Virtual Reality (2024) 28:125
https://doi.org/10.1007/s10055-024-01024-w

ORIGINAL ARTICLE

Perspective matters: a systematic review of immersive virtual reality to reduce racial prejudice

Sarah Higgins[1] · Stephanie Alcock[1] · Bianca De Aveiro[1] · William Daniels[2] · Harry Farmer[3] · Sahba Besharati[1]

Received: 6 November 2023 / Accepted: 10 June 2024
© The Author(s) 2024

Abstract

In the wake of the COVID-19 pandemic and the rise of social justice movements, increased attention has been directed to levels of intergroup tension worldwide. Racial prejudice is one such tension that permeates societies and creates distinct inequalities at all levels of our social ecosystem. Whether these prejudices are present explicitly (directly or consciously) or implicitly (unconsciously or automatically), manipulating body ownership by embodying an avatar of another race using immersive virtual reality (IVR) presents a promising approach to reducing racial bias. Nevertheless, research findings are contradictory, which is possibly attributed to variances in methodological factors across studies. This systematic review, therefore, aimed to identify variables and methodological variations that may underlie the observed discrepancies in study outcomes. Adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, this systematic review encompassed 12 studies that employed IVR and embodiment techniques to investigate racial attitudes. Subsequently, two mini meta-analyses were performed on four and five of these studies, respectively — both of which utilised the Implicit Association Test (IAT) as a metric to gauge these biases. This review demonstrated that IVR allows not only the manipulation of a sense of body ownership but also the investigation of wider social identities. Despite the novelty of IVR as a tool to help understand and possibly reduce racial bias, our review has identified key limitations in the existing literature. Specifically, we found inconsistencies in the measures and IVR equipment and software employed, as well as diversity limitations in demographic characteristics within both the sampled population and the embodiment of avatars. Future studies are needed to address these critical shortcomings. Specific recommendations are suggested, these include: (1) enhancing participant diversity in terms of the sample representation and by integrating ethnically diverse avatars; (2) employing multi-modal methods in assessing embodiment; (3) increasing consistency in the use and administration of implicit and explicit measures of racial prejudice; and (4) implementing consistent approaches in using IVR hardware and software to enhance the realism of the IVR experience.

Keywords Race · Racial prejudice · Immersive virtual reality · Embodiment · Perspective-taking · Implicit bias

✉ Harry Farmer
h.farmer@gre.ac.uk

✉ Sahba Besharati
sahba@besharati.com

Sarah Higgins
sarahhiggins262@gmail.com

Stephanie Alcock
stephanie.alcock@wits.ac.za

Bianca De Aveiro
biancadeaveiro@gmail.com

William Daniels
William.Daniels@wits.ac.za

1 Department of Psychology, School of Human and Community Development, University of the Witwatersrand, Johannesburg, South Africa

2 School of Physiology, University of the Witwatersrand, Johannesburg, South Africa

3 School of Human Sciences, Institute of Lifecourse Development, University of Greenwich, London, UK

Published online: 21 June 2024

1 Introduction

In the last 20 years, the use of immersive virtual reality (IVR) has gained a significant amount of traction in the social and cognitive neurosciences. By integrating feedback from sensory (seeing oneself in a virtual body or avatar) and motor signals (moving in real-time with the avatar), multisensory integration is achieved to create a full-body illusion (Blanke 2012; Slater et al. 2010). IVR, therefore, offers a powerful tool to manipulate the sense of body ownership (i.e., the feeling that your body belongs to you; Gallagher 2000), going far beyond the pioneering studies of the rubber hand illusion (Botvinick and Cohen 1998). Accordingly, IVR has had diverse applications in multidisciplinary fields, such as in neurorehabilitation (Demeco et al. 2023), education (Hodgson et al. 2019), and visual perception (for an overview, see Wilson and Soranzo 2015). Drawing on this embodied cognition framework, social neuroscientists have utilised IVR methods to understand the psychological mechanisms involved in feelings of prejudice, especially in relation to perceptions of race and implicit racial attitudes (Farmer and Maister 2017; Maister et al. 2015; Peck et al. 2013).

1.1 Implicit bias and the implicit association test (IAT)

Social attitudes are described as an individual's predisposition to behave in a particular way towards another individual or a group. These attitudes involve cognitive elements, including beliefs, evaluations, opinions, and social emotions (Poggi and D'Errico 2010). Implicit attitudes and/or biases refer to unconscious or automatic mental associations that are typically thought to arise as a product of one's internalised schemas and subsequently guide one to partake in discriminatory behaviours without conscious intent (FitzGerald and Hurst 2017). Conversely, explicit biases refer to preferences, beliefs, and attitudes that a person is consciously aware of and can identify (Dovidio and Gaertner 2010). Considering that implicit attitudes represent automatic associations, they were initially considered to be relatively uncontrollable in nature. Consequently, researchers hypothesised that external environmental cues could not alter them under any circumstances or conditions (Jost et al. 2004). However, current research highlights the social and contextual sensitivity of implicit attitudes, suggesting that they are highly malleable and receptive to even the subtlest environmental influences (Dasgupta and Greenwald 2001; Lowery et al. 2001).

Implicit attitudes are most frequently measured with the use of the IAT – with researchers relying heavily on this tool, particularly for race-based research (Schnabel et al. 2008; Banakou et al. 2016; Peck et al. 2013). Developed in 1998, the IAT represents an indirect measure of the associative strength between a bipolar target (i.e., self versus others) and a bipolar attribute (i.e., dark-skinned versus light-skinned; Greenwald et al. 1998; Schnabel et al. 2008). The speed and accuracy at which each task in the test is successfully categorised serve as indicators of the level of bias (Greenwald et al. 1998). While the continued use of the IAT in social-cognitive research is supported based on its behavioral predictive value, especially in sensitive contexts (Nock et al. 2010), and its applicability as an alternative to explicit self-reporting measures, which present with numerous limitations (Schnabel et al. 2008), several criticisms have been identified challenging its validity. Critiques of the test encompass concerns about its lack of psychometric quality, susceptibility to influence from social-contextual factors, and reinforcement of cultural stereotypes (Barden et al. 2004; Dasgupta and Greenwald 2001; Schimmack and Howard 2021). Nevertheless, it remains as a benchmark instrument of implicit bias in the field.

1.2 Implicit racial bias in IVR

Virtual reality (VR) as a general technology has many approaches and is a rapidly developing field. VR technology has evolved to develop a range of hardware and software that create realistic and typically three-dimensional computer-based interfaces that are becoming increasingly more affordable for public and scientific use (Anthes et al., 2016). However, the immersive nature of the virtual reality experience can differ, with immersive VR (as outlined above) requiring specific requirements, such as multisensory integration (see Slater and Sanches-Vives, 2016), to create a full-body illusion. In a recent study, Tassinari and colleagues (2022) conducted a systematic review of 64 studies using general VR methods to reduce different forms of prejudice (e.g., racial, gender, HIV-stigma, disability). The results highlighted the increased prevalence of utilising VR methods in studying inter-group contact and prejudice reduction. However, IVR more specifically, has emerged as a promising tool for investigating implicit biases, given that this methodology can induce a temporary illusion of embodying an outgroup virtual avatar. Similarly, IVR harnesses the potential to capture and analyse behavioural cues related to emotions and implicit attitudes, thereby offering a more immersive and ecologically valid environment for such investigations (for an overview, see Jacob-Dazarola et al. 2016).

This process involves experiencing the virtual body from a first-person perspective, enabling individuals to perceive the virtual body as their own. In doing so, IVR possibly reduces an individual's psychological distance to

abstract mental construals (Kalyanaraman & Bailenson, 2019), which can change distant attitude objects to near ones (Nikolaou et al., 2022). This transformation is essential for altering social attitudes as attitudes toward objects perceived as psychologically near tend to change more than those towards distant attitude objects (Trope & Liberman, 2010). Alterations in social attitudes, namely implicit biases, through the use of IVR, have been observed for age, disability, and gender. For example, in an attempt to support their Proteus Effect hypothesis (i.e., the tendency to alter behaviour and self-representation in response to adopting the persona of a virtual body), Yee and Bailenson (2007) showed that the embodiment of an elderly person can reduce implicit prejudices and negative stereotypes held against senior populations. Similar positive effects have been found for embodying disabled (Chowdhury et al. 2019) and different gendered avatars (Wu & Chen, 2022). Together these findings demonstrate the possible application of IVR technology in positively altering implicit biases.

Nevertheless, the literature surrounding the alteration of implicit racial attitudes seems to be incongruent. For example, research has demonstrated that perspective-taking tasks involving perceptual ownership of an alternative racial body in IVR can have a positive effect by reducing racial biases (Banakou et al. 2016; Farmer et al. 2012). Reduced in-group favouritism (Dovidio et al. 2004), implicit racial biases (Banakou et al. 2016; Forscher et al. 2019), and explicit stereotypes (Galinsky and Moskowitz 2000) have also been attributed to IVR embodiment interventions that alter perception. Moreover, several studies have documented a positive change in implicit racial bias towards Black individuals when White participants perceive an illusory rubber arm belonging to a Black virtual body as their own (Blanke et al. 2015; Blanke 2012; Ehrsson 2012; Farmer et al. 2012). Although the aforementioned studies are representative of partial-body-ownership illusions, many researchers have since investigated the effects of full-body-ownership illusions and have reached similar conclusions (Banakou et al. 2016; Behm-Morawitz et al. 2016; Groom et al. 2009; Hasler et al. 2017; Peck et al. 2013; Thériault et al. 2021).

In contrast to the aforementioned studies, research has reported either increased or minimal to no change in performance on measures of bias after being subjected to embodiment and/or interpersonal interaction in IVR (Groom et al. 2009; Hasler et al. 2017; Thériault et al. 2021). Moreover, Rossen et al. (2008) revealed that medical students were less likely to express empathy for dark-skinned virtual patients than for light-skinned patients. Using the same racial ingroup and outgroup, D'Errico et al. (2020) instructed participants to choose to help an ingroup (White) and outgroup (Black) virtual confederate, dressed either casually, as a businessman or a beggar. They found that empathy was greater when participants interacted with ingroup (White) virtual avatars compared to outgroup (Black) virtual avatars. Similarly, a study using interpersonal distance and physiological measures as an indication of racial prejudice showed that participants were fearful of racial minority avatars (Dotsch and Wigboldus 2008). Collectively, these studies indicate that racial biases may extend into virtual intergroup encounters, which is possibly attributed to race representing a particularly potent cue for prejudicial categorisation due to its visual salience (Cosmides et al. 2003). However, comprehending the context and characteristics of the participants is essential for a precise interpretation of their fear responses.

1.3 Shortcomings of IVR racial research

While research using IVR technology and racial biases has seen significant advancements, the inconsistent findings point towards notable shortcomings in this field. It has been suggested that the inconclusive findings could arise from several factors, including stereotypical depictions of minority groups in IVR scenarios, methodological imprecision, or a limited grasp of the socio-cognitive mechanisms and moderating variables influencing intergroup bias, especially within the context of embodiment (Chen et al. 2021). Similarly, research in this domain represents a racial issue in and of itself since researcher-related and methodological complications may contribute to the results and the interpretation thereof in the final publication. Persons involved in the research process (i.e., experimenters, authors, participants, and editors) are often systematically connected, with a notable trend of a few authors incorporating and examining diverse populations for the purpose of their study (Roberts et al. 2020). Therefore, it is imperative that we acknowledge the sensitivity of geographical and social context in racial research as well as the diversity of external factors that may contribute to the final results.

Furthermore, with an increase in researchers capitalising on the potential of IVR to reduce racial prejudice, it becomes imperative to consider additional methodological and measurement-related implications in the research process. These include the degree of immersion a participant may experience with their embodied avatar, the complexities of using IVR technology, and the details pertaining to the applied measurement of racial bias or prejudice. Therefore, this systematic review synthesises articles that have used IVR to investigate and alter implicit and/or explicit racial prejudice with the aim of understanding how virtual embodiment may contribute to our racial and social beliefs, opinions, and prejudices. More specifically, this review has a specific focus on determining the effectiveness of IVR interventions in altering racial prejudice. We also aim to conduct a critical analysis of the representativeness and external

validity of the included studies, identify both methodological and technological strengths and limitations inherent in race-related IVR studies, and explore how these factors may impact the outcomes of current research. Finally, we aim to outline the implications of these findings for future research in this domain. As a supplementary assessment, we performed a meta-analysis on the application of the IAT in investigating racial attitudes using the articles encompassed in this review.

Accordingly, the main research questions guiding this review were as follows:

RQ1 : What are the characteristics of the studies and degree of representation of the sample and avatars used in IVR studies?

RQ2 : What measures are used to assess implicit and explicit racial biases in IVR studies?.

RQ3 : To what extent do the results of the IAT differ across various studies focusing on IVR embodiment and racial implicit bias?

RQ4 : Has embodiment using IVR been successful in eliciting a reduction in racial prejudice?

Table 1 Summary of search terms and corresponding boolean operators

Primary terms	
Population (race/ethnicity)	*"race" OR "racial" OR "racism" OR*
	"ethnic OR "tribal" OR*
	"white" OR "black" OR "non-white" OR
	"dark skin" OR "light skin*" / "dark-skin*" OR "light-skin*" OR*
	"ingroup" OR "outgroup" OR
	"minority" OR "majority"
Virtual reality	*"VR" OR "virtual reality" OR*
	"immersive virtual reality" OR
	"immersive virtual environment" OR
	"virtual reality exposure therapy" OR
	"augmented reality" OR
	"embodiment" OR "embody"*
Prejudice	*"prejudice" OR*
	"discriminate" OR*
	"bias" OR*
	"stereotype" OR*
	"stigma" OR*
	"intergroup attitude"*

Note * = Represents a wildcard, matching any value(s) in a boolean expression

2 Methods

This review followed the Preferred Reporting Items for Systematic Reviews (PRISMA) guidelines and has been pre-registered in the PROSPERO database (https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42022325576). All deviations from the pre-registration are distinctly documented in this review. We undertook a comprehensive systematic review by searching through four distinct online databases, namely PsychINFO, Embase, MEDLINE, and Global Health (all of which were accessed via OvidSP - a search platform enabling access to a variety of international databases, journals, and books). Although none were found, we attempted to source additional records from the reference lists of selected articles during the screening process. The termination point for the database search was 03.10.2022. Table 1 illustrates the employed search terms.

The inclusion and exclusion criteria were outlined according to the Population, Intervention, Comparator and Outcomes (PICO) framework (Richardson et al. 1995). In terms of population, the review included studies examining neurologically healthy adults and/or children capable of using IVR technology and embodiment. As the intervention approach, participants were required to engage with IVR technology through full-body embodiment of an avatar from a different racial background (i.e., a race distinct from their own). Full-body embodiment refers to the perceptual and cognitive experience of feeling fully present and identified with a virtual body within a 3D virtual environment. This is achieved by providing participants with a virtual representation of their own body or another person's body, seen from a first-person perspective, and creating a sense of ownership and agency over that virtual body (Slater 2017). The results of the intervention strategy were evaluated in contrast to comparators, including non-treatment and treatment comparators, which refer to conditions in the studies that are used to compare the effectiveness of the intervention (i.e., full-embodiment of a differently-raced virtual body). The comparators comprised various conditions, including groups that (1) did not receive any intervention (non-treatment comparator); (2) experienced a non-IVR intervention (active comparators such as perspective-taking exercises, viewing videos, playing two-dimensional (2D) three-dimensional (3D) video games); (3) experience an IVR condition that involves embodying or interacting with an in-group avatar (i.e., an avatar identical to their own race); or (4) experience the partial embodied of a differently raced avatar (e.g., embodiment via the rubber hand illusion). These comparators serve as a baseline against which the effects of the intervention (i.e., full-embodiment of a differently-raced avatar) can be measured to distinguish its

effect. Specifically, comparators one and two isolate IVR and embodiment intervention effects. Comparator three isolates the influence of racial embodiment, and comparator four isolates the specific effects of full embodiment.

The included articles used either quantitative or qualitative measures of implicit and/or explicit racial bias/prejudice. This encompasses various measurements such as reaction time assessments, including the IAT, evaluations of racism-related attitudes, racial prejudice, or bias, response time evaluations, measurements of behavioural or interpersonal proximity, physiological indicators (i.e., skin conductance and heart rate measurements), and any other qualitative evaluations of racial prejudice. Articles were excluded based on whether they were classified as reviews or opinion pieces, were unpublished or had not yet been peer-reviewed. Additionally, articles targeting non-race-related prejudice (i.e., ageism, sexism, disease-related prejudice) and studies not written or published in English were excluded for the purpose of this systematic review.

Following the electronic database search, we transferred our search results to Zotero – a referencing management software – to remove duplicate articles. Our preliminary search results were then exported to Rayyan (Ouzzani et al. 2016) – a collaborative online tool for systematic reviewers – to screen articles according to the established set of inclusion and exclusion criteria. The screening process involved three independent phases: title and abstract screening, full-article screening, and conflict resolution. Two independent reviewers (SH & SA) performed an initial screening of the titles and abstracts in accordance with the aforementioned criteria. Subsequently, a full-text review was performed to confirm the eligibility of each article for inclusion. A third reviewer (BD) was then consulted to resolve any conflicts. The risk of bias was minimised using Rayyan's 'blind' option, which ensured that each reviewer could not view their collaborators' screening decisions. Additionally, we evaluated the methodological rigour for each paper utilising the Joanna Briggs Institute (JBI) critical appraisal tool (as detailed in the Supplementary Information). The relevant extracted information was organised into three data extraction tables (seen Tables 2, 3 and 4).

Thereafter, two random effects meta-analyses were conducted to augment the outcomes and strengthen the statistical power of the systematic review. Incorporating eight out of 12 eligible papers, the meta-analysis focused on implicit bias assessment through the IAT. However, the methods of measuring IAT scores varied among the studies. Specifically, five of the eight articles reported Post-IAT scores, indicating assessments conducted after the experimental conditions, while four others reported difference in IAT (dIAT) scores, reflecting the difference between Pre- and Post-test IAT scores. Consequently, two distinct mini meta-analyses were conducted: one utilising Post-IAT scores ($n=5$) and the other using dIAT scores ($n=4$) from these studies. Within the context of these meta-analyses, the effects observed across primary studies were transformed into standardised effect sizes. These effect sizes were calculated for each study by subtracting the average score of the experimental group, which embodied a racial outgroup avatar in IVR, from the average score of the control group (i.e., participants who either embodied their own-race avatars or engaged in different activities, such as perspective-taking interventions). The resulting value was divided by the combined standard deviations of both groups. The relevant data were extracted and organised in Microsoft Excel, with analyses and figures conducted using the "meta" package for R (Schwarzer et al. 2015).

3 Results

3.1 Search results

A total of 681 articles were first identified. After removing duplicates, 479 articles were screened according to their title and abstracts. Subsequently, 18 full-text papers were screened, 12 of which met the eligibility criteria and were included in the review. Figure 1 depicts the PRISMA flowchart, illustrating the progression of information across the various stages of the review.

3.2 Characteristics of the studies and samples

Seven studies, accounting for half of the total number, were published between 2020 and 2022, thereby demonstrating the growing interest and progress of IVR research in the cognitive sciences in investigating the question of racial prejudice. Of the remaining articles, five studies were published between 2013 and 2018, and the earliest publication was dated 2009. All studies except for one, which was conducted in Singapore (Chen et al. 2021), were undertaken in the global North, with the majority of studies originating from Spain ($n=4$) and the United States of America (USA; $n=3$). Most studies used a between-group design ($n=10$), with additional study designs including repeated measures ($n=1$) and mixed factorial research design ($n=1$).

Sample sizes ranged from 32 to 171 participants, and all studies consisted of young participants, with a mean age ranging from 21 to 38.5 years old. Every study had a higher percentage of female participants, with some studies including only female participants. Limited information was provided regarding the socio-economic status of participants, with most articles only reporting that their sample comprised university students ($n=9$). Additional data

Table 2 Details of included studies

Title	Authors	Location of study	Aim	Design	Type of interaction
Body swapping with a Black person boosts empathy: Using virtual reality to embody another	Thériault et al. (2021)	Canada	Investigating the effects of mental versus embodied perspective-taking on prejudice.	Between-subjects	Mental Perspective Taking (MPT) and Embodied Perspective Taking (EPT)
Increasing Self-Other Similarity Modulates Ethnic Bias in Sensorimotor Resonance to Others' Pain	Harjunen et al. (2021)	Finland	Investigate if increasing self-other similarity via virtual transfer to another colour body reduces ingroup bias in sensorimotor resonance.	Within-subjects	Embodiment
Virtual race transformation verses racial ingroup bias	Hasler et al. (2017)	Spain	Examine if there are differences in body ownership with respect to the embodiment of White people in a White or Black virtual body and if this changes racial bias and social mimicry.	Mixed factorial	Embodiment & interaction
Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias	Banakou et al. (2016)	Spain	Investigate if there is any sustained reduction in implicit bias that lasts at least 1 week after embodiment and if it is enhanced by multiple exposures.	Between-subjects	Embodiment & interaction
Putting yourself in the skin of a black avatar reduces implicit racial bias	Peck et al. (2013)	Spain	Determine if a sense of body ownership can be induced in a differently raced avatar, and to determine if this illusion could reduce negative implicit responses toward the other race.	Between-subjects	Embodiment & non-embodiment
Investigating the Influence of Intergroup Contact in Virtual Reality on Empathy: An Exploratory Study Using AltspaceVR.	Tassinari et al. (2022a)	Finland	Investigate how VR can be used to create positive intergroup contact with a member of a stigmatised outgroup and present the results of the effect of intergroup contact in VR on empathy.	Between-subjects	Embodiment & interaction
The Effect of VR Avatar Embodiment on Improving Attitudes and Closeness Toward Immigrants.	Chen et al. (2021)	Singapore	Investigate the effect of VR embodiment on attitudes and closeness toward an outgroup vs. ingroup in novel co-ethnic immigration context in Asia.	Between-subjects	Embodiment & interaction
Virtual body ownership and its consequences for implicit racial bias are dependent on social context.	Banakou et al. (2020)	Spain	Investigate if the social context of a virtual environment influences the extent to which participants experience a reduction in implicit bias	Between-subjects	Embodiment
The impact of virtual reality on implicit racial bias and mock legal decisions.	Salmanowitz (2018)	USA	Examine the influence of a short virtual reality paradigm on implicit racial biases and evaluations of legal scenarios.	Between-subjects	Embodiment & non-embodiment (sham version)
Contact in VR: Testing avatar customisation and common ingroup identity cues on outgroup bias reduction.	Alvidrez and Peña (2020)	USA	Explore how customising avatars to resemble self, together with the salience of common ingroup identity cues affected intergroup perceptions in a VR interaction.	Between-subjects	Embodiment
The influence of racial embodiment on racial bias in immersive virtual environments.	Groom et al. (2009)	USA	Determine how one's implicit bias is affected by the race of their avatar in an immersive virtual environment.	Between-subjects	Imagined embodiment & embodiment
Exploring the Effect of Cooperation in Reducing Implicit Racial Bias and Its Relationship With Dispositional Empathy and Political Attitudes.	Patané et al. (2020)	France	To examine the effects of some socio-cognitive variables that could contribute to enhancing or reducing the implicit bias	Between-subjects	Embodiment & interaction

Table 3 Sample demographics

Title	Authors	Total *N*	% female	Mean age (SD)	Ethnic composition	SES
Body swapping with a Black person boosts empathy: Using virtual reality to embody another	Thériault et al. (2021)	90	71	22.2 (3.0)	White (50%) Asian (26%) South Asian (17%) Other (7%)	Not stated
Increasing Self-Other Similarity Modulates Ethnic Bias in Sensorimotor Resonance to Others' Pain	Harjunen et al. (2021)	58	60.3	26.21 (6.44)	White (100%)	University students
Virtual race transformation verses racial ingroup bias	Hasler et al. (2017)	32	100	Not stated	White (100%)	University students
Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias	Banakou et al. (2016)	90	100	21.9 (4.45)	White (100%)	University students
Putting yourself in the skin of a black avatar reduces implicit racial bias	Peck et al. (2013)	60	100	Not stated	White (100%)	University students
Investigating the Influence of Intergroup Contact in Virtual Reality on Empathy: An Exploratory Study Using AltspaceVR.	Tassinari et al. (2022a)	64	Not stated	EXP: 20 CONT: 17	White (93.8%) Asian (6.2%)	University and some high school students
The Effect of VR Avatar Embodiment on Improving Attitudes and Closeness Toward Immigrants.	Chen et al. (2021)	171	58.5	22.43 (2.07)	Singaporean Chinese (100%)	University students
Virtual body ownership and its consequences for implicit racial bias are dependent on social context.	Banakou et al. (2020)	92	100	21.8	White (100%)	University students
The impact of virtual reality on implicit racial bias and mock legal decisions.	Salmanowitz (2018)	92	50	Not stated	White (100%)	51% University students. Mainly liberal and highly educated.
Contact in VR: Testing avatar customisation and common ingroup identity cues on outgroup bias reduction.	Alvidrez and Peña (2020)	135	82.3	20.47 (2.05)	White (17%) Hispanic (20.6%) Asian (52.5%) Other (9.9%)	University students
The influence of racial embodiment on racial bias in immersive virtual environments.	Groom et al. (2009)	98	60.2	Not stated	White (46.9%) Asian (21.4%) African American (16.4%) Hispanic (7.1%) Other (8.2%)	University students
Exploring the Effect of Cooperation in Reducing Implicit Racial Bias and Its Relationship With Dispositional Empathy and Political Attitudes.	Patané et al. (2020)	40	67.5	25.8 (5.46)	White (100%)	Not stated

on socio-economic status were provided by Salmanowitz (2018), who stated that their participants were predominantly liberal and highly educated. Finally, the majority of the included studies consisted of either an all-White sample ($n=7$) or a majority-White sample ($n=3$). Five studies included Asian participants, and three studies incorporated Hispanic or self-identified Other participants.

3.3 IVR interaction and embodied avatar

All studies included an embodiment condition. It is important to note while the intervention aspect of our inclusion criteria necessitated participants to engage in a full embodiment IVR condition, Harjunen et al. (2021) incorporated partial embodiment, specifically focusing on the hands. Nonetheless, since the participants were situated behind a virtual table with only their hands visible, the implication is that the entire body is connected. Consequently, this study was not excluded from our search. The main forms of interaction included a combination of both avatar embodiment and avatar interaction ($n=6$), embodiment-only conditions ($n=3$), and alternative conditions ($n=3$). The alternative conditions included mental perspective-taking (imagining taking the perspective of a research confederate) versus embodied perspective-taking ("body swapping" with the research confederate; Thériault et al. 2021); a sham condition, whereby

Table 4 Details of the methods, measures, and results

Title	Authors	Intervention method	Comparater	Avatar race	Prejudice-related measures	Main Results
Body swapping with a Black person boosts empathy: Using virtual reality to embody another	Thériault et al. (2021)	MPT (Participants were provided with a picture and instructed to imagine taking the perspective the Black confederate) EPT ("body swapping" with Black confederate to give the illusion of body ownership)	Mental perspective taking	Sex-matched Black confederate	Implicit Association Test (IAT) Symbolic Racism Scale Interpersonal Reactivity Index Questionnaire of Cognitive Affective Empathy	EPT group showed more empathy than controls. The EPT and MPT groups showed more self-other merging than controls. EPT and MPT groups are comparable to the control group on explicit and implicit racial prejudice measures.
Increasing Self-Other Similarity Modulates Ethnic Bias in Sensorimotor Resonance to Others' Pain	Harjunen et al. (2021)	Treatment of virtual hand: noxious vs. soft. Avatar skin colour: white vs. black vs. violet. Participant embodiment: white vs. black body.	Embodiment of ingroup (White) avatar hands.	Embodied white and black avatar & interaction with Black, white and violet virtual hand	Self-report measures and sensorimotor beta event-related desynchronization (ERD) measured using electroencephalography (EEG)	Ethnic bias exists in vicarious pain perception. Increasing self-other similarity through virtual body transfer effectively reduces this bias. Empathy is influenced by the congruence of skin colour between the observer and the observed agent.
Virtual race transformation verses racial ingroup bias	Hasler et al. (2017)	Embodied Black or White: each participant repeated the study twice with their respective virtual body race. The skin colour of the other agent differs across each trial.	Embodiment of ingroup (White) avatar	White: embodiment & interaction. Black: embodiment & interaction.	Pre- and Post-Racial IAT (PreIAT: 2–30 days before 1st exposure. PostIAT: after 2nd exposure). Mimicry: number of mimicked behaviours.	When embodied in a Black virtual body, White participants treat Black avatars as their novel in-group and White avatars become their novel out-group. This reversed in-group bias effect was obtained regardless of participants' level of implicit racial bias.
Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias	Banakou et al. (2016)	Embodied Black or White: participants embodied either a Black or White avatar and had either 1, 2 or 3 IVR exposures. Interaction: participants underwent a virtual Tai Chi lesson with either a White or Asian teacher.	Embodying and interacting with ingroup (White) virtual avatar	Embodiment: Black or White avatar. Interaction: Asian vs. White avatar	Pre- and Post-Racial IAT (PreIAT: 1 week before first virtual exposure. PostIAT: 1 week after their last virtual exposure).	Embodying a Black avatar consistently reduces implicit racial bias, regardless of exposure frequency. Bias also diminishes with more IVR exposures, regardless of the embodiment condition (White or Black).
Putting yourself in the skin of a black avatar reduces implicit racial bias	Peck et al. (2013)	Embodiment (involving visual-motor synchrony): White vs. Black vs. Purple. Non-Embodiment (asynchronous movement): Black	Embodying ingroup (White) and alien (purple) virtual avatar	Embodiment: White, Black or Alien (purple).	Pre- and Post-Racial IAT (PreIAT: 3 days before the experiment). PostIAT: immediately after the experiment).	White participants embodying a Black avatar showed less bias towards Black people than those with White or purple avatars.
Investigating the Influence of Intergroup Contact in Virtual Reality on Empathy: An Exploratory Study Using AltspaceVR.	Tassinari et al. (2022a)	Embodiment of own race (white). Interaction with White or Black avatars (playing a ball tossing game)	Interaction with White avatars.	Embodied: White Interaction: White or Black. Confederate: Middle Eastern	Batson et al. (1987) affective empathy scale. Four item scale of co-presence.	VR intergroup contact did not directly affect empathy. After VR exposure, distress and empathetic interest decreased similarly in both research groups compared to their pre-test scores.

Table 4 (continued)

Title	Authors	Intervention method	Comparater	Avatar race	Prejudice-related measures	Main Results
The Effect of VR Avatar Embodiment on Improving Attitudes and Closeness Toward Immigrants.	Chen et al. (2021)	Double or single condition: embody both PRC and SG. Single condition: embody either PRC or SG. Interaction: virtual SG customer.	SG only condition	Embodiment: PRC avatars (minority out-group) and/or SG avatars	Adapted Interpersonal Reactivity Index (Davis 1980). Adapted Alwin (1997) feeling thermometer. Adapted Aron et al. (1992) perceived closeness test.	Facing discrimination as an immigrant avatar enhances attitudes and closeness to the group. Empathy mediated this effect on improved attitudes and closeness towards immigrants.
Virtual body ownership and its consequences for implicit racial bias are dependent on social context.	Banakou et al. (2020)	Embodiment of own race (white). Interaction: Negative, neutral or positive)	Embodiment of White Avatar	Embodiment: Black & White Interaction: not stated	Pre- and Post-IAT Pre- and Post-Attitudes towards Blacks scale	Body ownership hinges on affective social context. For White participants in a Black avatar, shifts in bias depend on situations; negativity enhances bias more than neutrality or positivity. Black avatar, neutral setting linked to lower IAT scores than White avatar, neutral settings.
The impact of virtual reality on implicit racial bias and mock legal decisions.	Salmanowitz (2018)	VR Type: embodied or sham (experienced the virtual world without any connection to a physical body) Interaction: Virtual Defendant	Sham condition	Embodiment: Black avatar Interaction: White or Black defendant	IAT Symbolic Racism Scale Decisions in mock crime scenario (including 5 day follow up that involved another mock crime scenario and questions on: severity, likelihood of reoffending, a severity of sentence)	Embodying a Black avatar reduced racial bias, making participants judge ambiguous legal cases more conservatively than those in a virtual world without avatar embodiment.
Contact in VR: Testing avatar customisation and common ingroup identity cues on outgroup bias reduction.	Alvidrez and Peña (2020)	Embodiment: avatar self-resembling or avatar not self-resembling Interaction: common ingroup identity cues or not	Self resembling virtual avatar	Embodiment: not specified Engagement: Hispanic avatar	Social distance scale The Identification with the VR Group scale	Self-resembling avatars interacting with outgroup members alongside a shared ingroup identity reduced outgroup bias. Greater engagement presence led to more social distance. Common ingroup identity cues did not impact outcomes.
The influence of racial embodiment on racial bias in immersive virtual environments.	Groom et al. (2009)	Participant race: White vs. non-white Model race: Black vs. White Model representation: imagined vs. embodied.	Perspective taking: imagining a day in the life of Black person	White and Black across each experimental group	IAT Interpersonal distance The Racial Argument Scale Racism Scale	EPT affected implicit racial bias, varying with model race, while imagined perspective-taking had no impact. After embodying a Black avatar, bias against Black people increased. Imagining White or Black models resulted in similar, moderate bias levels.

Table 4 (continued)

Title	Authors	Intervention method	Comparater	Avatar race	Prejudice-related measures	Main Results
Exploring the Effect of Cooperation in Reducing Implicit Racial Bias and Its Relationship With Dispositional Empathy and Political Attitudes.	Patané et al. (2020)	Embodiment: Black avatar Group: Cooperation (cooperate with Black confederate in building activity) or Neutral (undertake building activity alone in the presence of Black confederate)	Neutral group	Embodiment: Black Interaction: Black	Pre- and Post-IAT Pre- and Post-IOS Interpersonal Reactivity Index scale	Cooperation group showed less bias than neutral. Reduced distance linked to lower neutral group bias, not cooperation. Empathy traits notably explained initial bias variance: perspective-taking, empathic concern, and personal distress predicted IAT scores.

participants experienced the virtual world but without any connection to a virtual body (Salmanowitz 2018); and a perspective-taking condition, which involved participants looking at a photograph of a model and imagining themselves as the model (Groom et al. 2009).

Two major forms of social group embodiment emerged from the analysis. These included the ingroup perspective (i.e., participants embodying same-race avatars) and the outgroup perspective (i.e., participants embodying outgroup avatars or different-race avatars). Most studies employed a combination of both ingroup and outgroup embodiment ($n=9$), the rest of which included outgroup-only embodiment ($n=3$). In both types of designs, the ingroup perspective typically involved White participants embodying White/light-skinned avatars, and the outgroup perspective usually entailed White participants embodying Black/dark-skinned avatars. Specifically, of the seven studies that comprised White-only participants, five involved participants embodying either their own race (White avatars) and/or Black avatars (Banakou et al. 2016, 2020; Harjunen et al. 2021; Hasler et al. 2017; Peck et al. 2013). The remaining two studies entailed participants embodying Black avatars only (Patané et al. 2020; Salmanowitz 2018). In contrast, Chen et al. (2021) included Singaporean Chinese (SC) participants (ingroup) who embodied both SC avatars and the People's Republic of China (PRC) Chinese avatars.

Two studies included neutral conditions, namely, alien (purple) avatars, with one study using an alien embodiment condition (Peck et al. 2013) and the other using an alien interaction condition (Harjunen et al., 2021). In addition to the alien interaction condition, Harjunen et al. (2021) included a condition that enabled the participants to interact with a White and Black virtual hand. Similarly, this condition was included in three other studies (Hasler et al. 2017; Patané et al. 2020; Salmanowitz 2018) in which participants could interact with a Black and White virtual avatar. The study conducted by Patané et al. (2020) included a black interaction condition only. Finally, three studies involved interaction with avatars of different ethnic groups, including Hispanic (Alvidrez et al., 2020), Asian (Banakou et al. 2016), and Middle Eastern descent (Tassinari et al. 2022a).

3.4 Outcome measures

Commonly investigated outcomes were implicit ($n=8$) and explicit ($n=5$) attitudes towards a particular target group. Only three studies used both implicit and explicit measures. Following the narrative of IVR being the ultimate "empathy machine" (Barbot and Kaufman 2020), several studies assessed empathy, pain perception, mimicry, evaluation of mock legal cases, and self-other overlap ($n=7$, using at least one such measure). Neurophysiological measures such as skin conductance, heart rate, and electroencephalography (EEG) were applied rarely, with only one study using EEGs to assess empathetic resonance to ethnic outgroup pain as measured by sensorimotor beta event-related desynchronisation (ERD; Harjunen et al. 2021).

As indicated in the literature, the IAT remained the predominant gauge of racial (implicit) prejudice ($n=8$), followed by the Interpersonal Reactivity Index (IRI; Davis 1980), which stood as the second most frequently employed assessment ($n=3$). The IRI is a multidimensional measure of empathy that comprises four subscales, namely, perspective taking (PT), empathetic concern (EC), fantasy (FS), and personal distress (PD). Unlike the IAT and the IRI, there is a wide range of different measures used across the studies for both explicit bias and empathy. Specifically, measures of explicit bias include the Symbolic Racism Scale (Thériault et al. 2021; Henry & Sears, 2002; Sears and Henry 2005), Feeling Thermometer (Chen et al. 2021; Salmanowitz 2018), Attitudes Towards Blacks (ATB) scale (Banakou et al. 2020), and the Racial Argument Scale (Groom et al. 2009). Measures of empathy include the IRI (Chen et al. 2021; Patané et al. 2020; Thériault et al. 2021; Batson et al. 1987), the Affective Empathy Scale (Tassinari et al. 2022a), and the Questionnaire of Cognitive Affective Empathy (Thériault et al. 2021). Finally, most studies ($n=10$) included some form of embodiment questionnaire.

Table 5 Immersive virtual reality equipment and avatar animation software

Title	Authors	Head Mounted Display	Resolution	Refresh Rate	Field of View (FOV)	Tracking System	Virtual Environment and Avatar Animation software
Body swapping with a Black person boosts empathy: Using virtual reality to embody another	Thériault et al. (2021)	Oculus Rift (DK2)	960 × 1080 per eye	75 Hz	100 degrees	PlayStation 3 camera for participants to view the confederate's body	The Machine to Be Another system developed by *BeAnotherLab*
Increasing Self-Other Similarity Modulates Ethnic Bias in Sensorimotor Resonance to Others' Pain	Harjunen et al. (2021)	Oculus Rift	1080 × 1200 per eye	90 Hz	110 degrees	Hand tracking -Leap Motion tracking	Photo realistic hand developed by Leap Motion program, which was integrated in male and female Genesis 2 characters in Daz Studio. Unity3D and Blender software are used for treatment animations
Virtual race transformation verses racial ingroup bias	Hasler et al. (2017)	Oculus Rift (DK2)	960 × 1080 per eye	75 Hz	100 degrees	Full-body tracking suit and Xsens MVN system real-time motion capture	Unity 3D software
Virtual Embodiment of White People in a Black Virtual Body Leads to a Sustained Reduction in Their Implicit Racial Bias	Banakou et al. (2016)	NVIS nVisor SX111	1280 × 1024 per eye	60 Hz.	111 degrees	Optitrack full body motion capture suit and the Motive software	3D Studio Max 2010 and Autodesk Character Generator for virtual avatars. Unity 3D for virtual environment.
Putting yourself in the skin of a black avatar reduces implicit racial bias	Peck et al. (2013)	NVIS nVisor SX111	1280 × 1024 per eye	60 Hz	111 degrees	Optitrack full body motion capture suit and NaturalPoint Arena motion capture software	HALCA for avatar animation
Investigating the Influence of Intergroup Contact in Virtual Reality on Empathy: An Exploratory Study Using AltspaceVR.	Tassinari et al. (2022a)	Oculus Quest 2	1832 × 1920 per eye	120 Hz	127 degrees (96 HFOV x 90 VFOV)	Headset tracking and hand controllers	Altspace built-in tools and Unity3D software for the virtual environments. Altspace for virtual characters.
The Effect of VR Avatar Embodiment on Improving Attitudes and Closeness Toward Immigrants.	Chen et al. (2021)	HTC Vive	1080 × 1200 per eye	90 Hz	110 degrees	Headset tracking and hand controllers	Not Stated
Virtual body ownership and its consequences for implicit racial bias are dependent on social context.	Banakou et al. (2020)	HTC Vive	2160 × 1200 per eye	90 Hz	100 degrees	Whole body tracking using VIVE hand waist and feet trackers	Unity3D software for the virtual environment. Adobe Fuse and Mixamo for virtual avatars
The impact of virtual reality on implicit racial bias and mock legal decisions.	Salmanowitz (2018)	HTC Vive	1080 × 1200 per eye	90 Hz	110 degrees	Headset tracking and hand controllers	Crime scene photos provide by SPACES, Inc.
Contact in VR: Testing avatar customisation and common ingroup identity cues on outgroup bias reduction.	Alvidrez and Peña (2020)	Oculus Rift	1080 × 1200 per eye	90 Hz	110 degrees	Headset tracking	vTime used as the virtual reality platform

Table 5 (continued)

Title	Authors	Head Mounted Display	Resolution	Refresh Rate	Field of View (FOV)	Tracking System	Virtual Environment and Avatar Animation software
The influence of racial embodiment on racial bias in immersive virtual environments.	Groom et al. (2009)	nVisor SX	1280 × 1024 per eye	60 Hz	63 degrees (50 HFOV x 38 VFOV)	Full body tracking using an optical tracking system (WorldViz PPT)	Avatars were created from photographs of students (400 × 400 pixels). Confederate's avatar's eyes blinked and mouth moved.
Exploring the Effect of Cooperation in Reducing Implicit Racial Bias and Its Relationship With Dispositional Empathy and Political Attitudes.	Patané et al. (2020)	Oculus Rift CV1	1080 × 1200 per eye	90 HZ	110 degrees	Head and upper body tracking using Perception Neuron body motion capture suit from Noitom	3D Studio Max 2018 on the Unity 3D software. Adobe Fuse CC for virtual avatar.

Notes Hz = Hertz; DK2 = Development Kit 2; HFOV = Horizontal Field of View; VFOV = Vertical Field of View. The full field of view was calculated for Tassinari et al. 2022a); Groom et al. (2009) using the $\sqrt{FOV = HFOV2 + VFOV2}$ formula. Information regarding resolution, refresh rate and/or field of view that was not provided by Alvidrez and Peña (2020), Chen et al. (2021), Hasler et al. (2017) and Salmanowitz (2018) was obtained based on the specification of the head-mounted display

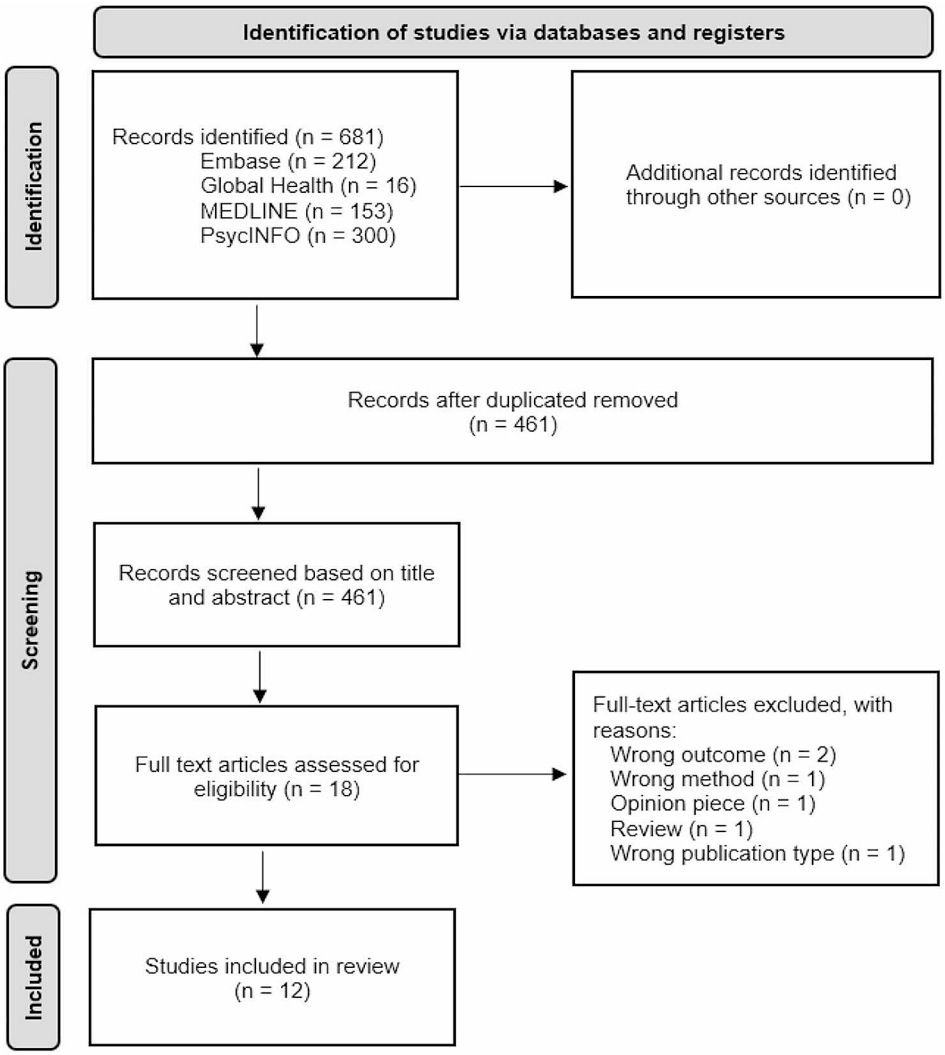

Fig. 1 PRISMA diagram depicting the flow of information through the four phases of the systematic review

3.5 IVR equipment and software

Differences in the IVR equipment and software were present (see Table 5). In terms of equipment, half of the studies used Oculus head-mounted displays, including the Rift Development Kit 2 (DK2; Hasler et al. 2017; Thériault et al. 2021), Rift (Alvidrez and Peña 2020; Harjunen et al. 2021) Rift CV1 (Patané et al. 2020) and the Quest 2 (Tassinari et al. 2022a). The HTC Vive was the second most used head-mounted display ($n=3$; Banakou et al. 2020; Chen et al. 2021; Salmanowitz 2018) and additional displays included the Navigation and Visualization Systems (NVIS), such as the NVIS nVisor SX111 (Banakou et al. 2016; Peck et al. 2013) and the nVisor SX (Groom et al. 2009). Across the included studies, the resolution ranged between 960 × 1080 (Thériault et al. 2021; Hasler et al. 2017) and 2160 × 1200 (Banakou et al. 2020) pixels per eye, the refresh rate ranged from 60 Hz (Banakou et al. 2016; Groom et al. 2009; Peck et al. 2013) to 120 Hz (Tassinari et al. 2022a) and the field of view was between 63 degrees and 127 degrees, thereby further demonstrating disparities in IVR systems. All studies used head-tracking systems fixed to the head-mounted display, with three studies using headset-only tracking (Alvidrez and Peña 2020; Harjunen et al. 2021; Thériault et al. 2021). Additional systems included full-body (Banakou et al. 2016, 2020; Groom et al. 2009; Hasler et al. 2017; Peck et al. 2013), head and upper body (Patané et al. 2020) and a combination of headset and hand controller tracking systems (Chen et al. 2021; Salmanowitz 2018; Tassinari et al. 2022a). Finally, while there are some commonalities in the IVR development platforms, namely Unity3D ($n=6$), 3D Studio Max ($n=2$), and Adobe Fuse ($n=2$), there are significant differences not only in the combination of software tools but also in the scenarios in which these tools are used to develop.

3.6 Research outcomes

Of the studies that measured implicit biases, five out of the eight studies reported a decrease in implicit bias scores (Banakou et al. 2016, 2020; Patané et al. 2020; Peck et al. 2013; Salmanowitz 2018). While these studies have demonstrated IVR's potential in the reduction of implicit racial bias, there is some conflicting evidence of the effect of embodiment on intergroup attitudes. In contrast to the findings of Hasler et al. (2017) and Thériault et al. (2021) - who reported no statistically significant differences in IAT scores across condition groups, suggesting a lack of measurable changes in implicit racial bias resulting from embodiment within their studied parameters - Groom et al. (2009) observed higher IAT scores among participants embodying Black avatars compared to those embodying white avatars. Similarly, Alvidrez and Peña (2020) demonstrated that participants categorised in the self-resembling avatars (ingroup) condition reported less perceived outgroup bias compared to participants who were required to customise avatars who looked physically different from themselves (outgroup). Three studies assessed explicit bias together with implicit bias (Banakou et al. 2020; Salmanowitz 2018; Thériault et al. 2021), two of which failed to obtain converging results, suggesting that these two types of measures are often discordant. While Thériault et al. (2021) found converging results for implicit and explicit bias, it is important to note that they coincided in that there were nonsignificant results for implicit and explicit biases. Thus, across all studies, no effect of IVR embodiment on explicit bias was found.

With regard to empathy, six studies demonstrated an increase in empathy or related components as a result of IVR embodiment. In particular, participants in an embodied perspective-taking condition showed an increase in empathy in comparison to the control group (Thériault et al. 2021). Moreover, cross-racial resemblance in IVR has been shown to increase mimicry (Hasler et al. 2017), modulate sensorimotor resonance to others' perceived pain (Harjunen et al. 2021), and lead to more conservative evaluations of legal cases (Salmanowitz 2018). Empathy significantly explained the variance observed in IAT scores, with perspective-taking, empathic concern, and personal distress being significant predictors of implicit bias. Chen et al. (2021) showed that empathy functioned as a mediator of IVR contact when it came to embodying outgroup members and that participants who placed greater importance on their various group memberships demonstrated stronger intervention effects (i.e., an increase in self-other overlap with the embodied outgroup). These findings correspond with those establishing that intergroup contact reduces prejudice via both affective mediators, namely, empathy and intergroup anxiety, and cognitive mediators, including perspective-taking, knowledge, and increased familiarity (for an overview, see Pettigrew and Tropp 2008). Nevertheless, IVR contact has also been shown to have no significant effect on empathy (Tassinari et al. 2022a).

3.7 Assessment of methodological quality

The Joanna Briggs Institute (JBI) critical appraisal checklist for randomised control trials (Tufanaru et al. 2017) was used to assess the methodological quality of the included studies. Two reviewers (SA & SH) evaluated the quality and, thus, the eligibility of each study. Any disagreements were resolved by a third reviewer (BD). A summary of this analysis is provided in the supplementary information (*Online Resource*1). The average quality score of the studies was 5.75, with a 95% confidence interval of 4.97 to 6.52,

and the maximum possible quality score was 10. Only five studies clearly stated if true randomisation was used for the randomisation of participants to treatment groups (Chen et al. 2021; Hasler et al. 2017; Peck et al. 2013; Tassinari et al. 2022a; Thériault et al. 2021). Similarly, it was unclear in most studies whether individuals delivering treatment were blind to treatment assignment ($n = 10$; Alvidrez et al., 2020; Banakou et al. 2016, 2020; Chen et al. 2021; Harjunen et al. 2021; Hasler et al. 2017; Patané et al. 2020; Peck et al. 2013; Salmanowitz 2018; Tassinari et al. 2022a) and whether the outcome assessors were blind to treatment assignment ($n = 9$; Alvidrez et al., 2020; Banakou et al. 2016, 2020; Chen et al. 2021; Hasler et al. 2017; Patané et al. 2020; Peck et al. 2013; Salmanowitz 2018; Tassinari et al. 2022a). Only one study (Salmanowitz 2018) made it clear that a follow-up was completed. However, all but one study (Thériault et al. 2021) scored highly for using the appropriate statistical tests. This decision was based on the numerous statistical tests used in the study, which the reviewers agreed would increase the likelihood of Type I errors. All papers measured treatment outcomes in the same way across treatment groups, and 10 articles (Alvidrez et al., 2020; Banakou et al. 2016, 2020; Groom et al. 2009; Harjunen et al. 2021; Hasler et al. 2017; Patané et al. 2020; Peck et al. 2013; Salmanowitz 2018; Thériault et al. 2021) clearly stated that treatment groups were treated identically. Taken together, insufficient clarity was provided in relation to whether the studies followed certain procedures that are characteristic of randomised control trials, including randomisation, the blinding of treatment assessors and outcome assessors. Similarly, these articles were vague in describing whether a participant follow-up procedure was conducted or completely failed to conduct this procedure altogether.

3.8 Meta-analyses of the IAT

The results of our first meta-analysis on dIAT scores conducted on four of the 12 studies (see Fig. 2) show that Cochran's Q statistic yielded a value of 7.43, suggesting a notable level of heterogeneity among the included studies. This indicates that the effect sizes across the included studies are not entirely consistent and could be influenced by factors beyond chance. The I^2 statistic, computed at 60%, underscores the extent of heterogeneity. This implies that approximately 60% of the observed variation in effect sizes could be attributed to real differences rather than random sampling errors. While this value indicates moderate heterogeneity, it further supports the notion that there are underlying factors contributing to the diversity of results. The p value ($p = .06$) does not quite reach statistical significance to reject homogeneity but still indicates that meaningful differences exist among the effect sizes. Testing for the main effect of condition, the meta-analysis revealed no significant overall decrease in IAT after embodiment of black vs. white avatars, $t(3) = -1.48$, $p = .24$. However, the high level of variance in effect sizes warrants a cautious interpretation of this result and prompts consideration of potential sources of heterogeneity within the studies.

Similarly, the results of our meta-analysis on Post IAT scores conducted on five of the 12 studies (as depicted in Fig. 3) show that a pronounced degree of heterogeneity exists within the analysed studies. The I^2 statistic, calculated at 90%, suggests that the observed variation in effect sizes could be attributed to genuine differences among the studies rather than mere random variability. The associated p value, which is less than 0.001, indicates a highly significant departure from homogeneity. Testing for the main effect of condition, the meta-analysis revealed no significant overall decrease in IAT after embodiment of black vs. white avatars, $t(4) = -1.05$, $p = .35$. Again, however, the substantial heterogeneity means that careful consideration of potential sources of variability among the studies becomes crucial to understanding the overall effect size and its implications. It is also important to recognise that our ability to draw robust conclusions is somewhat constrained by the relatively small number of studies available for analysis.

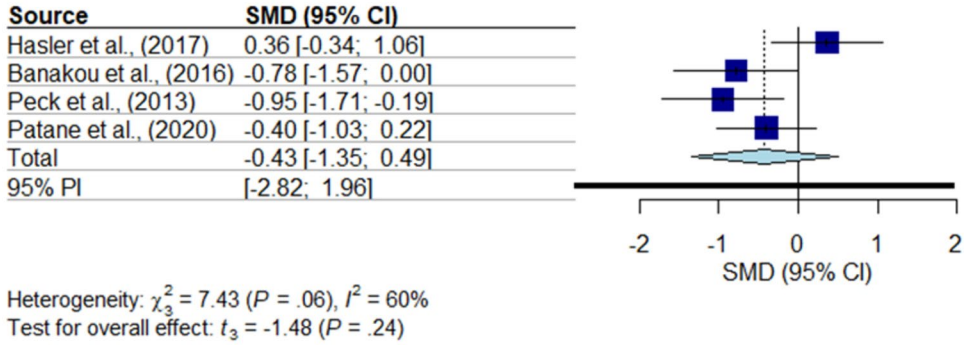

Fig. 2 Forest plot of the meta-analysis of the dIAT effects ($n = 4$) when comparing the experimental and control conditions

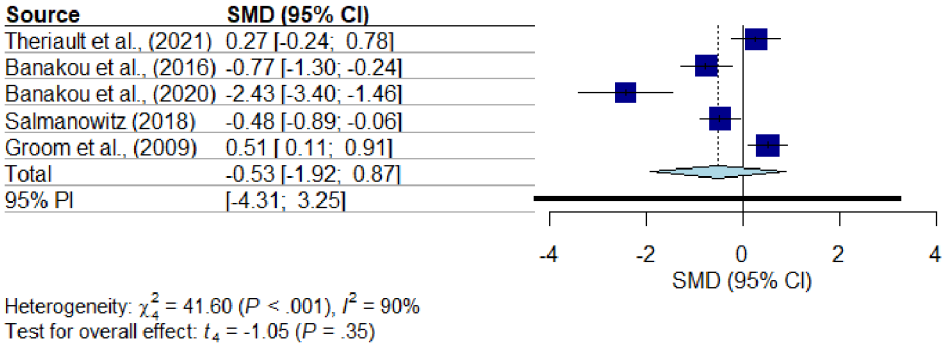

Fig. 3 Forest plot of the meta-analysis of the Post IAT effects ($n = 5$) when comparing the experimental and control conditions

4 Discussion

Social attitudes and evaluations play a crucial role in both real-life interactions and virtual environments. The use of IVR technology to induce body ownership illusions and investigate wider social identities has gained increased attention and significance in research. As users embody avatars, the representation of cognitive elements, beliefs, and evaluations becomes an integral part of their virtual identity. The success of avatar embodiment in VR studies is tied to the users' ability to project and perceive social attitudes, impacting how they are judged or how they judge others within the simulated environment (for an overview, see Poggi and D'Errico 2010). This systematic review aimed to comprehensively synthesise and examine the literature to elucidate the growing trend of IVR and its implications for research on racial prejudice. In particular, we sought to address four main research questions: (RQ1) What are the characteristics of the studies and the degree of representation of the sample and avatars used in IVR studies? (RQ2) What measures are used to assess implicit and explicit racial biases in IVR studies? (RQ3) To what extent do the results of the IAT differ across various studies focusing on IVR embodiment and racial implicit bias? (RQ4) Has embodiment using IVR been successful in eliciting a reduction in racial prejudice?

Regarding RQ1, demographic biases were demonstrated consistently in nearly all studies included in the review. All studies, except for one, were conducted in the global North and, thus, comprised participants from predominantly Western, Educated, Industrialised, Rich, and Democratic (WEIRD) settings and populations (Henrich et al. 2010). WEIRD populations often lack representation from a wide range of racial or ethnic backgrounds. As identified by Durrheim (2023), progressive calls such as 'WEIRD' exclude conversations around race, suggesting an extension of the term to *white and* Western populations (also see Besharati and Akinyemi 2023). As such, various measurement and sampling-related biases emerged in the review, underscoring the current limitations in external validity within the existing studies. Considering the complexity and diversity of human behaviour and attitudes within and between different racial groups, the results from the reviewed studies may portray an incomplete perspective. A lack of consideration of such cultural differences and social context may lead to a superficial understanding of the effects of the embodiment phenomena in IVR, reducing the depth of insights that this research is able to offer. Additionally, researchers originating from WEIRD locations may possess their own implicit assumptions regarding the cultural, socio-political, and racial norms of their own societies, which will inadvertently shape the research process and interpretation of results. Hence, we propose an increased emphasis on enhancing participant diversity. This can be achieved through collaboration with researchers originating from diverse geographic populations/locations, actively seeking participants beyond the confines of the hosting university or institution, and cultivating diverse leadership within the research teams overseeing these studies (see Peck et al. 2021 for an overview). Furthermore, researchers and reviewers should emphasize the importance of accurately quantifying the results of their studies in relation to the observed participant diversity (or lack thereof). This transparency could ensure a more nuanced interpretation of the findings and acknowledgment of the potential limitations in the generalizability of results.

Similar biases are observed in avatar representation. The existing literature has investigated either an ingroup perspective (i.e., embodiment of a same-race avatar) or an outgroup perspective (i.e., embodiment of a different race). In both types of designs, the ingroup perspective typically involved White participants embodying White avatars, while the outgroup perspective entailed White participants embodying Black avatars. Only one study included different ingroup-outgroup embodiment conditions other than Black and White ethnic groups. Within this study, Singaporean Chinese (SC) participants (ingroup) embodied both SC avatars and People's Republic of China (PRC) Chinese avatars (outgroup). Of the studies that recruited participants from different ethnic groups (i.e., Asian, Hispanic, or self-identified Other participants), most emulated

the above-mentioned patterns. That is, most still involved participants embodying either a combination of Black and White avatars (Groom et al. 2009; Tassinari et al. 2022a) or Black avatars only (Thériault et al. 2021). Hence, while these studies can be considered more inclusive by recruiting a more diverse sample, the embodiment condition is still limited to Black and White social groups. Moreover, even with the inclusion of a more varied participant pool, the predominant majority remains White in most cases, except for a single study in which the characteristics of the embodied avatar are unclear (Alvidrez and Peña 2020).

Our findings reveal that in addition to the overrepresentation of WEIRD participants, the majority of articles either exclusively or predominantly recruited White participants who embody either same-raced (White) avatars and/or Black avatars, the latter of which is consistently used as the representation of the social outgroup. While it has been argued that this choice is frequently influenced by demographic attributes — that is, White individuals being the dominant majority in the study region — the studies in the current review do not provide a theoretically driven justification or rationale for their choice of the study sample and avatar race. Additionally, race is often considered to be a sensitive or challenging area of empirical inquiry (Silverio et al. 2022), especially considering that the research area of focus is particularly controversial in that it involves one racial group embodying another. As a result, researchers might be hesitant to incorporate ethnically varied participants and avatars who could potentially embody marginalised social groups. This, in turn, could contribute to the observed patterns in the studies included. Nevertheless, the tendency to solely or predominantly draw on a White sample perpetuates and reifies the notion that this particular population sets the standard against which others are to be measured. IVR researchers can improve the inclusivity of their studies by integrating ethnically diverse avatar templates into their setups. An effective resource for this purpose is the open-source Virtual Avatar Library for Inclusion and Diversity (VALID) library, developed by Do et al. (2023). In summary, future research should be directed at diversifying not only the avatars embodied but also the research sample to better capture heterogeneity within diverse populations (Hatfield et al., 2022; Peck et al. 2021; Riches et al., 2023; Seaborn et al., 2023).

Although the included studies in this review had limited diversity in their sample and avatar embodiment, they demonstrated that embodiment could be induced for outgroup (e.g., Black or PRC group) avatars. In particular, ten out of the 12 studies included some form of embodiment questionnaire — assessing either immersion and feelings of presence in the virtual world or body ownership — thereby controlling for the success of the IVR experience. Nevertheless, there is a pressing need for further standardization of embodiment measures, such as the Embodiment Questionnaire developed by Peck and Gonzalez-Franco (2021) and the Alpha IVBO (Illusion of Virtual Body Ownership) scale developed by Roth et al. (2017). While a majority of studies rely on self-reported questionnaires to gauge the sense of embodiment in IVR, future researchers should consider a broader array of measures. This includes exploring biased sensory feedback, physiological indicators, locomotion patterns, and mental imagery tasks, as outlined in a comprehensive review by Guy et al. (2023). It is essential to recognize that an exclusive reliance on psychometric approaches for assessing embodiment may also overlook dimensions of depth in IVR. To enrich the understanding of embodiment phenomena, researchers should consider the integration of more qualitative methods, as suggested by Hassard (2023) and Lewis and Lloyd (2010). This holistic approach will contribute to a more comprehensive and nuanced exploration of the embodied experiences in immersive virtual reality settings.

In terms of the instruments used in the existing literature (RQ2), our results demonstrate that implicit bias, assessed via the IAT, and explicit bias, measured using a variety of instruments, are the most common constructs assessed. Additional constructs included empathy, typically assessed using the IRI, pain perception, mimicry, evaluation of mock legal cases, and self-other overlap. While the use of multiple traits and methods can provide a more comprehensive understanding, it may lead to inconsistencies in how reduced racial bias is defined and measured, which can limit the comparability and interpretation of research findings. In addition to varying constructs and instruments, differing scoring methods are occasionally implemented. For example, yet-to-be-verified standalone scores for empathy subscales (Wang et al. 2020) have been implemented (Patané et al. 2020; Thériault et al. 2021). The use of a variety of different tests and scoring methods brings into question the convergent validity of these measures, that is, the degree to which a test is related to other measures of the same construct and, consequently, the degree to which results are comparable across studies (Westen and Rosenthal 2003). Thus, throughout the examined studies, a wide range of prejudice-related measures, questionnaires, and tasks are administered without any apparent emerging standards in the field. Future research should, therefore, strive for consistency in conceptualisation and measurement of variables. A possible approach for IVR studies focusing on racial prejudice is to use the IAT, given its common use in the literature.

Regarding RQ3 and RQ4, there appears to be some consistency across studies using the IAT, as most studies indicated better performance on the IAT subsequent to the IVR condition. The consistency across studies that use different

samples and conditions possibly indicates that the IAT sufficiently assesses the intended construct (i.e., implicit racial bias), thereby supporting its use in future research. However, while the IAT may enable researchers to consistently study and predict individual behaviour over and above self-report measures, it has engendered some controversy in the literature. For example, the IAT faces challenges in reliability (Gawronski et al. 2017), implicit-criterion correlation and predictive ability (Greenwald et al. 2009; Kurdi et al. 2019; Oswald et al. 2013, 2015), ambiguous bias thresholds (Mitchell and Tetlock 2017), susceptibility to external influences and administration issues (Blanton et al. 2006; Greenwald et al. 2022; Ito et al. 2015). Therefore, future research should re-evaluate the appropriateness of the IAT and employ alternative methodologies, such as using improved scoring algorithms (Greenwald et al. 2009, 2015) and the application of pretest-posttest administration strategies to mitigate challenges associated with the IAT. Nevertheless, the results of our meta-analysis have shown that the IAT remains sensitive to experimental manipulations of embodiment in IVR settings.

Despite the apparent consistency, there are notable conflicting results, with studies demonstrating either no differences in IAT scores between condition groups (Hasler et al. 2017; Thériault et al. 2021) or an increase in IAT scores among participants embodying Black compared to White avatars (Groom et al. 2009). In addition to the variety of measures used and issues associated with the IAT, a possible alternative explanation is differences in equipment and software across studies. For example, head-mounted displays combined with full-body tracking deliver a more compelling rendering of the IVR modality compared to partial tracking (i.e., head and upper body; Patané et al. 2020) and the use of controllers, which enhances the sense of immersion by providing more accurate spatial awareness within the play area boundaries (Barbot and Kaufman 2020). Of the studies that used full-body tracking technology (Banakou et al. 2016, 2020; Groom et al. 2009; Hasler et al. 2017; Peck et al. 2013), only Hasler et al. (2017) yielded no effect. Although Groom et al. (2009) observe an increase in IAT scores, the study provides evidence to suggest that alterations in social attitudes occurred after the IVR condition. Therefore, full-body tracking technology may create a heightened sense of immersion and, therefore, potentially influence the effectiveness of IVR interventions.

Immersion is also possibly contingent on the quality of technology, which encompasses factors such as display resolution, refresh rate, and field of view (Barbot and Kaufman 2020). These features can provide more detailed and realistic visual representations, which contributes to the immersive experience (Sami Ur Rehman et al. 2023). Therefore, higher resolution and refresh rate, as well as an increased field of view, presumably enhance the immersive experience and, hence, the effectiveness of the IVR condition. Possible support for this idea becomes apparent when considering that the study with the highest display resolution (Banakou et al. 2020; 2160 × 1200 pixels per eye) showcased a reduction in implicit bias after the IVR condition, whereas the study with the lowest display resolution (Hasler et al. 2017; 960 × 1080 pixels per eye) yielded no effect. Nevertheless, the converse seems to occur for refresh rate and field of view, whereby no significant effect was observed subsequent to the IVR condition for the study with the highest refresh rate and field of view (Tassinari et al. 2022a; 120 Hz and 127 degrees, respectively) and a reduction of implicit bias subsequent to IVR expose was found in the studies with the lowest refresh rate (Banakou et al. 2016; Peck et al. 2013; 60 Hz) field of view (Groom et al. 2009; 63 degrees). It is important to highlight that the research of Tassinari and colleagues (2022) is notably distinguished from the other studies included in the review in terms of avatar visualisation. Specifically, the development platform (i.e., Altspace) comprises avatars that can be described as cartoon-like. This design style consists of non-realistic human proportions, is characterised by exaggerated body parts, such as larger or disproportionate physical features (i.e., head, eyes, nose and hands), and features non-typical human morphology (e.g., four fingers per hand; Weidner et al. 2023). More realistic rendering has been linked to improvement in embodiment and body ownership experiences, as well as presence and social presence (Weidner et al. 2023). Therefore, despite the high refresh rate, contributing to more natural and fluid movements of the virtual avatar and field of view, allowing a more natural perception of virtual environment, the limited realism of the avatar in Tassinari and colleagues' (2022) study may have contributed to the observed lack of effectiveness of the IVR intervention. Therefore, future studies should focus on using IVR technology and development software that may enhance a sense of realism by aligning the virtual experience with real-world expectations.

An alternative explanation for the lack of an effect in Hasler et al. (2017) study is that for the relevant element of embodied avatar (white/black) this study did use a between study design, i.e. all participants embodied either a white or black avatar. However, there was an additional factor of intergroup contact which was implemented via a within subject design (in one session participants performed a task with a white virtual character and in another with a black virtual character with IAT scores only being measured at the end of the second session. Since we know that intergroup contact can act to reduce prejudice (Pettigrew and Tropp 2008) it is possible that the effect of virtual embodiment on implicit bias was cancelled out by that of intergroup contact

which was experienced by all participants regardless of avatar body.

Taken together, in evaluating the success of embodiment through IVR technology in prompting a decrease in racial prejudice (research question four), our systematic review, coupled with the meta-analysis, highlighted the potential for IVR methods to be used to understand intergroup attitudes. Nevertheless, our findings also identify various factors that may impact research findings. In particular, inconsistencies in measures of prejudice may lead to instrument-specific findings as well as interpretation and comparability limitations across studies. Moreover, differences in IVR equipment specifications and virtual world development platforms may significantly influence the sense of realism experienced by participants, which could influence the effectiveness of IVR as an intervention to reduce racial bias. Our results also show that there is a large homogeneity of the population groups (sampled populations and embodied avatars), limiting research findings to WEIRD settings.

Finally, the current systematic review has notable strengths, namely, the inclusion of the quality appraisal procedure, the use of PICO and PRISMA guidelines, and the use of two independent reviewers together with a third reviewer to resolve conflict. These factors assisted in minimising errors and enhancing the power of the review. Nevertheless, notable limitations persist, including the scope of the literature search, which was only conducted in four major electronic databases (Embase, Global Health, MEDLINE and PsycINFO). Additionally, articles were only included if they were available in English. Therefore, the limited databases and the decision to solely incorporate English articles may have led to the oversight of additional relevant studies. The significance of language choice as a limitation of this review becomes particularly apparent, considering that a drawback we identified is the prevalence of studies conducted in WEIRD settings such as the USA, where English is the predominant language. Hence, the inclusion of articles contingent on whether they are in English possibly constitutes selection bias. Another potential limitation is the small number of articles included in the review. Nevertheless, the number of studies included in our review is the typical range used in most relevant, peer-reviewed systematic reviews, with the number of included studies ranging from 11 to 20 (Choi et al., 2023; Demeco et al. 2023; Neumann et al., 2018; Turbyne et al., 2021). Furthermore, while the relatively limited number of included articles may be indicative of stringent inclusion criteria, the number of studies included also reflects the existing state of the field in using IVR methods to reduce racial prejudice specifically, aligning with the primary aim of this review. Lastly, there is some heterogeneity across the studies in terms of control groups, interventions, and measures, which have the potential to affect study results (Bartolucci and Hillegass 2010). Nevertheless, heterogeneity may be attributed to the scope of the review, as it determines the extent to which the included articles are diverse.

4.1 Concluding comments

Recent research in the field has highlighted the potency of IVR as a tool for inducing and controlling embodiment while also effectively showing the potential to change attitudes and reduce racial bias, albeit temporarily. However, a cautionary note is also needed in this line of research, as not to undermine the complexities of racism and racist attitudes, as well as the post-colonial legacies involved in systemic prejudice. This review has highlighted that the sense of immersion and embodiment fostered by the IVR encounter renders it a powerful method to understand the embodied nature of perspective-taking and attitude change, but the methodology is still highly reductionist. IVR methods therefore have the potential to help examine the embodied mechanisms underlying multifaceted processes such as racial bias, but care is needed to understand these results within the wider socio-political and historical context in which they are embedded. Future studies drawing on more diverse sample groups and embodied avatars, while also using more standardised benchmarks for measures of implicit and explicit racial biases, will undoubtedly enrich and broaden the application of the use of IVR methods in racial prejudice research.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10055-024-01024-w.

Acknowledgements SB is a CIFAR Azrieli Global Scholar in the Brain, Mind and Consciousness Program and is funded by the Oppenheimer Memorial Trust and the National Research Foundation of South Africa (grant number: 150745). This work is based on the research supported by the National Institute for The Humanities and Social Sciences to SF.

Author contributions SH: Investigation, Data Curation, Methodology, Writing-Original Draft, Writing-Review & Editing, Visualization. SA: Formal Analysis, Writing-Original Draft, Writing-Review & Editing. BDA: Formal Analysis, Methodology, Writing-Original Draft, Writing-Review & Editing, Data Curation, Visualization. WD: Conceptualisation, Writing-Review & Editing. HF: Conceptualisation, writing-Review & Editing, Formal Analysis. SB: Conceptualisation, Funding, Formal Analysis, Writing-Original Draft, Writing-Review & Editing, Project Administration, Supervision.

Funding Open access funding provided by University of the Witwatersrand.

Declarations

Competing interests Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

5. References

Alvidrez S, Peña J (2020) Contact in VR: Testing Avatar Customisation and Common Ingroup Identity cues on Outgroup Bias reduction. Annual Rev Cybertherapy Telemedicine 18:89–93

Anthes C, García-Hernández RJ, Wiedemann M, Kranzlmüller D (2016), March State of the art of virtual reality technology. In 2016 IEEE aerospace conference (pp. 1–19). IEEE

Banakou D, Hanumanthu PD, Slater M (2016) Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. Front Hum Neurosci 10. https://doi.org/10.3389/fnhum.2016.00601

Banakou D, Beacco A, Neyret S, Blasco-Oliver M, Seinfeld S, Slater M (2020) Virtual body ownership and its consequences for implicit racial bias are dependent on social context. Royal Soc Open Sci 7(12):201848. https://doi.org/10.1098/rsos.201848

Bar-Haim Y, Ziv T, Lamy D, Hodes RM (2006) Nature and nurture in own-race face processing. Psychol Sci 17(2):159–163. https://doi.org/10.1111/j.1467-9280.2006.01679.x

Barbot B, Kaufman JC (2020) What makes immersive virtual reality the ultimate empathy machine? Discerning the underlying mechanisms of change. Comput Hum Behav 111:106431. https://doi.org/10.1016/j.chb.2020.106431

Barden J, Maddux WW, Petty RE, Brewer MB (2004) Contextual moderation of racial bias: the impact of social roles on controlled and automatically activated attitudes. J Personal Soc Psychol 87(1):5–22. https://doi.org/10.1037/0022-3514.87.1.5

Bartolucci AA, Hillegass WB (2010) Overview, strengths, and limitations of systematic reviews and meta-analyses. Evidence-Based Practice: Toward Optimizing Clin Outcomes, 17–33

Batson CD, Fultz J, Schoenrade PA (1987) Distress and empathy: two qualitatively distinct vicarious emotions with different motivational consequences. J Pers 55(1):19–39. https://doi.org/10.1111/j.1467-6494.1987.tb00426.x

Behm-Morawitz E, Pennell H, Speno AG (2016) The effects of virtual racial embodiment in a gaming app on reducing prejudice. Communication Monogr 83(3):396–418. https://doi.org/10.1080/03637751.2015.1128556

Besharati S, Akinyemi R (2023) Accelerating African neuroscience to provide an equitable framework using perspectives from West and Southern Africa. Nat Commun 14(1):8107. https://www.nature.com/articles/s41467-023-43943-3

Blanke O (2012) Multisensory brain mechanisms of bodily self-consciousness. Nat Rev Neurosci 13(8):556–571. https://doi.org/10.1038/nrn3292

Blanke O, Slater M, Serino A (2015) Behavioral, neural, and computational principles of bodily self-consciousness. Neuron 88(1):145–166. https://doi.org/10.1016/j.neuron.2015.09.029

Blanton H, Jaccard J, Gonzales PM, Christie C (2006) Decoding the implicit association test: implications for criterion prediction. J Exp Soc Psychol 42(2):192–212

Botvinick M, Cohen J (1998) Rubber hands 'feel' touch that eyes see. Nature 391(6669):756–756. https://doi.org/10.1038/35784

Chen VH, Ibasco GC, Leow VJ, Lew JY (2021) The effect of VR avatar embodiment on improving attitudes and closeness toward immigrants. Front Psychol 12. https://doi.org/10.3389/fpsyg.2021.705574

Chowdhury TI, Shahnewaz Ferdous SM, Quarles J (2019) VR Disability Simulation Reduces Implicit Bias Towards Persons With Disabilities. IEEE Transactions on Visualization and Computer Graphics. PP. 1–1. https://doi.org/10.1109/TVCG.2019.2958332

Cosmides L, Tooby J, Kurzban R (2003) Perceptions of race. Trends Cogn Sci 7(4):173–179. https://doi.org/10.1016/s1364-6613(03)00057-3

D'Errico F, Leone G, Schmid M, D'Anna C (2020) Prosocial virtual reality, empathy, and EEG measures: a pilot study aimed at monitoring emotional processes in intergroup helping behaviors. Appl Sci 10(4):1196. https://doi.org/10.3390/app10041196

Dasgupta N, Greenwald AG (2001) On the malleability of automatic attitudes: combating automatic prejudice with images of admired and disliked individuals. J Personal Soc Psychol 81(5):800–814. https://doi.org/10.1037/0022-3514.81.5.800

Davis MH (1980) A multidimensional approach to individual differences in empathy. JSAS Catalog Sel Documents Psychol 10:85

Demeco A, Zola L, Frizziero A, Martini C, Palumbo A, Foresti R, Buccino G, Costantino C (2023) Immersive virtual reality in Post-stroke Rehabilitation: a systematic review. Sensors 23(3):1712. https://doi.org/10.3390/s23031712

Do TD, Zelenty S, Gonzalez-Franco M, McMahan RP (2023) Valid: a perceptually validated virtual Avatar Library for inclusion and diversity. Front Virtual Real. 4https://doi.org/10.3389/frvir.2023.1248915

Dotsch R, Wigboldus DHJ (2008) Virtual prejudice. J Exp Soc Psychol 44(4):1194–1198. https://doi.org/10.1016/j.jesp.2008.03.003

Dovidio JF, Gaertner SL (2010) Intergroup bias. In: Fiske ST, Gilbert DT, Lindzey G (eds) Handbook of social psychology. John Wiley & Sons, Inc., pp 1084–1121. https://doi.org/10.1002/9780470561119.socpsy002029

Dovidio JF, ten Vergert M, Stewart TL, Gaertner SL, Johnson JD, Esses VM, Riek BM, Pearson AR (2004) Perspective and prejudice: antecedents and mediating mechanisms. Pers Soc Psychol Bull 30(12):1537–1549. https://doi.org/10.1177/0146167204271177

Durrheim K (2023) Conversational silencing of racism in psychological science: toward decolonization in practice. Perspect Psychol Sci. https://doi.org/10.1177/17456916231182922

Ehrsson HH (2012) The concept of body ownership and its relation to multisensory integration. New Handb Multisensory Process 775–792. https://doi.org/10.7551/mitpress/8466.003.0067

Farmer H, Maister L (2017) Putting ourselves in another's skin: using the plasticity of self-perception to enhance empathy and decrease prejudice. Soc Justice Res 30(4):323–354. https://doi.org/10.1007/s11211-017-0294-1

Farmer H, Tajadura-Jimenez A, Tsakiris M (2012) Beyond the colour of my skin: how skin colour affects the sense of body-ownership. Conscious Cogn 21(3):1242–1256

FitzGerald C, Hurst S (2017) Implicit bias in healthcare professionals: a systematic review. BMC Med Ethics 18(1). https://doi.org/10.1186/s12910-017-0179-8

Forscher PS, Lai CK, Axt JR, Ebersole CR, Herman M, Devine PG, Nosek BA (2019) A meta-analysis of procedures to change implicit measures. J Personal Soc Psychol 117(3):522–559. https://doi.org/10.1037/pspa0000160

Galinsky AD, Moskowitz GB (2000) Perspective-taking: decreasing stereotype expression, stereotype accessibility, and in-group

favoritism. J Personal Soc Psychol 78(4):708–724. https://doi.org/10.1037/0022-3514.78.4.708
Gallagher S (2000) Philosophical concepts of the self: implications for cognitive sciences. Trends Cogn Sci 4:14–21
Gawronski B, Morrison M, Phills CE, Galdi S (2017) Temporal stability of implicit and explicit measures: a longitudinal analysis. Pers Soc Psychol Bull 43(3):300–312
Greenwald AG, McGhee DE, Schwartz JL (1998) Measuring individual differences in implicit cognition: the Implicit Association Test. J Personal Soc Psychol 74(6):1464–1480. https://doi.org/10.1037/0022-3514.74.6.1464
Greenwald AG, Poehlman TA, Uhlmann EL, Banaji MR (2009) Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. J Personal Soc Psychol 97(1):17
Greenwald AG, Banaji MR, Nosek BA (2015) Statistically small effects of the implicit association test can have societally large effects. J Personal Soc Psychol 108(4):553–561. https://doi.org/10.1037/pspa0000016
Greenwald AG, Brendl M, Cai H, Cvencek D, Dovidio JF, Friese M, Hahn A, Hehman E, Hofmann W, Hughes S, Hussey I, Jordan C, Kirby TA, Lai CK, Lang JWB, Lindgren KP, Maison D, Ostafin BD, Rae JR, Wiers RW (2022) Best research practices for using the Implicit Association Test. Behav Res Methods 54(3):1161–1180. https://doi.org/10.3758/s13428-021-01624-3
Groom V, Bailenson JN, Nass C (2009) The influence of racial embodiment on racial bias in immersive virtual environments. Social Influence 4(3):231–248. https://doi.org/10.1080/15534510802643750
Guy M, Normand J-M, Jeunet-Kelway C, Moreau G (2023) The sense of embodiment in virtual reality and its assessment methods. Front Virtual Real. 4https://doi.org/10.3389/frvir.2023.1141683
Harjunen VJ, Sjö P, Ahmed I, Saarinen A, Farmer H, Salminen M, Järvelä S, Ruonala A, Jacucci G, Ravaja N (2021) Increasing self–other similarity modulates ethnic bias in sensorimotor resonance to others' pain. Soc Cognit Affect Neurosci 17(7):673–682. https://doi.org/10.1093/scan/nsab113
Hasler BS, Spanlang B, Slater M (2017) Virtual race transformation reverses racial in-group bias. PLoS ONE 12(4). https://doi.org/10.1371/journal.pone.0174965
Hassard AS The Sense of Embodiment in Virtual Reality Explored Through Presence, Narrative Immersion and, Effort P (2023) [https://qmro.qmul.ac.uk/xmlui/handle/123456789/84265]
Henrich J, Heine SJ, Norenzayan A (2010) Most people are not WEIRD. Nature 466:29–29
Hodgson P, Lee VWY, Chan JCS, Fong A, Tang CSY, Chan L, Wong C (2019) Immersive virtual reality (IVR) in higher education: development and implementation. In: tom MC, Dieck, Jung T (eds) Augmented reality and virtual reality: the power of AR and VR for Business. Springer International Publishing, pp 161–173. https://doi.org/10.1007/978-3-030-06246-0_12
Ito TA, Friedman NP, Bartholow BD, Correll J, Loersch C, Altamirano LJ, Miyake A (2015) Toward a comprehensive understanding of executive cognitive function in implicit racial bias. J Personal Soc Psychol 108(2):187
Jacob-Dazarola R, Ortíz Nicolás JC, Cárdenas Bayona L (2016) Behavioral measures of emotion. Emot Meas 101–124. https://doi.org/10.1016/b978-0-08-100508-8.00005-9
Jost JT, Banaji MR, Nosek BA (2004) A decade of System Justification Theory: accumulated evidence of conscious and unconscious bolstering of the Status Quo. Political Psychol 25(6):881–919. https://doi.org/10.1111/j.1467-9221.2004.00402.x
Kurdi B, Seitchik AE, Axt JR, Carroll TJ, Karapetyan A, Kaushik N, Tomezsko D, Greenwald AG, Banaji MR (2019) Relationship between the Implicit Association Test and intergroup behavior: a meta-analysis. Am Psychol 74(5):569
Lewis E, Lloyd DM (2010) Embodied experience: a first-person investigation of the Rubber Hand Illusion. Phenomenology Cogn Sci 9(3):317–339. https://doi.org/10.1007/s11097-010-9154-2
Lowery BS, Hardin CD, Sinclair S (2001) Social influence effects on automatic racial prejudice. J Personal Soc Psychol 81(5):842–855. https://doi.org/10.1037/0022-3514.81.5.842
Maister L, Slater M, Sanchez-Vives MV, Tsakiris M (2015) Changing bodies changes minds: owning another body affects social cognition. Trends Cogn Sci 19(1):6–12. https://doi.org/10.1016/j.tics.2014.11.001
Mitchell G, Tetlock PE (2017) Popularity as a poor proxy for utility: the case of implicit prejudice. Psychol Sci under Scrutiny: Recent Challenges Proposed Solutions, 164–195
Nock MK, Park JM, Finn CT, Deliberto TL, Dour HJ, Banaji MR (2010) Measuring the suicidal mind. Psychol Sci 21(4):511–517. https://doi.org/10.1177/0956797610364762
Oswald FL, Mitchell G, Blanton H, Jaccard J, Tetlock PE (2013) Predicting ethnic and racial discrimination: a meta-analysis of IAT criterion studies. J Personal Soc Psychol 105(2):171
Oswald FL, Mitchell G, Blanton H, Jaccard J, Tetlock PE (2015) Using the IAT to predict ethnic and racial discrimination: small effect sizes of unknown societal significance. SSRN Electron J. https://doi.org/10.2139/ssrn.2564290
Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A (2016) Rayyan—a web and mobile app for systematic reviews. Syst Reviews 5:1–10. https://doi.org/10.1186/s13643-016-0384-4
Patané I, Lelgouarch A, Banakou D, Verdelet G, Desoche C, Koun E, Salemme R, Slater M, Farnè A (2020) Exploring the effect of cooperation in reducing implicit racial bias and its relationship with dispositional empathy and political attitudes. Front Psychol 11. https://doi.org/10.3389/fpsyg.2020.510787
Peck TC, Gonzalez-Franco M (2021) Avatar embodiment. A standardized questionnaire. *Frontiers in Virtual Reality*, *1*. https://doi.org/10.3389/frvir.2020.575943
Peck TC, Seinfeld S, Aglioti SM, Slater M (2013) Putting yourself in the skin of a black avatar reduces implicit racial bias. Conscious Cogn 22(3):779–787. https://doi.org/10.1016/j.concog.2013.04.016
Peck TC, McMullen KA, Quarles J (2021) DiVRsify: Break the cycle and develop VR for everyone. IEEE Comput Graph Appl 41(6):133–142. https://doi.org/10.1109/mcg.2021.3113455
Pettigrew TF, Tropp LR (2008) How does intergroup contact reduce prejudice? meta-analytic tests of three mediators. Eur J Social Psychol 38(6):922–934. https://doi.org/10.1002/ejsp.504
Poggi I, D'Errico F (2010) Cognitive modelling of Human Social Signals. Proceedings of the 2nd International Workshop on Social Signal Processing. https://doi.org/10.1145/1878116.1878124
Richardson WS, Wilson MC, Nishikawa J, Hayward RS (1995) The well-built clinical question: a key to evidence-based decisions. ACP J Club 123(3):A12–A13
Roberts SO, Bareket-Shavit C, Dollins FA, Goldie PD, Mortenson E (2020) Racial inequality in psychological research: Trends of the past and recommendations for the future. Perspect Psychol Sci 15(6):1295–1309. https://doi.org/10.1177/1745691620927709
Rossen B, Johnsen K, Deladisma A, Lind S, Lok B (n.d.) (eds) Virtual humans elicit skin-tone bias consistent with real-world skin-tone biases. *Intelligent Virtual Agents*, 237–244. https://doi.org/10.1007/978-3-540-85483-8_24
Roth D, Lugrin J-L, Latoschik ME, Huber S (2017) Alpha Ivbo - construction of a scale to measure the illusion of virtual body ownership. Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. https://doi.org/10.1145/3027063.3053272
Salmanowitz N (2018) The impact of virtual reality on implicit racial bias and mock legal decisions. J Law Biosci 5(1):174–203. https://doi.org/10.1093/jlb/lsy005

Sami Ur Rehman M, Abouelkhier N, Shafiq MT (2023) Exploring the effectiveness of immersive virtual reality for Project Scheduling in Construction Education. Buildings 13(5):1123. https://doi.org/10.3390/buildings13051123

Schimmack U, Howard A (2021), July 29 The race Implicit Association test is biased. Replicability Index. Retrieved from https://replicationindex.com/2021/07/28/the-race-implicit-association-test-is-biased/

Schnabel K, Asendorpf JB, Greenwald AG (2008) Assessment of individual differences in implicit cognition. Eur J Psychol Assess 24(4):210–217. https://doi.org/10.1027/1015-5759.24.4.210

Schwarzer, G., Carpenter, J. R., & Rücker, G. (2015). Meta-analysis with R (Vol. 4784). Cham: springer.

Sears DO, Henry PJ (2005) Over thirty years later: a contemporary look at symbolic racism. Adv Exp Soc Psychol 37:95–150. https://doi.org/10.1016/S0065-2601(05)37002-X

Silverio SA, Sheen KS, Bramante A, Knighting K, Koops TU, Montgomery E, November L, Soulsby LK, Stevenson JH, Watkins M, Easter A, Sandall J (2022) Sensitive, challenging, and difficult topics: experiences and practical considerations for qualitative researchers. Int J Qualitative Methods 21:16094069221124740. https://doi.org/10.1177/16094069221124739

Slater M (2017) Implicit learning through embodiment in immersive virtual reality. In: Liu D, Dede C, Huang R, Richards J (eds) Virtual, augmented, and mixed realities in Education. Smart Computing and Intelligence. Springer, Singapore. https://doi.org/10.1007/978-981-10-5490-7_2

Slater M, Sanchez-Vives MV (2016) Enhancing our lives with immersive virtual reality. Front Rob AI 3:74. https://doi.org/10.3389/frobt.2016.00074

Slater M, Spanlang B, Sanchez-Vives MV, Blanke O (2010) First person experience of body transfer in virtual reality. PLoS ONE 5(5):e10564. https://doi.org/10.1371/journal.pone.0010564

Tassinari M, Aulbach MB, Jasinskaja-Lahti I (2022) Investigating the influence of intergroup contact in virtual reality on empathy: an exploratory study using AltspaceVR. Front Psychol 12. https://doi.org/10.3389/fpsyg.2021.815497

Tassinari M, Aulbach MB, Jasinskaja-Lahti I (2022a) The use of virtual reality in studying prejudice and its reduction: a systematic review. PLoS ONE 17(7):e0270748

Thériault R, Olson JA, Krol SA, Raz A (2021) Body swapping with a black person boosts empathy: using virtual reality to embody another. Q J Experimental Psychol 74(12):2057–2074. https://doi.org/10.1177/17470218211024826

Tufanaru C, Munn Z, Aromataris E, Campbell J, Hopp L (2017) Systematic reviews of effectiveness. Joanna Briggs Institute reviewer's manual. The Joanna Briggs Institute, Adelaide, Australia, pp 3–10

Wang Y, Li Y, Xiao W, Fu Y, Jie J (2020) Investigation on the rationality of the extant ways of scoring the interpersonal reactivity index based on confirmatory factor analysis. Front Psychol 11:1086

Weidner F, Boettcher G, Arboleda SA, Diao C, Sinani L, Kunert C, Raake A (2023) A systematic review on the visualization of avatars and agents in AR & VR displayed using head-mounted displays. IEEE Trans Vis Comput Graph. https://doi.org/10.1109/TVCG.2023.3247072

Westen D, Rosenthal R (2003) Quantifying construct validity: two simple measures. J Personal Soc Psychol 84:608–618. https://doi.org/10.1037/0022-3514.84.3.608

Wilson CJ, Soranzo A (2015) The Use of Virtual Reality in Psychology: A Case Study in Visual Perception. *Computational and Mathematical Methods in Medicine*, *2015*, 151702. https://doi.org/10.1155/2015/151702

Yee N, Bailenson J (2007) The proteus effect: the effect of transformed self-representation on behavior. Hum Commun Res 33(3):271–290. https://doi.org/10.1111/j.1468-2958.2007.00299.x

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.