	-2	-1
Q64	…every service user should be treated exactly the same	-5	2	5
Q65	…it is important NHS staff don't feel blamed or labelled as racist	-4	-1	2

Note. Factor scores are colour coded to aid interpretation. The colours represent a scale from dark red (for lowest ranked scores of -5), through orange, yellow, and light green, to dark green (for highest ranked scores of +5).

Consensus statements

If the *z*-scores for a Q-statement were not significantly different across the three factors, that statement was considered a consensus statement (Ramlo, 2008). Whilst consensus statements can increase the correlations between factors and make distinguishing factors more difficult (Skorpen et al., 2012), it can also be useful to identify shared perspectives across opinion groups. Keeping ethnic minority members of staff in their roles 'for the long term' (Q62) and 'broad brush solutions' (Q9) were ranked very low across all factors. Staff completing questionnaires about how comfortable they felt talking about race (Q34) and celebrating the achievements of ethnic minority staff (Q50) were also ranked relatively low across the board. On the other hand, acknowledging the impact of poor socioeconomic factors on the health of ethnic minority people (Q52) and incorporating multicultural perspectives into the culture of the NHS (Q25) were generally ranked relatively highly across all factors.

More ethnic minority mentors (Q23), placing ethnic minority staff and service users as 'experts' in equality and diversity (E&D) (Q6), acknowledging that medical diagnoses can be more stigmatising for certain communities (Q51), and upskilling the white workforce to understand the unspoken rules of ethnic minority cultures (Q56) were not prioritised in either direction across any of the factors. The factor scores for identified consensus statements are shown in Table 9.

Distinguishing All Statements

A statement was considered a 'distinguishing all statement' if there was a significant difference between the *z*-scores across all three factors for an individual Q-statement (Ramlo, 2008) Identified distinguishing all statements are shown in Table 10. Distinguishing all statements are explored further when considering each opinion group in turn.

Table 9

Consensus Statements and the Difference in z-scores Between Each Factor

Q-statement	F1-F2	F1-F3	F2-F3
Q6. …ethnic minority staff and service users should be considered 'experts' in equality and diversity.	-0.211	-0.543	-0.331
Q9. …broad brush solutions need to be applied as standard across the NHS.	-0.129	0.212	0.341
Q23. …more ethnic minority mentors should be available.	-0.040	0.231	0.271
Q25. …multicultural perspectives must be incorporated into the culture of the NHS.	-0.263	-0.384	-0.121
Q34. …all NHS staff should complete questionnaires about how comfortable they feel talking about race.	0.397	-0.022	-0.419
Q50. …we should celebrate the achievements of ethnic minority staff more.	-0.231	-0.020	0.211
Q51. …it should be acknowledged that medical diagnoses can be more stigmatising for certain ethnic minority communities.	0.094	0.164	0.070
Q52. …the negative impact of poor socioeconomic factors on ethnic minority people's health should be acknowledged.	0.102	0.459	0.358
Q56. …the White workforce needs to be upskilled to understand the unspoken rules of ethnic minority cultures.	0.277	-0.036	-0.313
Q62. …ethnic minority staff must be kept in their roles for the long term	0.176	0.315	0.139

*$p < .05$, **$p < .01$, ***$p < .001$

Table 10

Distinguishing All Statements and the Difference in z-scores Between Each Factor

Q-statement	F1-F2	F1-F3	F2-F3
Q1. …a director of equality should be appointed in all NHS Trusts	0.486*	-0.946***	-1.432***
Q2. …a member of staff with an ethnic minority background should always be included on NHS interview panels	1.395*	-1.411*	-2.806*

Q4. …data on the ethnicity of all NHS service users and staff must be collected	-0.5*	0.640*	1.141***
Q5. …independent organisations should be the ones assessing levels of racial inequality in the NHS	0.565*	1.246***	0.682*
Q7. …NHS staff not born in Britain should have the opportunity to be educated on what different social mannerisms mean in British culture	-1.567*	-0.709*	0.858**
Q16. …every member of staff should have generic equality and diversity training	-1.481*	-2.507*	-1.027***
Q17. …greater involvement of ethnic minority people in healthcare research is needed	-1.065***	0.872**	1.937*
Q22. …all NHS staff should be required to attend regular team meetings where racial inequality is discussed	0.621**	-1.071***	-1.692*
Q29. …ethnic minority staff / service users should be provided with extra resources and opportunities to compensate for their disadvantage	0.74**	1.718*	0.979**
Q30. …staff need to accept they may not realise they're racist	1.163*	-1.055***	-2.218*
Q38. …research is required to understand where ethnic minority service users experience the most racial inequality whilst receiving NHS treatments	-1.039***	0.823**	1.862*
Q41. …the outcomes of ethnic minority service users should be tracked across their entire healthcare journey	-0.534*	1.321***	1.855*
Q42. …staff in the NHS must be antiracist (actively oppose racism)	1.914*	0.649*	-1.266***
Q59. …there should be compulsory questions about equality and diversity in job interviews	0.813***	-1.651*	-2.463*
Q60. …healthcare treatments need to be tailored to effectively meet the needs of ethnic minority service users	0.94***	1.758*	0.818**
Q64. …every service user should be treated exactly the same	-3.389*	-4.124*	-0.735*
Q65. …it is important NHS staff don't feel blamed or labelled as racist	-1.448*	-2.422*	-0.974**

*$p < .05$, **$p < .01$, ***$p < .001$

Factor 1: 'Prioritise Responsibility and Reflection'

Factor 1 accounted for 20% of the total variance and was significantly associated with twelve Q-sorts (R1, R2, R3, R5, R6, R8, R9, R11, R16, R19, R22, and R28). Overall, Factor 1 appeared to reveal a perspective where staff willingness to take responsibility was considered

most effective for addressing racial inequality. The most strongly endorsed Q-statements in this factor (rounded *z*-score = ± 4/5) are shown in Table 11 along with their averaged factor scores.

Highly ranked distinguishing statements centred around individual-level strategies such as staff's willingness to address racial inequality (Q10: +4), to be 'antiracist' (Q42: +3), and to recognise their own biases (Q43: +5). Q-statements interpreted as a distraction from individual responsibility (Q9: -4, Q64: -5), placing responsibility onto one group (Q7: -5, 20: -5), putting individual views aside (Q33: -3), and minimising a sense of individual responsibility (Q65: -4) were ranked the lowest in this Factor. Participants in Factor 1 reported to have ranked statements as 'most effective' because "…*if we cannot admit that our ignorance or lack of openness could be causing part of the problem then nothing will change*", "*[they] centre on personal responsibility*" and *"…[they] highlight the importance of everyone within a service acknowledging their own role/biases…"*. On the other hand, statements that placed "*responsibilities on ethnic minorities to fix the problem*" or loaded "*responsibility to one group"* were considered least effective. Responsibility extended from individual staff members to the NHS as an organisation, with participants commenting that effectively addressing racial inequality will require "*system level changes"* and "*cultural change*".

Table 11

Rounded Z-scores for Items Most Strongly Associated with Factor 1

Q-Statement	Average Factor Score
Q8…ethnic minority staff should be better supported when racism is directed at them from service users	+5
Q43…staff must be willing to recognise their biases	+5
Q57…there must be measurable action not just conversations, research and policies	+5
Q10…each member of staff must personally think addressing racial inequality is important	+4

Q26…healthcare interventions must be appropriately adapted for the needs of ethnic minority cultures and communities	+4
Q60…healthcare treatments need to be tailored to effectively meet the needs of ethnic minority service users	+4
Q61…the lived experience of ethnic minority staff and service users should be learnt from	+4
Q9…broad brush solutions need to be applied as standard across the NHS	-4
Q31…treatment should be withheld from service users who are racist to staff	-4
Q62…ethnic minority staff must be kept in their roles for the long term	-4
Q65…it is important NHS staff don't feel blamed or labelled as racist	-4
Q7…NHS staff not born in Britain should have the opportunity to be educated on what different social mannerisms mean in British culture	-5
Q20…addressing racial inequality should be solely left to NHS leaders to deal with	-5
Q64…every service user should be treated exactly the same	-5

Factor 1 also emphasised the importance of continually learning from and improving upon previous strategies to appropriately meet the needs of service users (Q8: +5, Q17: +3, Q26: +4, Q39: +3, Q57: +5, Q60: +4, Q61: +4). Comments elaborated that this should be approached reflectively "G*eneralist, broad, tokenistic…*" solutions (Q9: -4) were considered least effective, with participants further elaborating that "*things have already been implemented, e.g., questions about equality and diversity interviews, but [that] doesn't mean that someone actually considers these issues important*". Comments also suggested that if statements detailed actions that had been "*applied before"* but hadn't been "*effective*", or were "*tokenistic*", these were also ranked as least effective.

Distinguishing Statements for Factor 1. For each Q-statement, if one factor had a significantly higher or lower *z*-score compared to the other two factors, it was considered a distinguishing statement for that specific factor. Distinguishing statements for Factor 1 are shown in Table 12 and were considered along with relevant distinguishing all statements.

Suggestions that staff should be antiracist (Q42: +3) and that racial inequality should be acknowledged and accepted (Q49: +2) were ranked significantly higher in Factor 1 than in the other two factors. On the other hand, Factor 1 ranked Q3 (+2), Q7 (-5), Q13 (-3), Q20 (-5), Q33 (-3), and Q65 (-4) significantly lower than the other two factors. It could be interpreted that these statements did not align with staff taking individual responsibility (e.g., Q65: it is important NHS staff don't feel blamed or labelled as racist). Q57 (+5; measurable action) and Q61 (+4; learning from lived experience) were also ranked higher compared to the other factors. Considering the distinguishing all Q-statements, Q60 (+4) which detailed tailoring services to meet the need of ethnic minority service users was ranked significantly higher in Factor 1 whereas Q64 (-5) which stated every service user should be treated the same was ranked significantly lower than the other two factors.

Overall, this analysis further added to the interpretation that individual responsibility and bespoke action tailored to meet the needs of ethnic minority people were considered important in Factor 1 and as such this distinguished Factor 1 from the other two factors.

Associations Between Factor and Demographic Characteristics. *T*-tests were performed on the factor loadings for Factor 1 by demographics across all the 28 respondents (see Table 13). Significant differences were observed for gender, ethnicity, and job role, with a stronger positive endorsement of Factor 1 being associated with females, White British participants, and psychologists.

Table 12

Distinguishing Statements for Factor One

Q-statement	F1-F2	F1-F3	F2-F3
Q3. …each NHS Trust should have a central equality and diversity team for staff to access resources when needed	-0.463*	-0.824**	-0.361

Q-statement	F1-F2	F1-F3	F2-F3
Q13. …all NHS staff should be trained in the routine collection of service user ethnicity	-1.373*	-0.797**	0.576
Q20. …addressing racial inequality should be solely left to NHS leaders to deal with	-0.691**	-1.044***	-0.353
Q33. …personal views on racial inequality need to be put aside	-0.831***	-0.934***	-0.103
Q49. …the existence of racial inequality in the NHS must first be accepted before progress can be made	0.852***	1.070***	0.217
Q57. …there must be measurable action not just conversations, research, and policies	1.183*	0.855**	-0.328
Q61. …the lived experience of ethnic minority staff and service users should be learnt from	0.542*	0.820**	0.278

*$p < .05$, **$p < .01$, ***$p < .001$

Table 13

Differences in Factor 1 Factor Loadings by Demographic Variables Across All Of The Respondents

Demographic	Level	*N*	Average factor loading for Factor 1 (*SD*)	*t*	*p*
Age	Younger (18-35)	11	0.419 (*SD* = 0.268)	0.737	.471
	Older (46-65)	9	0.337 (*SD* = 0.217)		
Gender	Male	13	0.275 (*SD* = 0.243)	2.729	.011
	Female	15	0.489 (*SD* = 0.170)		
Ethnicity	White British	20	0.466 (*SD* = 0.190)	3.209	.004
	Ethnic Minority	8	0.199 (*SD* = 0.221)		
Job Role	Psychological practitioner	15	0.490 (*SD* = 0.191)	2.750	.011
	Other role	13	0.274 (*SD* = 0.224)		

EBE	EBE Non-EBE	4 24	0.310 (*SD* = 0.250) 0.403 (*SD* = 0.230)	-0.702	.463
Job Length	5 years or less 6years or more	18 10	0.409 (*SD* = 0.240) 0.356 (*SD* = 0.220)	0.576	.0.285

Factor 2: 'Research Focused'

Factor 2 accounted for 19% of the total variance and was significantly associated with eight Q-sorts (R10, R13, R17, R18, R21, R23, R25, and R26). Overall, Factor 2 appeared to reveal a perspective in which research was ranked most effective for addressing racial inequality whereas individual-level strategies were rank low. Several post-sort comments centred on the potential problems of making adjustments for ethnic minority staff and service users. The most strongly endorsed items (rounded *z*-score = ± 4/5) are shown in Table 14 along with their factor scores.

Based upon the comments and the ranking of statements in Factor 2, "*more research into the area"* (Q12: +3, Q17: +5, Q18: +4, Q24: +5, Q38: +4, Q39: +3), including co-production (Q17: +5 and Q11: +4), was considered one of the most effective strategies for addressing racial inequality. This could also be linked to the high ranking of appropriately adapting healthcare interventions for ethnic minority service users (Q26: +4).

Statements which were interpreted to be related to making adjustments for ethnic minority staff or service users were queried because "*other staff will feel effected*" and were ranked as least effective. This theme included the low ranking of "*punitive measures*"

(Q32: - 4), prioritising ethnic minority staff career progression (Q54: -5 and Q62: -5) and providing ethnic minority staff/service users with extra resources (Q29: -3). One comment expressed a concern that certain strategies could "*alienate the Non-BME people*" and "*reverse the direction of racism*". Statements which suggested spotlighting staff's

Table 14

Rounded Z-scores for Items Most Strongly Associated with Factor 2

Q-Statement	Average Factor Score
Q8…ethnic minority staff should be better supported when racism is directed at them from service users	+5
Q17…greater involvement of ethnic minority people in healthcare research is needed	+5
Q24…more research is needed on the health effects of racial inequality	+5
Q11…ethnic minority staff and service users should be continually involved in decision making	+4
Q18…interventions applied to address racial inequality will need to have measurable outcomes	+4
Q26…healthcare interventions must be appropriately adapted for the needs of ethnic minority cultures and communities	+4
Q38…research is required to understand where ethnic minority service users experience the most racial inequality whilst receiving NHS treatments	+4
Q9…broad brush solutions need to be applied as standard across the NHS	-4
Q19…it must be accepted that it's not just 'a few bad apples' who are racist	-4
Q32…Trusts that are not meeting equality targets should be penalised	-4
Q34…all NHS staff should complete questionnaires about how comfortable they feel talking about race	-4
Q20…addressing racial inequality should be solely left to NHS leaders to deal with	-5
Q54…the career development of ethnic minority staff must take priority over that of White staff	-5
Q62…ethnic minority staff must be kept in their roles for the long term	-5

relationship to race and racial inequality (Q19: -4, Q22: -3, Q34: -4, Q48: -3) were ranked some of the lowest. Indeed, comments from the post-sort questionnaire echoed a sense that focusing on staff's views would be unrealistic "…*staff should change their views. Ok but how? Ideally yes but how would it happen?*".

Distinguishing Statements for Factor 2. Distinguishing statements for Factor 2 are shown in Table 15. Distinguishing all statements were also considered. Q4 (+2), Q11 (+4), Q12 (+3), Q17 (+5), Q18 (+4), Q24 (+5), Q38(+4), and Q41(+2) which all related to research, were ranked significantly higher in Factor 2 than in the other two factors.

Conversely, statements interpreted to be related to making adjustments based on someone's ethnicity were ranked significantly lower in Factor 2 than in the other two factors. For example, adjusting recruitment processes (Q2: -2, Q36: -3) and including more ethnic minority staff in senior levels of the NHS (Q40: -2) were all ranked significantly lower. Statements that linked to individual/interpersonal-level strategies including considering staff's relationship to race and racial inequality (Q19: -4, Q43: -1), discussing racial inequality as a team (Q22: -3), staff being antiracist (Q42: -2), and assessing teams for how they respond to issues relating to race (Q48: -3) were also ranked significantly lower in Factor 2 than in the other two factors.

Overall, this analysis further added to the interpretation that research was considered the most effective strategy where individual-level strategies and to some degree making adjustments for ethnic minority staff were least effective.

Associations Between Factor and Demographic Characteristics. *T*-tests were performed on the factor loadings for Factor 2 by demographic characteristics across all 28 respondents (see Table 16). No significant differences were observed.

Table 15

Distinguishing Statements for Factor 2

Q-statement	F1-F2	F1-F3	F2-F3
Q11. …ethnic minority staff and service users should be continually involved in decision making	-0.976 ***	-0.047	0.929 **
Q12. …NHS leaders should be demanding more evidence about what is effective for addressing racial inequality	-0.676 **	0.130	0.806 **
Q14. …the percentage of ethnic minority and White NHS staff should be the same as that of the general population	-0.725 **	-0.135	0.590*
Q15. …each ethnic minority service user should have an assessment to see how accessible NHS services are for ~~them~~	-0.603 **	0.110	0.713*
Q18. …interventions applied to address racial inequality will need to have measurable outcomes	-1.274*	0.058	1.332 ***
Q19. …it must be accepted that it's not just 'a few bad apples' who are racist	2.000*	-0.077	-2.077*
Q24. …more research is needed on the health effects of racial inequality	-1.891*	-0.350	1.541*
Q31. …treatment should be withheld from service users who are racist to staff	-0.561*	0.065	0.626*
Q36. …recruitment processes should place more importance on an applicant's underlying competencies over their qualifications	1.346*	-0.318	-1.664 *
Q40. …senior levels of the NHS should include more ethnic minority people	1.096 ***	-0.226	-1.322 ***
Q43. …staff must be willing to recognise their biases	1.769*	0.468	-1.301 ***
Q47. …strategies must be completely focused on addressing racism that is deeply embedded in the NHS as an organisation	0.612 **	-0.419	-1.031 ***
Q48. …all NHS teams should be assessed on how they respond to issues relating to race	1.407*	-0.311	-1.718 *
Q63. …the long history of racism in the NHS should not be forgotten	1.019 ***	0.399	-0.621 *

Table 16

Differences in Factor 2 Factor Loadings by Demographic Variables Across All of The Respondents

Demographic	Level	*N*	Average factor loading	*t*	*p*
Age	Younger (18-35)	11	0.351 (*SD* = 0.195)	-0.634	.534
	Older (46-65)	9	0.419 (*SD* = 0.283)		
Gender	Male	13	0.3968 (*SD* = 0.288)	-0.884	.385
	Female	15	0.311 (*SD* = 0.228)		
Ethnicity	White British	20	0.342 (*SD* = 0.223)	-0.242	.407
	Ethnic Minority	8	0.373 (*SD* = 0.345)		
Job Role	Psychological practitioner	15	0.3786 (*SD* = 0.265)	0.613	.545
	Other role	13	0.3182 (*SD* = 0.254)		
EBE	EBE	4	0.184 (*SD* = 0.240)	-1.429	.165
	Non-EBE	24	0.378 (*SD* = 0.253)		
Job Length	5 years or less	18	0.318 (*SD* = 0.234)	-0.899	.188
	6 years or more	10	0.409 (*SD* = 0.297)		

(rounded *z*-score = ± 4/5) are shown in Table 17 along with their factor scores.

Statements that centred around policy and organisation-level solutions, such as ensuring the diversity of staff on interview panels (Q2: +5), compulsory E&D questions in interviews (Q59: +4), staff access to E&D resources (Q1: +3; Q3: +4), zero-tolerance policies (Q58: +5), and E&D training for staff (Q16: +4, Q27: +3) were ranked highly in Factor 3. Although placing all

Table 17

Rounded Z-scores for Items Most Strongly Associated with Factor 3

Q-Statement	Average Factor Score
Q2…a member of staff with an ethnic minority background should always be included on NHS interview panels	+5
Q58…there should be a zero-tolerance policy on racial inequality	+5
Q64…every service user should be treated exactly the same	+5
Q3…each NHS Trust should have a central equality and diversity team for staff to access resources when needed	+4

Q16…every member of staff should have generic equality and diversity training	+4
Q30…staff need to accept they may not realise they're racist	+4
Q59…there should be compulsory questions about equality and diversity in job interviews	+4
Q20…addressing racial inequality should be solely left to NHS leaders to deal with	-4
Q45…ethnic minority service users' attitudes towards using NHS services should be surveyed	-4
Q54…the career development of ethnic minority staff must take priority over that of White staff	-4
Q62…ethnic minority staff must be kept in their roles for the long term	-4
Q9…broad brush solutions need to be applied as standard across the NHS	-5
Q29…ethnic minority staff / service users should be provided with extra resources and opportunities to compensate for their disadvantage	-5
Q31…treatment should be withheld from service users who are racist to staff	-5

responsibility onto NHS leaders were ranked very low (Q20: -4).

Assessing the experiences of ethnic minority service users was ranked low and included the low ranking of assessing ethnic minority patient's accessibility to services (Q15: -3), assessing the efficacy of treatments for ethnic minority service users (Q39: -3), tracking health outcomes of ethnic minority service users (Q41: -2), and surveying ethnic minority service users' attitudes towards services (Q45: -4) | .

Furthermore, statements that centred around making adjustments for ethnic minority people was also prominent. This interpretation stemmed from the low ranking of statements which suggested ethnic minority service users and staff should have extra resources and that ethnic minority staff careers should be prioritised (Q29: -5; Q54: -4; Q55: -3: Q62: -4). This theme was complemented by the high ranking of Q64 (+5) which stated all service users 'should be treated exactly the same'.

Distinguishing Statements for Factor 3. Distinguishing statements are shown in Table 18. Distinguishing all statements were also considered.

Q1 (+3) which suggested a director of equality should be appointed, Q16 (+4) and Q27 (+3) which supported staff training, Q58 (+2) which considered a zero-tolerance policy, and Q59 (+4) which outlined job interviews should include questions on equality and diversity were ranked significantly higher in Factor 3 than in the other two factors. On the other hand, statements that suggested allocating resources to meet service users' needs in a culturally appropriate way, providing additional resources to ethnic minority service users and staff, providing high quality interpreting services and providing allocated workers to engage with ethnic minority communities, were ranked significantly lower in Factor 3 than in other factors (Q26: 0; Q28: -1; Q29: -5; Q37: -2; Q39: -3; Q45: -4; Q55: -3). Considering the distinguishing all statements, Factor 3 ranked Q64 (+5), which suggested every service user should be treated the same, significantly higher than in the other two factors.

Overall, this analysis further added to the interpretation that policy and organisational-level solutions were considered most effective by participants in Factor 3. Making adjustments for ethnic minority service users based on their ethnicity was ranked significantly lower in Factor 3 and this distinguished Factor 3 from the other two factors.

Table 18

Distinguishing Statements for Factor 3

Q-statement	F1-F2	F1-F3	F2-F3
Q8. …ethnic minority staff should be better supported when racism is directed at them from service users	-0.098	0.829**	0.927**
Q10. …each member of staff must personally think addressing racial inequality is important	0.071	0.747**	0.677*
Q21. …it will be important to create a safe space for ethnic minority staff before expecting them to talk about racism	-0.237	0.654*	0.891**
Q26. …healthcare interventions must be appropriately adapted for the needs of ethnic minority cultures and communities	0.092	1.145***	1.053***

Q27. …NHS professionals should have specific training on racism that incorporates real life scenarios	-0.177	-1.198***	-1.020***
Q28. …services should have allocated workers to engage people from hard-to-reach ethnic minority communities	-0.386	0.665*	1.051***
Q37. …serious efforts are needed to ensure access to high quality interpreting services	-0.297	1.290***	1.587*
Q39. …the efficacy of NHS treatments for ethnic minority groups must be reviewed	0.022	2.034*	2.013*
Q45. …ethnic minority service users' attitudes towards using NHS services should be surveyed	-0.239	1.195***	1.434***
Q55. …NHS services should be able to allocate funding individually to each service user to meet their needs in the way that feels most culturally appropriate	-0.068	0.963***	1.031***
Q58. …there should be a zero-tolerance policy on racial inequality	0.124	-1.356***	-1.479*

Associations Between Factor 3 Loadings and Demographic Characteristics. *T*-tests were performed on the factor loadings for Factor 3 by demographic characteristics across all 28 respondents (see Table 19). Significant differences were observed for ethnicity and job role with a stronger positive endorsement of Factor 3 being associated with those who were not psychological practitioners.

Table 19

Differences in Factor 3 Factor Loadings by Demographic Variables Across All of The Respondents

Demographic	Level	*N*	Average factor	*t*	*p*
Age	Younger (18-35)	11	0.196 (*SD* = 0.288)	-6.580	.519
	Older (46-65)	9	0.271 (*SD* = 0.202)		
Gender	Male	13	0.397 (*SD* = 0.288)	-0.985	.334
	Female	15	0.311 (*SD* = 0.228)		

Ethnicity	White British	20	0.215 (*SD* = 0.230)	-1.235	.114
	Ethnic Minority	8	0.341 (*SD* = 0.279)		
Job Role	Psychological	15	0.147 (*SD* = 0.191)	-2.629	.014
	Other role	13	0.370 (*SD* = 0.256)		
EBE	EBE	4	0.375 (*SD* = 0.251)	1.094	.284
	Non-EBE	24	0.230 (*SD* = 0.245)		
Job Length	5 years or less	18	0.269 (*SD* = 0.270)	0.515	.305
	6 years or more	10	0.218 (*SD* = 0.207)		

Discussion

Q-methodology was used to explore NHS stakeholders' perspectives on how to address racial inequality most effectively in the NHS. Studying perspectives, including areas of consensus and difference, can help to understand and shift stagnation. Five people took part in stage 1, creating the Q-set, and twenty-eight people took part in stage 2, the Q-sort. A solution outlining three different opinion groups was identified. All three opinion groups suggested racial inequality in the NHS could be addressed most effectively in a different way.

The themes that emerged appeared to reflect strategies already recommended or employed to address racial inequality in the NHS (Kapadia et al., 2022; NHS England, 2022b; The NHS Staff Council, 2021). However, the aim of this study was not necessarily to generate new solutions, rather it was to explore the range of stakeholder perspectives on such strategies and to consider barriers to change. Factor 1 prioritised the high ranking of statements centring on staff being motivated to address their biases, Factor 2 prioritised research, and Factor 3 prioritised organisational and policy level solutions. Some of these themes conflicted across groups. Whilst Factor 1 participants ranked strategies centring on individual biases highly, Factor 2 participants ranked these as some of the lowest. 'Equity' on the other hand was a

theme that featured across all three opinion groups but was related to differently by each factor. Equity, taking individual circumstances and needs into account to inform actions, can be considered different to equality which refers to treating people the same (Espinoza, 2007). The ranking of Q64 (‘every service user should be treated exactly the same’), demonstrated this as it was ranked significantly differently across Factor 1 (-5), Factor 2 (+2), and Factor 3 (+5). Of the ten consensus statements identified, most were strategies felt to be least effective, or not prioritised either way, and no particular themes were identified. Only two statements were ranked moderately high across all factors (Q25: 3, 3, 3 and Q52: 3, 2, 1). This suggests participants agreed more about what they would find unhelpful than what would be most effective. This interpretation echoes the narrative summarised by Adeyemi (2019), that whilst strategies have been tried before and considered insufficient, agreement about what will work is less clear. Differences were also found across demographics. Psychological practitioners were much more highly associated with Factor 1 than other roles. This is understandable given the reflective nature of a psychological practitioner, including the importance of considering personal views and biases, as well as ‘difference’ and ‘sameness’ (Nolte, 2017; Prasko et al., 2012). Factor 3 on the other hand was more associated with non-psychological roles and flagged participants included those in corporate roles, which is in line with the defining theme of this opinion group.

As applied in other Q-studies (DeCourville & Hafer, 2001), these findings and interpretations could be used to inform strategies for addressing racial inequality in the NHS. For example, knowing that all opinion groups ranked ‘broad brush’ solutions as very low means it could be beneficial to avoid such strategies, and knowing that different professions align with particular opinion groups to differing degrees may mean it could be beneficial for professional bodies to support the NHS in addressing racial inequality. However, the

distinguished differences and conflicting themes between opinion groups could be a barrier to directly apply these results to address racial inequality. For example, applying solutions in line with Factor 1 participants' high ranking of strategies centring on individual staff bias is unlikely to be supported by participants in Factor 2 which ranked these strategies low.

The contrast of perspectives across opinion groups is considered one of the most important findings in this study. Each opinion group prioritised a different strategy, ranked the same themes differently, and agreed more about what would be least effective than most effective. Indeed, identifying the different subcultures within an organisation has been highlighted as crucial for organisational culture change (Scott et al., 2003). As described previously, without shared goals, beliefs, and motivations, organisational change is limited (NHS England, 2018) and without incorporating the needs and feedback of stakeholders' strategies are unlikely to be effective (Altman et al., 2018). If the finding that stakeholders have distinctly different perspectives on how to address racial inequality can help to explain the perpetual cycle of attempting to address racial inequality but falling short, this is extremely valuable to make explicit and to actively apply going forward.

To translate these findings into practice, it is recommended that recognition of the multiple, and potentially conflicting, perspectives of stakeholders is required by all levels of the NHS organisation. Whilst this may seem intuitive, it is not something routinely acknowledged in strategic documents, reports, or policies. Instead one size fits all recommendations are often provided (Kapadia et al., 2022; NHS England, 2022a). The current findings suggest broadly applied policies, reports, training, and targets are unlikely to align with the overall range of opinions across stakeholder populations. This was a view shared by all opinion groups in this study. Rather, interventions to address racial inequality in the NHS may be most effective if they are tailored to meet a range of stakeholder's needs.

This may require a complete culture change for the NHS as an organisation. How this would be achieved and the plausibility of delivering this within an organisation as large as the NHS which has statutory standards to meet is unknown (Scott et al., 2003) and may need to be the focus of future research. It is hoped that the findings from this study will be a helpful starting place to better understand the underlying processes that contribute to the felt sense of 'stuckness' around addressing racial in the NHS (Skorpen et al., 2012).

As this study was a novel approach to researching how to address racial inequality in the NHS, both strengths and limitations were identified. It is hoped that both are beneficial for informing future research. The purpose of a concourse is to facilitate participants' expressions of their perspectives. If the concourse is not derived from a breadth and depth of sources, this limits the effectiveness of the Q-sorts. A strength of this study was the broad range of themes captured in the development of the Q-set. Two methods were utilised, a literature search and stakeholder interviews. This is considered a particularly in-depth level of concourse development (Gough et al., 2014; Watts & Stenner, 2012). Furthermore, the diversity of the interviewees was also considered ok across demographic areas. However, building a different concourse (e.g., from wider or narrower stakeholder pools), or simply updating the concourse, is likely to result in the identification of different opinions groups and different interactions. This is a recommendation for future research.

A limitation of the current study was that the demographics of the participants who completed the Q-sorts were disproportionate, with there being significantly less EbEs, ethnic minority participants, and non-psychological practitioner participants. Whilst ethnic minority participants were represented to the same proportion as the working population (NHS Digital, 2019), Q-methodology does not necessarily aim to be representative, but rather to capture important perspectives (Watts & Stenner, 2012). A strength of the Q-set development stage

was not only completing interviews with stakeholders but being able to draw upon media sources written by many different people including those from a range of ethnicities and backgrounds. This increased the breadth, depth, and richness of the concourse supporting its utilisation and ability to be used as a vessel for expressing participants' perspectives. Including participants who were White British was considered a strength of this study. In the UK, as racial discrimination is directed from White British people and systems dominated by White British people, White British people need to be part of the solution (Flintoff et al., 2015). Therefore it was considered helpful and important to identify the perspectives of White British people in relation to the solutions. It was hoped that understanding White British people's perspectives on what would be most or least effective would give insight into their motivation and opinions on engaging with such strategies. However, considering that racial inequality was the topic of this study, understanding White British perspectives needed to be balanced with capturing the voices of ethnic minority stakeholders as well. The proportion of participants identifying with an ethnic minority group is disappointing. The decision to retain all White British participants introduces the limitation that White British perspectives are captured more than ethnic minority perspectives. This limits the breadth, depth, and richness of this data as well as the opportunity to observe similarities and differences across ethnic groups. Fundamentally, as much as it is important to capture the perspectives of White British people to better understand the barriers to addressing racial inequality, it is crucially important to understand the strategies that those with lived experience feel are effective and appropriate solutions. Limitations in recruitment that are likely to have contributed to this discrepancy between the number of White British participants and ethnic minority participants were identified. Firstly, participants were recruited through local collaborators that could reach a wide pool of people as well as through word of mouth through the researcher's contacts.

However, as participants were able to participate in the Q-sort autonomously and anonymously, the ability to monitor or gate keep who was participating was limited. This may also have contributed to the high number of psychological practitioners who were likely recruited via snowball sampling; a common method used in Q-methodology (Webler et al., 2007). However, due to the issue of ethnic minority people being underrepresented in certain professions such as Clinical Psychology (York, 2019), snowball sampling is likely to have added to the discrepancy. Target sampling can be used in Q-methodology to seek out participants with certain demographics or characteristics to ensure a balance is found. This would be a recommendation for future research to consider sensitively. Future research should consider whether anonymous participation is appropriate, or whether an increased level of gatekeeping would be helpful. This further related to the use of online Q-sort software.

A further recruitment limitation in the current study, was suggested by those with lived experience that NHS staff identifying with an ethnic minority group may find engaging in research within the NHS conflicting due to the context of systemic racism and the White British ethnicity of the researcher may also have been a limiting factor in this. The researcher being White British may have reduced the sense of psychological safety for potential participants identifying with an ethnic minority group. On the other hand, it may have increased the psychological safety of engaging in research about racism for White British people. This could have contributed to majority of participants identifying as White British. Furthermore, a White British person may bring with them their own biases and blind spots (Gordon, 2005), with it being acknowledged that if Whiteness and power imbalances are not consciously acknowledged, it can reduce how much is learnt from research studies (Phillippo & Nolan, 2022). Mio and Iwamasa (1993) outlined key lessons to be applied for White researchers initiating cross-cultural studies. Collaborative working and a respect of both

White and Ethnic Minority motivations for research was suggested. White researchers were encouraged not to ‘run away’ from being a part of the research field, but rather to learn from ethnic minority people, gain insight into particular issues for ethnic minority people, and to be aware of the potential resentment a White researcher may evoke. It was considered that this will be important for the research to have a positive effect. Building upon these reflections in future research will be crucial. Whilst Q-methodology used in this study inherently drew on those with lived experience, a second researcher or a consultant identifying as an ethnic minority person should be considered going forward to provide additional insight and direction from a research design position.

This had its strengths, in that it enabled participants to take part remotely within a COVID context when health concerns were still present, and its limitations, such as being unable to gauge the engagement and understanding of participants. To build upon the current study, in-person Q-sorts should be aimed for when possible. If online Q-sorts are used, the researcher may want to consider being present remotely via video conferencing facilities to provide support and ascertain feedback from participants about using the online Q-sort software. Whether in person or remotely, facilitating live post-sort interviews should be considered a priority for future research. Indeed, more information on why participants completed the Q-sorts in the way they did would have been highly beneficial for supporting interpretations in the current study.

Conclusion

Overall this study showed that stakeholders held different perspectives on how to address racial inequality most effectively in the NHS. Whilst the number of perspectives identified in this study may not be exhaustive, the perspectives identified do represent those within the study population. Defining, consensus, and distinguishing themes could be used to

develop targeted interventions. However, it is hoped that considering the significance distinct opinion groups is recognised as they could be a substantial barrier to applying solutions broadly. Effectively incorporating the needs of stakeholders and supporting shared motives, beliefs, and goals to facilitate organisational change will need greater thought.

Previous reports and research have applied quantitative methods to explore the impact of racial inequality in the NHS (Kapadia et al., 2022; NHS, 2022). Using Q-methodology to explore stakeholder's perspectives is considered a novel approach, and this study can be considered an initial exploration of applying this method. As such, further research is urged to build upon of these findings and improve upon its limitations. Priority should be given to capturing the perspectives of more EbEs and ethnic minority participants, particularly at the Q-sort stage. It may be that further Q-sorts within specific professions, Trusts, or teams would be useful to support extrapolation of the results to those groups. Alternatively, if more innovative, practical solutions are desired, a study utilising a Design Think approach could be a follow up study. The Design think framework could use the understanding of stakeholder's perspectives developed in this study to empathise with stakeholders, and then support the innovation and implementation of new ideas.

References

Adebowale, V., & Rao, M. (2020). Racism in medicine: Why equality matters to everyone. *British Medical Journal, 368.* https://doi.org/10.1136/bmj.m530

Adeyemi, A. (2019, October 14). The future of race in the NHS. *NHS England.* https://www.england.nhs.uk/blog/the-future-of-race-in-the-nhs/ [last accessed 19th December 2022]

Altman, M., Huang, T. T. K., & Breland, J. Y. (2018). Design thinking in health care. *Preventing Chronic Disease*, *15*(117). https://doi.org/10.5888/pcd15.180128

Anekwe, L. (2020). Ethnic disparities in maternal care. *British Medical Journal, 368*(442). https://doi.org/10.1136/bmj.m442

Bamrah, J. S., & Chakravorty, I. (2022). Racism in the NHS: Let BMA report be last needed so we can return to tackling NHS crisis. *British Medical Journal*. *378*. https://doi.org/10.1136/bmj.o2064

Beagan, B. L., Sibbald, K. R., Bizzeth, S. R., & Pride, T. M. (2022). Systemic racism in Canadian occupational therapy: a qualitative study with therapists. *Canadian Journal of Occupational Therapy*, *89*(1), 51–61. https://doi.org/10.1177/00084174211066676

Ben, J., Cormack, D., Harris, R., & Paradies, Y. (2017). Racism and health service utilisation: A systematic review and meta-analysis. *PloS One, 12*(12). https://doi.org/10.1371/journal.pone.0189900

Broderick, F., Deignan, T., & Combes, H. (2017). An Exploratory study into the traumatic impact of advanced cancer among patients and partners. *Operant Subjectivity*, *39*(3/4), 1–24. https://doi.org/10.15133/j.os.2017.009

Brown, S. R. (1980). *Political Subjectivity: Applications of Q Methodology in Political Science*. Yale University Press.

Cave, L., Cooper, M. N., Zubrick, S. R., & Shepherd, C. C. J. (2020). Racial discrimination and child and adolescent health in longitudinal studies: A systematic review. *Social Science & Medicine, 250*. https://doi.org/10.1016/j.socscimed.2020.112864

Chan, K. (2018). *A design thinking mindset beyond the public health model*. World Medical & Health Policy, *10*(1), 111- 119. https://www.doi.org/10.14707/ajbr.210102

Chand, K. (2018, April 20). Racism is still rife in the NHS. *The British Medical Journal Opinion*. https://blogs.bmj.com/bmj/2018/04/20/kailash-chand-racism-is-still-rife-in-the-nhs/ [last accessed 19th December 2022]

Combes, H., Hardy, G., & Buchan, L. (2004). Using Q-methodology to involve people with intellectual disability in evaluating person-centred planning. *Journal of Applied Research in Intellectual Disabilities, 17*(3), 149–159. https://doi.org/10.1111/j.1468-3148.2004.00191.x

Crown. (2018). *Modernising the Mental Health Act – final report from the independent review*. GOV.UK. https://www.gov.uk/government/publications/modernising-the-mental-health-act-final-report-from-the-independent-review

DeCourville, & Hafer. (2001). Attitudes toward affirmative action programs: A Q Methodological Study. *Operant Subjectivity, 24*(4), 183–200. https://doi.org/10.15133/j.os.2001.010

Deitch, E., Barsky, A., Butz, R., Chan, S., Brief, A., & Bradley, J. (2003). Subtle yet significant: The existence and impact of everyday racial discrimination in the workplace. *Human Relations, 56,* 1299–1324. https://doi.org/10.1177/00187267035611002

Dilbeck. (2017). *The SAGE Encyclopaedia of communication research methods*. SAGE Publications, Inc.

Donner, J. (2001). Using Q-sorts in participatory processes: An introduction to the methodology. In Krueger, R. A., Casey, M. A., Donner, J., Kirsch, S., & Maack, J. N. (Eds). *Social Analysis Selected Tools and Techniques*. Social Development Department, The World Bank. https://www.academia.edu/2398312/Using_q_sorts_in_participatory_processes_An_introduction_to_the_methodology [last accessed 19th December 2022]

Dyer. (2019, July 12). How can we have the same outcomes when we're not having the same experiences? *NHS England.* https://www.england.nhs.uk/blog/how-can-we-have-the-same-outcomes-when-were-not-having-the-same-experiences/ [last accessed 19th December 2022]

Esmail, A., & Carnall, D. (1997). Tackling racism in the NHS: We need action not words. *British Medical Journal*, *314*(7081). https://doi.org/10.1136/bmj.314.7081.618

Espinoza, O. (2007). Solving the Equity-Equality Conceptual Dilemma: A New Model for Analysis of the Educational Process. *Educational Research, 49,* 343–363. https://doi.org/10.1080/00131880701717198

Festinger, L. (1957). *A theory of cognitive dissonance* (1st ed.). Stanford University Press.

Flintoff, A., Dowling, F., & Fitzgerald, H. (2015). Working through whiteness, race and (anti) racism in physical education teacher education. *Physical Education and Sport Pedagogy*, *20*(5), 559–570. https://doi.org/10.1080/17408989.2014.962017

Fluckinger, C. D. (2014). Big Five measurement via Q-Sort: An alternative method for constraining socially desirable responding. *SAGE Open*, *4*(3). https://doi.org/10.1177/2158244014547196

Forrest-Bank, S. S., & Jenson, J. M. (2015). The relationship among childhood risk and protective factors, racial microaggression and ethnic identity, and academic self-

efficacy and antisocial behavior in young adulthood. *Children and Youth Services Review*, *50*, 64–74. https://doi.org/10.1016/j.childyouth.2015.01.005

Gallagher, K., & Porock, D. (2010). The use of interviews in Q methodology: Card content analysis. *Nursing Research*, *59*(4), 295–300. https://doi.org/10.1097/NNR.0b013e3181e4ffff

Gay, W., & Bamford, D. (2007). A case study into the management of racial diversity within an NHS teaching hospital. *International Journal of Public Sector Management*, *20*(4), 257–271. https://doi.org/10.1108/09513550710749996

Glanz, K., Rimer, B. K., & Viswanath, K. (Eds.). (2008). *Health behavior and health education: theory, research and practice* (4th ed.). John Wiley & Sons.

Goretzko, D., Pham, T. T. H., & Bühner, M. (2021). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology*, *40*(7), 3510–3521. https://doi.org/10.1007/s12144-019-00300-2

Gordon, J. (2005). White on White: Researcher Reflexivity and the Logics of Privilege in White Schools Undertaking Reform. *The Urban Review*, *37*(4), 279–302. https://doi.org/10.1007/s11256-005-0015-1

Gough, H., Tweed, A., & Sorensen, J. (2014). Priorities of staff and service users in mental health assessments: A Q methodological study. *Operant Subjectivity*, *37*(4), 23–40. https://doi.org/10.15133/j.os.2014.014

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464

Guschke, B. L., & Christensen, J. F. (2021). From individual to organisational bias: A norm-critical proposition for unconscious bias intervention in organizations. *Women, Gender & Research, 32*(3*)*, 86-105. https://doi.org/10.7146/kkf.v31i3.128974

Harmon-Jones, E. (Ed.). (2019). C*ognitive dissonance: Re-examining a pivotal theory in psychology* (2nd ed.). American Psychological Association.

Hassen, N., Lofters, A., Michael, S., Mall, A., Pinto, A. D., & Rackal, J. (2021). Implementing anti-racism interventions in healthcare settings: A scoping review. *International Journal of Environmental Research and Public Health, 18*(6). https://doi.org/10.3390/ijerph18062993

Herrington, N., & Coogan, J. (2011). Q methodology: An overview. *Research in Teacher Education, 1*(2). https://doi.org/10.15123/uel.8604v

Iacobucci, G. (2020). NHS blood unit is "systemically racist" and "psychologically unsafe," investigation finds. *British Medical Journal, 369*. https://doi.org/10.1136/bmj.m2308

Johnson, M. B. (2012). Experimental test of social norms theory in a real-world drinking environment. *Journal of Studies on Alcohol and Drugs, 73*(5), 851–859. https://doi.org/10.15288/jsad.2012.73.851

Jolliff. (2019, August 8). Inclusive leadership: How ready are we to learn? *The King's Fund*. https://www.kingsfund.org.uk/blog/2019/08/inclusive-leadership

Kaltiso, S.-A. O., Seitz, R. M., Zdradzinski, M. J., Moran, T. P., Heron, S., Robertson, J., & Lall, M. D. (2021). The impact of racism on emergency health care workers. *Academic Emergency Medicine*, *28*(9), 974–981. https://doi.org/10.1111/acem.14347

Kapadia, D., Zhang, J., Salway, S., Nazroo, J., Booth, A., Villarroel-Williams, N., Becares, L., & Esmail, A. (2022). *Ethnic Inequalities in Healthcare: A Rapid Evidence Review*. NHS Race & Health Observatory. https://www.nhsrho.org/publications/ethnic-

inequalities-in-healthcare-a-rapid-evidence-review/ [last accessed 29th December 2022]

Kar, P. (2020). To tackle racism, the NHS needs policies with teeth. *British Medical Journal, 369*. https://doi.org/10.1136/bmj.m2583

Kline. (1994). *An easy guide to factor analysis*. Routledge. https://doi.org/10.4324/9781315788135

Kline, R. (2014). *The snowy white peaks of the NHS: A survey of discrimination in governance and leadership and the potential impact on patient care in London and England*. Middlesex University. https://doi.org/10.22023/mdx.12640421.v1

Kowalski, C. M., Rogoza, R., Vernon, P. A., & Schermer, J. A. (2018). The Dark Triad and the self-presentation variables of socially desirable responding and self-monitoring. *Personality and Individual Differences, 120*, 234–237. https://doi.org/10.1016/j.paid.2017.09.007

Ladan, M. A., Wharrad, H., & Windle, R. (2018). Towards understanding healthcare professionals' adoption and use of technologies in clinical practice: Using Q-methodology and models of technology acceptance. *Journal of Innovation in Health Informatics, 25*(1), 965. https://doi.org/10.14236/jhi.v25i1.965

Leonard, R., Sundblad, D., & B. Thomas, D. (2021). Attribution and ideology in American politics: causal reasoning, political cognition, and partisan polarization in the age of trump. *Operant Subjectivity, 42*, 86–109. https://doi.org/10.15133/j.os.2020.005

Lippert-Rasmussen, K. (2020). *Making Sense of Affirmative Action*. Oxford University Press.

Mann, C. (2014). *The effect of mindfulness on racial stereotype activation and application.* [Doctoral dissertation, Bryn Mawr College]. https://doi.org/10.22023/mdx.12640421.v1

McHugh, N., Baker, R., Biosca, O., Ibrahim, F., & Donaldson, C. (2019). Who knows best? A Q methodology study to explore perspectives of professional stakeholders and community participants on health in low-income communities. *BMC Health Services Research*, *19*(1). https://doi.org/10.1186/s12913-019-3884-9

McLaughlin, J. E., Wolcott, M. D., Hubbard, D., Umstead, K., & Rider, T. R. (2019). A qualitative review of the design thinking framework in health professions education. *BMC Medical Education*, *19*(1), 98. https://doi.org/10.1186/s12909-019-1528-8

Mio, J. S., & Iwamasa, G. (1993). To Do, or Not to Do: That is the Question for White Cross-Cultural Researchers. *The Counseling Psychologist*, *21*(2), 197–212. https://doi.org/10.1177/0011000093212001

Morton, F., & Altschul, D. (2019). Data reduction analyses of animal behaviour: Avoiding Kaiser's criterion and adopting more robust automated methods. *Animal Behaviour*, *149*, 89–95. https://doi.org/10.1016/j.anbehav.2019.01.003

Nair, M., Kurinczuk, J. J., & Knight, M. (2014). Ethnic variations in severe maternal morbidity in the UK- a case control study. *PloS One*, *9*(4). https://doi.org/10.1371/journal.pone.0095086

NHS. (2022). *NHS England Workforce Race Equality Standard 2021*. NHS England. https://www.england.nhs.uk/publication/workforce-race-equality-standard-2021/

NHS Digital. (2019). *NHS Workforce Statistics—March 2019*. NHS Digital. https://digital.nhs.uk/data-and-information/publications/statistical/nhs-workforce-statistics/nhs-workforce-statistics---march-2019-provisional-statistics

NHS England. (2016, June 14). *NHS Workforce Race Equality Standard – full video (14/6/16)* [Video]. YouTube. https://www.youtube.com/watch?v=JBY8tlJs8xo

NHS England. (2018). *The change model guide*. NHS England. https://www.england.nhs.uk/publication/the-change-model-guide/ [last accessed 19th December 2022]

NHS England. (2022a). *NHS workforce race equality standard*. NHS England. https://www.england.nhs.uk/about/equality/equality-hub/workforce-equality-data-standards/equality-standard/ [last accessed 19th December 2022]

NHS England. (2022b). *NHS England and NHS improvement equality objectives for 2022/23 and 2023/24*. NHS England. https://www.england.nhs.uk/about/equality/objectives-for-22-23-and-23-24/ [last accessed 19th December 2022]

Noon, M. (2018). Pointless diversity training: unconscious bias, new racism and agency. *Work, Employment and Society*, *32*(1), 198–209. https://doi.org/10.1177/0950017017719841

Norfolk, Suffolk and Cambridgeshire & Strategic Health Authority. (2003). *Independent inquiry into the death of David Bennett*. http://image.guardian.co.uk/sys-files/Society/documents/2004/02/12/Bennett.pdf [last accessed 19th December 2022]

Osborne, J. (2019). What is rotating in exploratory factor analysis? *Practical Assessment, Research, and Evaluation*, *20*(1). https://doi.org/10.7275/hb2g-m060

Paradies, Y., Ben, J., Denson, N., Elias, A., Priest, N., Pieterse, A., Gupta, A., Kelaher, M., & Gee, G. (2015). Racism as a determinant of health: A systematic review and meta-analysis. *PLOS ONE*, *10*(9). https://doi.org/10.1371/journal.pone.0138511

Phillippo, K. L., & Nolan, J. L. (2022). White-on-white research: A study of white qualitative researcher positionality among white participants. *International Journal of Qualitative Studies in Education*, *0*(0), 1–19. https://doi.org/10.1080/09518398.2022.2061728

Prochaska, J. O., & DiClemente, C. C. (1982). Transtheoretical therapy: Toward a more integrative model of change. *Psychotherapy: Theory, Research & Practice*, *19*, 276–288. https://doi.org/10.1037/h0088437

Pruneddu, A. (2011). *Q-Sortware* [Computer Software]. Q-Sortware. https://www.qsortware.net/

Pruneddu, A. (2013). *Implicit person theories and Q-sort: Personality change in emerging adults* [Phd, University of York]. https://etheses.whiterose.ac.uk/6503/

QMethod Software. (2022). *Q Methodology: An Overview* [Computer software]. QMethod Software. https://qmethodsoftware.com/q-methodology/

Ramlo, S. E. (2008). Determining the various perspectives and consensus within a classroom using Q methodology. *AIP Conference Proceedings*, *1064*(1), 179–182. https://doi.org/10.1063/1.3021248

Roberts, J. P., Fisher, T. R., Trowbridge, M. J., & Bent, C. (2016). A design thinking framework for healthcare management and innovation. *Healthcare*, *4*(1), 11–14. https://doi.org/10.1016/j.hjdsi.2015.12.002

Roland G., F. (2011). Racial inequality in the 21st century: The declining significance of discrimination. In D. Card & O. Ashenfelter (Eds.), *Handbook of Labor Economics* (855–971). Elsevier. https://doi.org/10.1016/S0169-7218(11)02408-7

Ross, Jabbal, Chauhan, Maguire, Randhawa, & Dahir. (2020, July 7). *Workforce race inequalities and inclusion in NHS providers*. The King's Fund. https://www.kingsfund.org.uk/publications/workforce-race-inequalities-inclusion-nhs [last accessed 19th December 2022]

RStudio Team. (2020). *RStudio: Integrated Development for R. RStudio*. [Computer software] RStudio Team.

Salway, S., Holman, D., Lee, C., McGowan, V., Ben-Shlomo, Y., Saxena, S., & Nazroo, J. (2020). Transforming the health system for the UK's multiethnic population. *British Medical Journal, 368*. https://doi.org/10.1136/bmj.m268

Schuetzler, R., Grimes, G., Giboney, J., & Nunamaker, J. (2018). The influence of conversational agents on socially desirable responding. *Proceedings of the 51st Hawaii International Conference on System Sciences*, 283–292.

Scott, T., Mannion, R., Davies, H. T. O., & Marshall, M. N. (2003). Implementing culture change in health care: Theory and practice. *International Journal for Quality in Health Care, 15*(2), 111–118. https://doi.org/10.1093/intqhc/mzg021

Shabila, N. P., Al-Tawil, N. G., Al-Hadithi, T. S., & Sondorp, E. (2014). Using Q-methodology to explore people's health seeking behavior and perception of the quality of primary care services. *BMC Public Health, 14*(1). https://doi.org/10.1186/1471-2458-14-2

Silva, R., Oliveira, D., Santos, D. P., Santos, L. F. D., Wilson, R. E., & Bedo, M. (2020). Criteria for choosing the number of dimensions in a principal component analysis: An empirical assessment. *Anais*, 145–150. https://doi.org/10.5753/sbbd.2020.13632

Skorpen, F., Arstad Thorsen, A., Råholm, M.-B., Forsberg, C., & Rehnsfeldt, A. (2012). Caring With Dignity: A Q-methodological study of staff members' experiences in a psychiatric context. *Operant Subjectivity, 36*(1), 1–24. https://doi.org/10.15133/j.os.2012.001

Stephenson, W. (1935). Technique of factor analysis. *Nature*, *136*(3434). https://doi.org/10.1038/136297b0

Surash, S. (2020). Racial equality in the NHS: We need objective data to support subjective experience. *British Medical Journal, 371*. https://doi.org/10.1136/bmj.m4496

Sylvester, J. (2000). Q methodology and social marketing: interpreting racial attitudes toward health care. *Operant Subjectivity*, *23*(2), 52–73. https://doi.org/10.15133/j.os.2000.003

Tate, S. A., & Page, D. (2018). Whiteliness and institutional racism: Hiding behind (un)conscious bias. *Ethics and Education*, *13*(1), 141–155. https://doi.org/10.1080/17449642.2018.1428718

The NHS Staff Council. (2021). *Equality, diversity and inclusion training*: *A good practice guide*. NHS Employer. https://www.nhsemployers.org/publications/equality-diversity-and-inclusion-training [last accessed 19th December 2022]

Thielmann, I., Heck, D. W., & Hilbig, B. E. (2016). Anonymity and incentives: An investigation of techniques to reduce socially desirable responding in the Trust Game. *Judgment and Decision Making*, *11*(5), 527–536.

Tinsley, H. E. A., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology*, *34*, 414–424. https://doi.org/10.1037/0022-0167.34.4.414

Tobon, A. L., Flores, J. M., Taylor, J. H., Johnson, I., Landeros-Weisenberger, A., Aboiralor, O., Avila-Quintero, V. J., & Bloch, M. H. (2021). Racial implicit associations in psychiatric diagnosis, treatment, and compliance expectations. *Academic Psychiatry*, *45*(1), 23–33. https://doi.org/10.1007/s40596-020-01370-2

Vakola, M., Tsaousis, I., & Nikolaou, I. (2004). The role of emotional intelligence and personality variables on attitudes toward organisational change. *Journal of Managerial Psychology*, *19*, 88–110. https://doi.org/10.1108/02683940410526082

Watkins. (2018). Exploratory factor analysis: A guide to best practice. *Journal of Black Psychology*, *44*(3), 219-246. https://doi.org/10.1177/0095798418771807

Watts, S., & Stenner, P. (2005). Doing Q methodology: Theory, method and interpretation. *Qualitative Research in Psychology*, *2*(1), 67–91. https://doi.org/10.1191/1478088705qp022oa

Watts, S., & Stenner, P. (2012). *Doing Q methodological research: theory, method & interpretation: theory, method & interpretation* (1st ed.). SAGE Publications Ltd.

Webler, T., Danielson, S., & Tuler, S. (2007). *Guidance on the Use of Q Method for Evaluation of Public Involvement Programs at Contaminated Sites*. Social and Environmental Research Institute.

Wight, J. (2022). Racism in the NHS: Let's not rely on flawed surveys. *British Medical Journal*, *378*. https://doi.org/10.1136/bmj.o1752

Willburne, J., Wagstaff, D., Pomykal Franz, D., & Polly, D. (2021). Mathematics teachers' perceptions of practice: A Q-methodology study. *Operant Subjectivity*, *42*, 58–85. https://doi.org/10.15133/j.os.2020.003

Williams, C., & Bunn, S. (2022). *Mental Health Act Reform – Race and Ethnic Inequalities*. UK Parliament Post. https://post.parliament.uk/research-briefings/post-pn-0671/ [last accessed 19th December 2022]

York, K. (2019, December 2). *BAME representation and psychology*. British Psychological Society. https://www.bps.org.uk/psychologist/bame-representation-and-psychology

Zabala, A. (2014). Qmethod: A package to explore human perspectives using Q methodology. *R Journal*, *6*(2), 163-173. https://doi.org/10.32614/RJ-2014-032

3. PUBLIC DISSEMINATION DOCUMENT

Literature Review: The Relationship Between Perceived Racial Discrimination and Anxiety: A Comparative Meta-Analysis

The relationship between racial discrimination and poor mental and physical health has been examined and described over many research studies. To help provide an overview of all these findings, Paradies et al (2015) completed a meta-analysis. A meta-analysis is used to statistically combine data from multiple studies. This is useful for summarising the overall message of an evidence base, as well as for identifying similarities and differences across studies. Paradies et al (2015) found that racial discrimination was associated with a variety of poor mental and physical health outcomes, including both anxiety and depression. This is significant as anxiety and depression are the two most prevalent mental health difficulties found cross-culturally (Antunes et al., 2018; Baxter et al., 2012; Lo et al., 2020). An updated meta-analysis investigating the relationship between racial discrimination and depression was more recently completed (Britt-Spells et al., 2018). No such update was found for anxiety at the time of this review. The aim of the current review was to provide an updated meta-analysis on the relationship between racial discrimination and anxiety. Paradies et al (2015) was used as a comparative review.

A systematic search of several electronic journal databases was completed using key words relevant to the research topic. Studies published after the Paradies et al. (2015) search ended (October 2013) to June 2020 were systemically screened to find ones that were relevant for use in the current review. From the 2,384 initial studies screened, fifty-five primary studies that reported results on the relationship between racial discrimination and anxiety were included in the final review. When the results reported by each study were averaged, a small but statistically significant relationship was found between racial discrimination and anxiety.

However, individually the results reported by each study were significantly different from one another. This suggested that variations across individual studies (e.g., which questionnaire each study used to measure anxiety) had an influence on the relationship found by each study. Statistical tests were used to investigate this. These tests found that variations across studies did appear to have an impact on the results found by each study. For example, the relationship between PRD and anxiety was found to be statistically higher in under 18s than over 18s. The quality of each study also impacted on the relationship identified. Rated against certain criteria, the quality of the studies was found to fluctuate, and many had several issues which meant their quality was considered questionable. Studies that were rated as having have a high risk of bias in certain areas reported the relationship between racial discrimination and anxiety as significantly lower than studies rated as having a lower risk of bias.

Compared to Paradies et al (2015) this review found a similar overall relationship between racial discrimination and anxiety. It is hoped that the findings from this review can help to inform education, policy development, social care, and mental health services. Due to the identified variations between studies, it could be that this review found an underestimation of the true relationship between PRD and anxiety. It is hoped that the recommendations from this review can help to encourage future research to improve study quality, include more underrepresented groups, and refine the questionnaires used to identify a more accurate estimation of the relationship. Systematic reviews and meta-analyses should continue to be updated regularly to track any changes in the reported relationship, as well as the impact of any improvements to study quality.

References

Antunes, A., Frasquilho, D., Azeredo-Lopes, S., Neto, D., Silva, M., Cardoso, G., & Caldas-de-Almeida, J. M. (2018). Disability and common mental disorders: Results from the World Mental Health Survey Initiative Portugal. *European Psychiatry*, *49*, 56–61. https://doi.org/10.1016/j.eurpsy.2017.12.004

Baxter, Scott, K., Vos, T., & Whiteford, H. (2012). Global prevalence of anxiety disorders: A systematic review and meta-regression. *Psychological Medicine*, *43*(5), 897-910. https://doi.org/10.1017/S003329171200147X

Britt-Spells, A. M., Slebodnik, M., Sands, L. P., & Rollock, D. (2018). Effects of perceived discrimination on depressive symptoms among black men residing in the united states: A meta-analysis. *American Journal of Men's Health*, *12*(1), 52–63. https://doi.org/10.1177/1557988315624509

Lo, S. M., Wong, H. C., Lam, C. Y., & Shek, D. T. L. (2020). Common mental health challenges in a university context in Hong Kong: a study based on a review of medical records. *Applied Research in Quality of Life*, *15*(1), 207–218. https://doi.org/10.1007/s11482-018-9673-5

Paradies, Y., Ben, J., Denson, N., Elias, A., Priest, N., Pieterse, A., Gupta, A., Kelaher, M., & Gee, G. (2015). Racism as a determinant of health: a systematic review and meta-analysis. *PLOS ONE* , *9*(10). https://doi.org/10.1371/journal.pone.0138511

Empirical Research Paper: Perspectives on Addressing Racial Inequality in the NHS

The impact of racial inequality on NHS staff and service users has long been recognised (Esmail & Carnall, 1997; Kline, 2014; Wight, 2022) Despite efforts to address racial inequality, disparities are still evident (Bamrah & Chakravorty, 2022; Esmail & Carnall, 1997; NHS England, 2022a). The Workforce Race Equality Standard (WRES) collects data on NHS staff's experiences and outcomes to support equal access to career opportunities and fair treatment. The most recent WRES reported inequalities across recruitment, training, representation, career-progression, experiences of unfair treatment, and the highest ever self-reported level of discrimination since the first WRES report in 2015 (Adebowale & Rao, 2020; Iacobucci, 2020; NHS, 2022). Furthermore, ethnic minority service users are more likely to report negative healthcare experiences (Ben et al., 2017). The severity of this extends to the higher detention of ethnic minority service users under the Mental Health Act and increased maternal mortality rates for ethnic minority women (Crown, 2018; Mann, 2014; Nair et al., 2014).

Ethnic minority people's continued experiences of racial inequality have raised the question as to whether the broad range of solutions that have been employed are effective or helpful (Gay & Bamford, 2007; Hassen et al., 2021; Kar, 2020). However, the complex and nuanced nature of individual behaviour and organisational change means there are many potential barriers to these solutions being effective. For example, promoting individual members of staff to tackle their own implicit biases is limited if the system around them is systemically racist (Noon, 2018). A sense that many solutions have been tried and fallen short is summarised by Adeyemi (2019):

"…what hashtag, conference, workshop, diversity training, evidence-gathering process have we not engaged with to bring about change? What is that internalised barrier that keeps the NHS from addressing the problem once and for all?" (Adeyemi, 2019).

New approaches being applied to health care innovation research (Chan, 2018; McLaughlin et al., 2019) suggests that the most effective solutions are found after developing a strong understanding of stakeholders' (e.g., staff and service user) needs and challenges (Altman et al., 2018; Roberts et al., 2016). Research into the perspectives of stakeholders on how they feel racial inequality in the NHS could be most effectively addressed is limited (Anekwe, 2020). The aim of this study was to use Q-methodology, an approach that facilitates the scientific exploration of people's opinions, to fill this gap in the research (Combes et al., 2004; Herrington & Coogan, 2011; Ladan et al., 2018; QMethod Software, 2022; Stephenson, 1935). Q-methodology is made up of four main stages a) gathering as many opinions as possible on a given topic (e.g., 'more ethnic minority mentors should be available') and writing those opinions onto cards, b) recruiting participants who have important perspectives on the topic, c) asking participants to sort the opinion cards based on their own perspective, d) statistically analysing the data. In this study, both written materials and interviews were used to create 65 opinion cards. Twenty-eight people sorted those cards along a scale of 'least effective' to 'most effective'. This card sort was completed online independently by participants.

Looking at the way all 28 people sorted the cards; 3 main opinion groups were identified. Group 1 prioritised the individual responsibility of staff to reflect on their own biases and behaviours (e.g., each member of staff must personally think addressing racial inequality is important). Group 2 spotlighted research as being most effective for addressing

racial inequality (e.g., more research is needed on the health effects of racial inequality). Group 3 prioritised corporate solutions (e.g., the use of zero tolerance polices). Certain solutions were rated similarly across all opinion groups (e.g. that broad brush solutions were least effective) and certain solutions were rated differently across all groups (e.g. every service user should be completed the same).

These findings could be used to develop specific strategies to address racial inequality. For example, implementing more corporate solutions as suggested by Group 3. However, as each opinion group had different ideas about what would be most effective, and some perspectives were conflicting across groups, meeting the needs of all stakeholder's is unlikely if a one size fits all approach is taken. This could be a reason as to why so many solutions have already been tried but racial inequality continues to perpetuate in the NHS. As such, a major recommendation of this research was for policy makers, leadership teams, staff, and service users to acknowledge that stakeholders are likely to have differing, and at times conflicting, opinions. This should be taken this into account when trying to address racial inequality. As using Q-methodology and exploring stakeholders' perspectives was considered a novel approach, it is recommended that further research is conducted and built upon in the further.

References

Adebowale, V., & Rao, M. (2020). Racism in medicine: Why equality matters to everyone. *British Medical Journal, 368*. https://doi.org/10.1136/bmj.m530

Adeyemi, A. (2019, October 14). The future of race in the NHS. *NHS England.* https://www.england.nhs.uk/blog/the-future-of-race-in-the-nhs/ [last accessed 19th December 2022]

Altman, M., Huang, T. T. K., & Breland, J. Y. (2018). Design thinking in health care. *Preventing Chronic Disease, 15*(117). https://doi.org/10.5888/pcd15.180128

Bamrah, J. S., & Chakravorty, I. (2022). Racism in the NHS: Let BMA report be last needed so we can return to tackling NHS crisis. *British Medical Journal, 378*. https://doi.org/10.1136/bmj.o2064

Ben, J., Cormack, D., Harris, R., & Paradies, Y. (2017). Racism and health service utilisation: A systematic review and meta-analysis. *PloS One, 12*(12). https://doi.org/10.1371/journal.pone.0189900

Chan, K. (2018). A design thinking mindset beyond the public health model. *World Medical & Health Policy, 10*(1), 111- 119. https://www.doi.org/10.14707/ajbr.210102

Combes, H., Hardy, G., & Buchan, L. (2004). Using Q-methodology to involve people with intellectual disability in evaluating person-centred planning. *Journal of Applied Research in Intellectual Disabilities, 17*(3), 149–159. https://doi.org/10.1111/j.1468-3148.2004.00191.x

Crown. (2018). *Modernising the Mental Health Act – final report from the independent review*. GOV.UK. https://www.gov.uk/government/publications/modernising-the-mental-health-act-final-report-from-the-independent-review

Esmail, A., & Carnall, D. (1997). Tackling racism in the NHS: We need action not words. *British Medical Journal*, *314*(7081). https://doi.org/10.1136/bmj.314.7081.618

Gay, W., & Bamford, D. (2007). A case study into the management of racial diversity within an NHS teaching hospital. *International Journal of Public Sector Management*, *20*(4), 257–271. https://doi.org/10.1108/09513550710749996

Hassen, N., Lofters, A., Michael, S., Mall, A., Pinto, A. D., & Rackal, J. (2021). Implementing anti-racism interventions in healthcare settings: A scoping review. *International Journal of Environmental Research and Public Health*, *18*(6). https://doi.org/10.3390/ijerph18062993

Herrington, N., & Coogan, J. (2011). Q methodology: An overview. *Research in Teacher Education*, *1*(2). https://doi.org/10.15123/uel.8604v

Iacobucci, G. (2020). NHS blood unit is "systemically racist" and "psychologically unsafe," investigation finds. *British Medical Journal, 369l*. https://doi.org/10.1136/bmj.m2308

Kar, P. (2020). To tackle racism, the NHS needs policies with teeth. *British Medical Journal*, *369*. https://doi.org/10.1136/bmj.m2583

Kline, R. (2014). *The snowy white peaks of the NHS: A survey of discrimination in governance and leadership and the potential impact on patient care in London and England*. Middlesex University. https://doi.org/10.22023/mdx.12640421.v1

Ladan, M. A., Wharrad, H., & Windle, R. (2018). Towards understanding healthcare professionals' adoption and use of technologies in clinical practice: Using Q-methodology and models of technology acceptance. *Journal of Innovation in Health Informatics*, *25*(1), 965. https://doi.org/10.14236/jhi.v25i1.965

Mann, C. (2014). *The effect of mindfulness on racial stereotype activation and application.* [Doctoral dissertation, Bryn Mawr College]. https://doi.org/10.22023/mdx.12640421.v1

McLaughlin, J. E., Wolcott, M. D., Hubbard, D., Umstead, K., & Rider, T. R. (2019). A qualitative review of the design thinking framework in health professions education. *BMC Medical Education*, *19*(1), 98. https://doi.org/10.1186/s12909-019-1528-8

Nair, M., Kurinczuk, J. J., & Knight, M. (2014). Ethnic variations in severe maternal morbidity in the UK- a case control study. *PloS One*, *9*(4). https://doi.org/10.1371/journal.pone.0095086

NHS. (2022). *NHS England Workforce Race Equality Standard 2021*. NHS England. https://www.england.nhs.uk/publication/workforce-race-equality-standard-2021/

NHS England. (2022a). *NHS workforce race equality standard*. NHS England. https://www.england.nhs.uk/about/equality/equality-hub/workforce-equality-data-standards/equality-standard/ [last accessed 19th December 2022]

Noon, M. (2018). Pointless diversity training: unconscious bias, new racism and agency. *Work, Employment and Society*, *32*(1), 198–209. https://doi.org/10.1177/0950017017719841

QMethod Software. (2022). *Q Methodology: An Overview* [Computer software]. QMethod Software. https://qmethodsoftware.com/q-methodology/

Stephenson, W. (1935). Technique of factor analysis. *Nature*, *136*(3434). https://doi.org/10.1038/136297b0

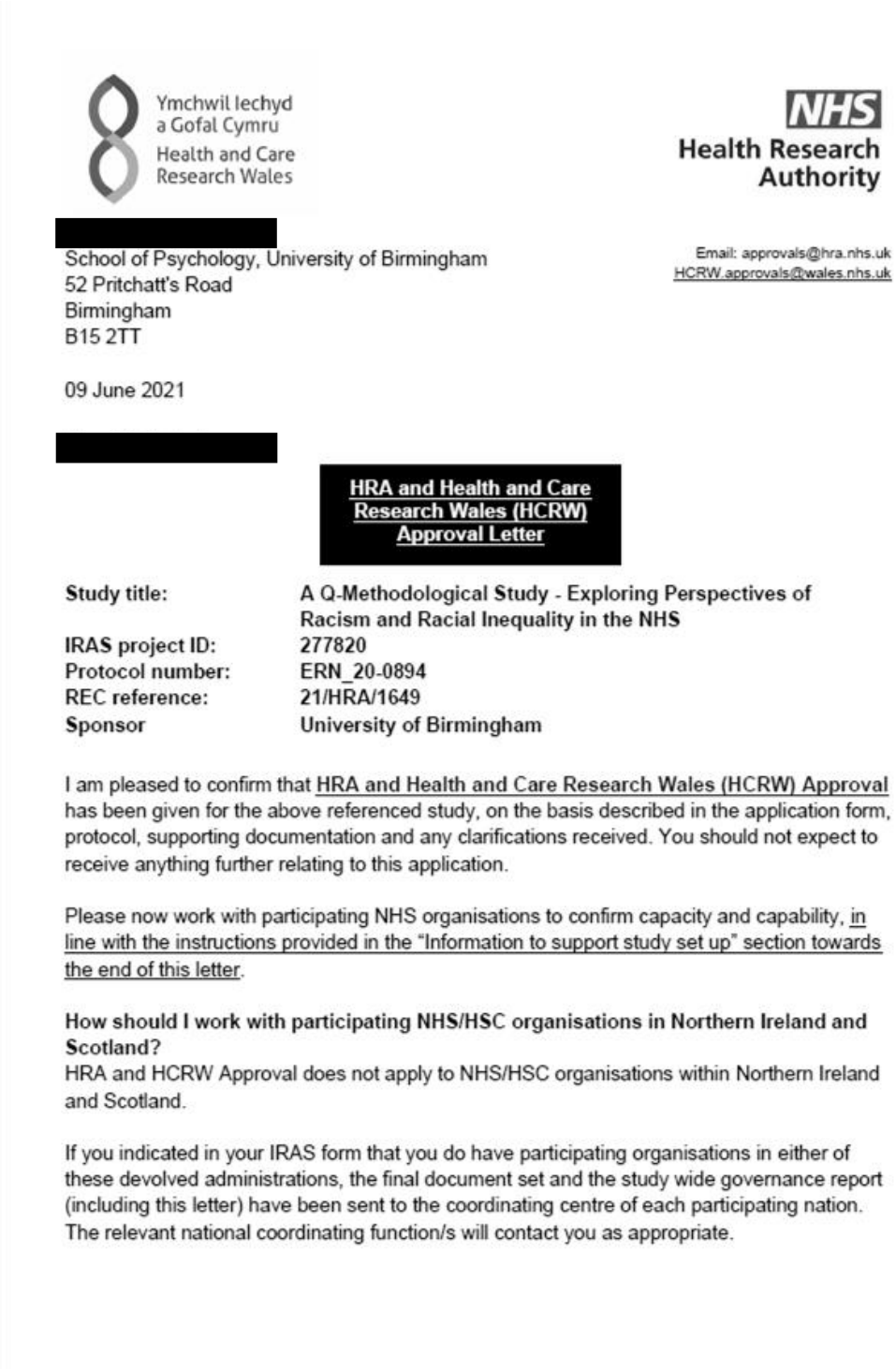
Wight, J. (2022). Racism in the NHS: Let's not rely on flawed surveys. *British Medical Journal*, *378*. https://doi.org/10.1136/bmj.o1752

Appendices

Appendix A – Letter of Ethnical Approval

Figure A

Health Research Authority Approval Letter



Ymchwil Iechyd a Gofal Cymru
Health and Care Research Wales

NHS
Health Research Authority

[redacted]
School of Psychology, University of Birmingham
52 Pritchatt's Road
Birmingham
B15 2TT

Email: approvals@hra.nhs.uk
HCRW.approvals@wales.nhs.uk

09 June 2021

[redacted]

HRA and Health and Care Research Wales (HCRW) Approval Letter

Study title:	**A Q-Methodological Study - Exploring Perspectives of Racism and Racial Inequality in the NHS**
IRAS project ID:	**277820**
Protocol number:	**ERN_20-0894**
REC reference:	**21/HRA/1649**
Sponsor	**University of Birmingham**

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 **277820**. Please quote this on all correspondence.

Yours sincerely,
Mathew Barnes

Approvals Specialist

Email: approvals@hra.nhs.uk

Copy to: *Dr Birgit Whitman*

Appendix B – Interview Guide

"[Q-sort] interviews are best done in a semi-structured format. Allow conversations to flow freely so that the interviewee can raise the points most important to them and frame the issues in the way that they ordinarily think. The goal is to generate a database of natural-language statements about the topic." (Webler, Danielson & Tuler, 2007)

- What are your understandings or feelings about the terms racial inequality and racism?

- Is RI and Racism something that needs addressing in the NHS?
    - *What is the biggest problem?*
    - *Is this something that can't be fixed?*
    - *Prevention – on-going – post*

- In your opinion, what would improve racial equality and reduce racism/racial inequality in the NHS?
    - *Are there particular areas that need to be spotlighted?*
    - *Are you aware of any initiatives or policies aimed at addressing racism/racial inequality in the NHS?*
        - *What needs to be done more?*
        - *What are their weaknesses or short-comings?*
        - *Are there things that have been tried that haven't worked*
        - *Are there any strategies or initiatives in other organisations that you think would be applied helpfully.*

- Are there nuances to your suggestions?/Would your suggestions need to be completed in a particular way?

- How would we know interventions were making a difference? / What change would you like to see in the next 5 years? How could this be measured?

- What other thoughts/beliefs/opinions have you come across?
    - *What similar thoughts to your own have you come across?*
    - *Have you come across thoughts that are different to your own?*