Valuing Diversity:

An Undervalued Outcome of and Potential Precursor to Intergroup Contact

Lukas Wallrich

Department of Psychology

Goldsmiths, University of London

Thesis submitted for the degree of Doctor of Philosophy

March 2021

Supervised by Dr Keon West and Prof Adam Rutland

Declaration of Authorship

I, Lukas Wallrich, hereby declare that this thesis and the work presented in it is entirely my own. Where I have consulted the work of others, this is always clearly stated.

Signed: _____

Date:

Acknowledgements

Firstly, I want to thank my supervisors, Dr Keon West and Prof Adam Rutland, who have provided the guidance, encouragement, freedom and feedback that made this endeavour possible. I also want to express my gratitude to other academics who provided valuable ideas and encouragement, in particular to Dr Michele Birtel, to Dr Shelley McKeown-Jones, to Dr Sally Palmer, and to anonymous reviewers in various journals. Such thanks also go to members of the Equalab and DIPlab research groups whose feedback helped to shape my research plans and thinking throughout.

I also need to express my gratitude to the *Studienstiftung des deutschen Volkes* (German Academic Scholarship Foundation). Their generous funding enabled me to pursue this project, while the opportunities for exchange with other scholarship holders provided valuable ideas for my research.

I am also grateful to the staff members of The Challenge Network who supported me in my desire to pursue applied research with real-world impact in spite of many conflicting pressures. In particular, I want to say thank you to Anthony Hawkins, Mila Evanics, Arnika Knight and Janine Fleming.

Most importantly, I am indebted to the many participants who took part in my studies or in the surveys that I relied on. Their willingness to respond to questionnaires that are often long and rarely exciting is what much of our scientific progress depends on. In that vein, I also want to thank Dr Rose Meleady, Dr Michelle Birtel, Dr Nicola Abbot and Dr Daniel Frings for their help with data collection.

On a personal level, I would not have been able to persist through the difficult times that appear to be part of most PhD journeys without the love and encouragement of my partner, Anna Hoffmann, whose critical questions hopefully also helped me to improve the communication of my ideas. Similarly, I want to thank my parents for the trust, love and support they have offered throughout.

Abstract

As diversity continues to increase across Western societies, prejudice persists. While public debate continues, meta-analytic evidence shows that identity-conscious approaches – those that recognise and appreciate differences – lead to reduced prejudice and better intergroup relations than identity-blind approaches. However, majority-status groups often prefer not to see colour or other differences. Therefore, the question of how a valuing of diversity can be developed matters. Intergroup contact might help, given that it is one of the best-established routes to prejudice reduction and an improvement of intergroup relations, yet to date, it has rarely been linked to diversity beliefs.

This thesis assesses the relationship between diversity beliefs and intergroup contact from various angles. It starts by asking about the primary direction of influence. So far, valuing diversity has been conceptualised as both a precursor to and outcome of intergroup contact, yet longitudinal or experimental research is very rare. Using longitudinal data, I show that positive and negative contact predict changes in the valuing of diversity over time, while the reverse paths are weaker and not statistically significant. From there, I test whether changes in valuing diversity can serve as a mediator that explains various effects of intergroup contact and show that such changes are particularly relevant when it comes to understanding the association between intergroup contact and cognitive outcomes. Regarding implications for practice, I show that a large-scale contact intervention increases the valuing of diversity, particularly when participants enter the intervention with high self-expansion orientation and engage in conversations about differences. However, a targeted intervention to promote the valuing of diversity through promoting conversations about the value of differences yielded mixed results, suggesting that the pathways to change are different for majority- and minority-status participants and that further intervention research is needed.

As highlighted through a single-paper meta-analysis at the end, the thesis provides consistent evidence for a model according to which positive and negative contact experiences shape the valuing of diversity, which then in turn shapes outgroup attitudes. I also present some evidence (both longitudinal and cross-sectional) that valuing diversity might increase intergroup approach intentions, and might thus result in increased intergroup contact over a longer timeframe. This suggests that conditions for the emergence of a virtuous cycle with self-reinforcing increases in positive contact and in valuing diversity might be created. With that, I highlight a novel pathway by which intergroup contact can contribute to the improvement of intergroup relations. Apart from advancing theory, this can inform the design of contact interventions.

Table of Contents

List of Tables

Appendices:

List of Figures

Appendices:

CHAPTER 1:

Intergroup contact and diversity beliefs

This dissertation is concerned with the challenges raised by increasing diversity in Western societies. It considers how intergroup contact and diversity ideologies can interact to contribute to social integration across groups. In this chapter, I consider the evidence to date and set out my research hypotheses in that context.

1.1. The need for social integration

Take Switzerland – a model democracy at the heart of Europe – and its famous controversy about black and white sheep in election campaigns (Church, 2008). Regarding the context, 25.1% of the Swiss population are foreigners, yet only about 2% of them are allowed to acquire citizenship each year (Federal Statistical Office, 2020) . While most Swiss support economic rights for foreigners, such as access to welfare, a majority opposes their participation in the political process (Federal Statistical Office, 2018). An election campaign poster depicting a white sheep kicking a black sheep over the border proved to be an election winner for the Swiss Populist Party for a decade, even though it also galvanised opposition (Quito, 2016). Similarly, visible minarets were banned by a popular vote in 2009, at a time when a vast majority of voters had never had the opportunity to see one with their own eyes, given that there were only four minarets in the whole country (Abdeleli, 2019). In a recent vote, this was followed by a ban on full-face coverings (e.g., burqas and niqabs, BBC News, 2021), shortly after a study estimated that no more than 30 Muslim women in the entire country regularly wore such coverings (Tunger-Zanetti, 2021). These decisions on the part of majority-status Swiss voters reflect common preferences with regard to diversity: generally, assimilation of foreigners, in which they adopt the mores of the majority population, is widely preferred over multiculturalism, where diversity is celebrated (Guimond et al., 2014). However, somewhat unsurprisingly, few minority-status group members

anywhere embrace assimilation, as that means giving up their own heritage (Teney, 2011; Verkuyten, 2005; Verkuyten & Martinovic, 2006). Similar dynamics take place elsewhere, be it in the form of the Muslim travel ban enacted by President Trump (N. Islam, 2018), the new citizenship law in India that disadvantages Muslims (Khan, 2020) or the "All Lives Matter" backlash to the "Black Lives Matter" movement (West et al., 2021). They provide the motivation for this research project.

Most Western societies are becoming increasingly diverse in terms of ethnicity, religion, wealth and (visible) sexuality, in a process that has been described as a "diversification of diversity" leading to a new state of "super-diversity" (Vertovec, 2007). While explicit expressions of racial prejudice may be declining (Dovidio & Gaertner, 2000), prejudicial attitudes are still widespread and have not changed in many countries since the 1990s (Hadler, 2012). Furthermore, members of minority-status groups experience discrimination in various domains, including the labour market (Blackaby et al., 2005) and the education system (Gillborn et al., 2017). In hiring, for instance, there is a significant gap in call-back rates between Black and White applicants, which, at least in the United States, has not narrowed in the past 30 years (Quillian et al., 2017).

This inequality takes place in a context of continuing segregation. An attraction to people who are similar (*homophily*) combined with various contextual factors leads to social networks that are similar in terms of employment status (Cappellari & Tatsiramos, 2011), social class, ethnicity and age (all in Social Integration Commission, 2014a). In the UK, for example, Britons have 40-50% fewer interactions with people from a different ethnicity than they would have if there was no social segregation, and this difference is strongest for adolescents (Social Integration Commission, 2014a). For example, 25% of English White adolescents attend schools that are at least 99% White and 50% of White English adolescents live in neighbourhoods where at least 95% of the

adolescents are White (compared to a national average of 87%, Burgess, Wilson, & Lupton, 2005). Combined with stark status differences between groups, this is one of the reasons why members of minority groups have weaker access to opportunities and a weaker voice in collective decision making.

It is important to note that the term social integration here is not used in the original sense, dating back to Durkheim (1951), that concerns dynamics within social groups. While this clearly is an important driver of mental and physical health (e.g., Seeman, 1996), integration across groups (which is what I shall refer to as social integration) has an additional importance. Putnam (1995) conceptualised this as bridging social capital and identified it as a requirement for equitable political participation. Later studies have identified it as a critical foundation for peace (Varshney, 2001), generalised social trust (Li et al., 2005; Uslaner, 2010), a tolerant society (Schmidt & Weick, 2017), and the liberal welfare state (Banting & Kymlicka, 2006). In addition, it has been associated with greater chances of finding productive employment (Bentolila et al., 2010). Conversely, the lack of social integration across groups has become widely recognized as imposing economic costs, reducing collective well-being and limiting a country's ability to resolve pressing challenges (Social Integration Commission, 2014b). Furthermore, it imposes individual costs. For instance, with regard to children and adolescents, the lack of cross-group friendships is associated with lower social competence (Eisenberg et al., 2009), increased relational victimization (Kawabata & Crick, 2011), lower resilience and higher vulnerability (Bagci et al., 2014; Graham et al., 2014), and lower academic performance (Bagci et al., 2017; Hallinan & Williams, 1990). Therefore, measures to increase intergroup connectivity and reduce social segregation deserve attention.

1.2. The effects of diversity

The increase of diversity has spurred much debate and research regarding its

consequences. On a societal level, increasing diversity has been suggested to lead to social isolation and a decline in trust, at least in the short term, which is most vividly expressed by Putnam in his description of the situation in the United States (2007):

> *"Inhabitants of diverse communities tend to withdraw from collective life, to distrust their neighbors, regardless of the color of their skin, to withdraw even from close friends, to expect the worst from their community and its leaders, to volunteer less, give less to charity and work on community projects less often, to register to vote less, to agitate for social reform more, but have less faith that they can actually make a difference, and to huddle unhappily in front of the television. [...]* **Diversity, at least in the short run, seems to bring out the turtle in all of us.**" *(emphasis added, pp. 150-151)*

In the UK, this resonates with a perspective widely held among White British people that "Britishness" is in decline due to an increase in diversity (Ethnos, 2006). Some research has replicated Putnam's findings that ethnically diverse neighbourhoods have lower levels of social trust, but various important caveats have been pointed out. For instance, Letki (2008) showed that the association becomes very small in the UK once neighbourhoods' socio-economic status is controlled for. Overall, it appears that the negative consequences of diversity come about through segregation and a lack of contact, not through diversity per se (Alesina & Zhuravskaya, 2011; Uslaner, 2010).

With regard to smaller groups, diversity has been hailed as bringing benefits in the form of higher creativity (McLeod et al., 1996; Stahl et al., 2010), which has been suggested to result in greater profitability of diverse companies (Herring, 2009). Conversely, diversity has been found to increase conflict within teams (Stahl et al., 2010) and be associated with more stereotyping (Leslie et al., 2008), thereby

undermining team performance. Taken together, these findings and debates have invigorated the debate surrounding diversity beliefs, i.e., answers to the questions of how diversity should be evaluated, and how it should be dealt with.

1.3. How to deal with diversity? Competing norms

In the face of increasing diversity, societies, organizations and individuals need to adopt a perspective as to how to best respond to diversity. On the organizational level, these have been labelled most frequently as diversity approaches, (Jansen et al., 2016; Stevens et al., 2008); on the personal level, they have been termed diversity ideologies (Rattan & Ambady, 2013). In both contexts, the key question is whether to value differences, or whether to downplay them. Leslie and colleagues (2020) use the analogy of salad bowls and melting pots, with the former representing identity-conscious and the latter identity-blind approaches.

Identity-conscious approaches are usually grouped together under the label of *multiculturalism,* while *colour-blindness* and *assimilation* are the most common identity-blind approaches.[1] These will be introduced and briefly discussed in turn before I present evidence regarding their respective impacts on intergroup relations.

1.3.1. Multiculturalism

Multiculturalism is based on the contention that differences between cultural groups should be acknowledged and recognized as something valuable, with the potential to enrich societies and groups (Stevens et al., 2008). In the UK, multiculturalism specifically emerged as a response to the influx of non-White immigrants from the 1950s onwards (Vertovec, 2007). In terms of its practical implementation, multiculturalism led to funding community organizations, focusing on

---

[1] Leslie and colleagues' meta-analysis (2020) includes meritocracy as a third identity-blind approach. However, in their discussion they highlight that meritocratic norms have often also been included in scales measuring multiculturalism, so that this seems to be a separate dimension rather than an alternative approach.

a positive portrayal of minority groups in the media, monitoring diversity and aspiring towards equality in employment, education, housing and representation and launching initiatives in education and other sectors that promote tolerance, respect and equality (Vertovec, 2007, 2010). In the late 20th century, the value of multiculturalism had become widely accepted and included into the set of politically correct beliefs, leading Nathan Glazer to begin his critique of multiculturalism with the premise that ''we are all multiculturalists now'' (Glazer, 1998). However, that consensus has become more contested since, which I will discuss below.

1.3.2. Assimilation

Assimilation is usually based on the fundamental idea that intergroup relations are best served if the minority groups 'behave like the dominant group' (Leslie et al., 2020, p. 457). However, some proponents suggest that this mischaracterises assimilation and that the ideology actually calls for "every cultural group to make its unique contribution to the final product [i.e. the American]." (Taylor, 2014, p. 2). Nevertheless, this egalitarian framing of assimilation appears implausible considering existing intergroup power hierarchies (c.f. Alba & Duyvendak, 2019). In any case, assimilation holds that the best way to deal with diversity is to reduce it, which stands in contrast with most people's desire to maintain their distinctive identities (Crisp et al., 2006).

1.3.3. Colour-blindness

The third set of diversity beliefs, colour-blindness, combines the assumption behind assimilation, according to which diversity comes in the way of interaction, with the notion that there should be no hierarchy of cultures or ethnicities. Therefore, the foundational belief underpinning a colour-blind approach to diversity is that the demands of equality and fairness are best served when group memberships and group differences are ignored in decision making and impression formation (Apfelbaum et al.,

2012; Stevens et al., 2008). This was attractive, at least initially, to many social psychologists who see categorization as the foundation of discrimination and outgroup derogation (Park & Judd, 2005; c.f. Tajfel, 1969). A demand for colour-blindness was brought to prominence by the US Civil Rights Movement (Rattan & Ambady, 2013), yet is now frequently seen as serving the interests of Whites concerned about anti-White bias (Norton & Sommers, 2011) and those attempting to ignore or fail to address existing racial discrimination (Awad et al., 2005). This view has found forceful expression in an opinion of U.S. Supreme Court Justice John Robert, who argued against affirmative action by stating that "[t]he way to stop discrimination on the basis of race is to stop discriminating on the basis of race" (*Parents Involved in Community Schools v. Seattle School Dist. No. 1*, 2007). However, that is unlikely to be the case; instead, colour-blindness can plausibly be held to entrench exclusion and inequality, as will be discussed in the following sections.

1.4. Impacts of diversity beliefs

Different sets of diversity beliefs are associated with different intergroup outcomes, as highlighted by a recent meta-analysis covering 167 samples and 296 effect sizes (Leslie et al., 2020). In their data, an endorsement of assimilation was consistently associated with negative intergroup outcomes, particularly with greater prejudice (ρ = .34 [.23, .46]), greater discrimination (ρ = .40 [.23, .58]), more negative stereotyping (ρ = .44 [.29, .59]) and less support for pro-diversity policies (ρ = -.38 [-.52, -.25]). Multiculturalism, conversely, was consistently associated with positive intergroup outcomes, particularly with less prejudice (ρ = -.32 [-.37, -.27]), less discrimination (ρ = -.22 [-.30, -.14]), less negative stereotyping (ρ = -.39 [-.52, -.27]) and greater support for pro-diversity policies (ρ = .57 [.41, .73]). The pattern for colour-blindness fell in between; embracing it was associated with somewhat lower prejudice (ρ = -.07 [-.15, -.003]), no significant difference in discrimination (ρ = -.08 [-.24, .08]), less negative

stereotyping (ρ = -.20 [-.31, -.10]), but lower policy support (ρ = -.25 [-.43, -.08]).

Regarding prejudice and negative stereotyping, multiculturalism was associated with

significantly stronger effects than colour-blindness ($p < .01$), while the associations with

discrimination did not differ ($p = .80$). An earlier and smaller meta-analysis (Whitley &

Webster, 2018) found similar associations and additionally showed a negative

relationship between multiculturalism and implicit prejudice.

1.4.1. Changes in interactions

Various studies have directly compared colour-blind and multicultural primes.

These were excluded by Leslie et al. (2020) as they lack a separate control, but they can

highlight the importance of the choice between these sets of beliefs. Holoien and

Shelton (2012), for instance, showed that when White participants were primed with

colour-blindness rather than multiculturalism, they exhibited greater behavioural

prejudice in a subsequent interaction with ethnic minority participants, who then

performed worse on a cognitive task. Conversely, cultural openness, defined as an

interest in other cultures, and thereby conceptually similar to an endorsement of multi-

culturalism, has been shown to predict bystander intervention intentions on the part of

British school children (Abbott & Cameron, 2014). Through longitudinal data,

intercultural competence, which includes paying attention to intergroup differences, has

been shown to lead to less negative and potentially also to more positive contact (the

latter effect only reached marginal significance, Meleady et al., 2020).

More generally, endorsement of colour-blindness has been found to induce a

focus on the self, i.e., on controlling cognition to ignore social categories, while

multicultural beliefs induce a focus on the other, i.e., on discovering and appreciating

differences, which then resulted in more positive interactions in the multiculturalism

condition (Vorauer et al., 2009). Confirming the same link in a real-world setting, a

survey of 3,758 employees of a healthcare organization revealed that White employees'

endorsement of colour-blindness predicted greater perception of racial bias in the organization by minority-ethnic employees', while an endorsement of multiculturalism on the part of White employees had the opposite effect (Plaut et al., 2009).

This is likely explained by the fact that we all see colour, so that striving for colour-blindness requires effort and introduces awkwardness. A variety of studies have shown that ethnicity is perceived quickly and automatically by both adults (Ito & Urland, 2003) and children as young as 3 months (Bar-Haim et al., 2006). Correspondingly, when tasked with sorting a set of pictures into two ethnic categories (Black and White), White American college students achieved 99.1% accuracy at high speed. However, compared to Black students, who achieved the same speed and accuracy, they underestimated their performance, likely driven by a motivation to appear unprejudiced (Norton et al., 2006, Study 1). This reluctance to acknowledge a perception of race can impede interracial cooperation, as shown in a study that tasked participants with identifying which of a set of face cards a confederate was holding, by asking as few yes-no questions as possible. Even though asking about skin colour was an obviously effective question, more than 1/3 of participants refrained from asking it when the confederate was Black. This avoidance was predicted by their endorsement of colour-blindness, and led to them being perceived as less friendly by their Black interaction partners (Norton et al., 2006, Study 2).

Conversely, a norm of valuing diversity can have beneficial effects on both majority- and minority-status participants, as a recent large-scale intervention study highlighted (Murrar et al., 2020). In that study, exposure to peers' pro-diversity attitudes led to a greater sense of inclusion and even to a narrowing of the ethnic achievement gap among minority-ethnic participants, while it led to more positive attitudes towards outgroups among majority-status participants. Consequently, minority-status participants exposed to the intervention reported being treated more inclusively by their

fellow-students.

1.4.2. Changes in policy support

The fact that colour-blindness was associated with decreased support for pro-diversity policies in Leslie et al. (2020), which often address inequities between groups, might be seen as surprising. Colour-blindness is frequently articulated as a principle of distributive justice, i.e., as requiring that socio-economic status should be independent of ethnicity. In that case, colour-blindness might be expected to predict support for redistribution that enhances equity between ethnic groups. However, when Whites are exposed to intergroup threat (i.e. to demands that threatens their privilege), colour-blindness is often transformed into a principle of procedural justice that prohibits the addressing of group-based inequalities in as far as this entails a recognition that these groups exist (Knowles et al., 2009). However, multiculturalism has also been criticised as insufficient when it comes to creating support for social change.

1.4.3. Critical challenges to multiculturalism

Once a concept comes to be embraced by everyone, across political divides, there will always be concerns that it loses its edge, i.e., its potential to lead to social transformation. With regard to multiculturalism, this has been vigorously articulated by Stuart Hall (2001):

> *"Over the years the term 'multiculturalism' has come to reference a diffuse, indeed maddeningly spongy and imprecise, discursive field: a train of false trails and misleading universals. Its references are a wild variety of political strategies. Thus conservative multiculturalism assimilates difference into the customs of the majority. Liberal multiculturalism subordinates difference to the claims of a universal citizenship. Pluralist multiculturalism corrals*

difference within a communally segmented social order. Commercial

multiculturalism exploits and consumes difference in the spectacle of

the exotic 'other.' Corporate multiculturalism manages difference in

the interests of the centre." (p. 3)

A similar concern arises from the work by Bell and Hartmann (2007). Trying to make sense of a survey finding according to which nearly half of Americans reported that diversity is 'mostly a strength for the country', they conducted in-depth interviews with 166 respondents. While most reported positive sentiments towards diversity, they found it difficult to talk about diversity and inequality – instead, most of the discourse consisted of what the authors termed "happy talk" (p. 895), which has been suggested to stand in the way of action to address inequality.

Furthermore, multiculturalism is compatible with high levels of neutral stereotyping; Leslie and colleagues' (2020) did not find a significant association, $\rho = .13$ [-.06, .32], but their results still highlighted the possibility that multiculturalism might lead to greater stereotyping in some domains, given that it implies that group membership has informational value. This goes along with the implicit focus on communities as "never-changing, socially bounded entities" (Vertovec, 2010, p. 85), which prevents a full recognition of intersectionality and is problematic because stereotypes of any valence can reduce identification with the outgroup and increase stress and vigilance during intercultural interactions (Hong et al., 2009).

1.4.3.1. Political concerns and contradictions

While psychological and sociological research continues to support multiculturalism, the powers that be at times disagree vehemently. In the early 2010s, two of Europe's most prominent political leaders pronounced the end of multiculturalism in brief succession. After Angela Merkel stated that in Germany, the

"approach [to build] a multicultural [society] and to live side-by-side and to enjoy each other has failed, utterly failed" (BBC News, 2010), David Cameroon concurred that in the UK "state multiculturalism has failed" as well (BBC News, 2011). This gave expression to a backlash against multiculturalism from the Right, who feared social breakdown in a time of heightened fears over domestic Islamist threats, but coincided with challenges from the Left who argued that multiculturalism serves to obscure class-based inequalities (Vertovec & Wessendorf, 2010).

As part of the concerns about multiculturalism, many countries have introduced (or bolstered) language requirements for immigrants, added tests on national history and cultural practices into the process for obtaining citizenship, restricted the wearing of religious symbols – particularly the hijab and niqab – in public settings, and otherwise increased demands placed on immigrants to embrace the national culture (Vertovec, 2010). While this usually does not directly affect established and naturalized ethnic minorities, it can still be argued to reflect and intensify a shift in norms regarding diversity and unity. Such proposals and policy shifts have been described as highlighting the shift from multiculturalism to assimilation (Vasta, 2007).

Apart from the ideological backlash, multiculturalism in its traditional manifestation had come under strain as it was based on the notion that society is made up of a limited number of clearly delineated and well-organized groups of immigrants or ethnic minorities. In the face of contemporary "super-diversity", where sub-groups have multiplied and boundaries have been blurred, alternative approaches had to be considered (Vertovec, 2007). One such idea is *interculturalism*, a new approach that emphasizes the dynamic nature of culture and cultural exchange alongside an emphasis on the valuing of cultural diversity. This might be able to better account for intersectionality and reduce stereotyping, even though the empirical evidence to date only shows marginal differences between multiculturalism and interculturalism

(Yogeeswaran et al., 2020).

Nevertheless, it appears that much of the multiculturalism debate has become about charged terminology rather than substance. For instance, in France, Macron proclaims that "it is important that the French whose parents or grandparents came to our country can preserve the knowledge of their culture and their language, and have the ability to contribute their culture to the Republic. […] This is valuable opportunity for all of us." (my translation; elysee.fr, 2020) while also stating that ('American') multiculturalism must be rejected (Tharoor, 2020). For the purpose of this dissertation, I will consider the fact that a recognition and appreciation of diversity brings benefits vis-à-vis other diversity beliefs as established, while generally steering away from its common but highly charged name.

1.5. Diverging preferences

A final challenge regarding diversity beliefs is that preferences diverge between ethnic minority and majority-status groups. White Americans endorse colour-blindness more than Black Americans, while the pattern for multiculturalism is the reverse (Ryan et al., 2007). Consequently, while members of ethnic minorities feel more comfortable and included in settings that highlight the value of diversity, members of the cultural majority report a preference for colour-blind approaches. This is both the case for national policy (Levin et al., 2012) and for organisations (Jansen et al., 2016; Plaut et al., 2011). In the latter case, however, colour-blind norms can resonate with minority-ethnic applicants *if* the organization espousing them is ethnically diverse (Purdie-Vaughns et al., 2008); coming from an ethnically homogeneous organization, such norms were associated with drastically reduced trust.

One reason for the diverging preferences might be that White people can feel left-out by multiculturalism. Typically, they are not expected to be part of the diversity

showcased in cultural celebrations and the like (Harris, 2013); thus some might feel that members of ethnic minorities get attention and/or benefits at their expense. In that case, advocating colour-blindness might be a rational strategy to eliminate perceived anti-White bias (Norton & Sommers, 2011).Therefore, an inclusive multiculturalism might be needed that explicitly values the diversity that Whites (and, among them, particularly heterosexual men) can bring to the table (Stevens et al., 2008).

Also, it should be noted that colour-blindness is not universally desired by most majority-status group members: many people wish to recognise diversity in cuisine and cultural expression yet revert to colour-blindness when confronted with a potentially charged situation. For instance, a survey of 27,000 EU citizens found that 72% of EU citizens believed that people with a different background (ethnic, religious, or national) enrich the cultural life of their country, with this opinion being held by a majority in all EU countries. 83% further expressed that they valued intercultural contacts (European Commission, 2007). However, when intergroup threat (i.e., threat to the majority-status group's privilege) arises, colour-blindness becomes more attractive (Knowles et al., 2009).

One reason that has been shown to underpin some Whites' preference for colour-blind approaches is the sense that not seeing race makes it easier not to be seen as racist. However, such strategic use of colour-blindness leads to increased negative non-verbal behaviour (Apfelbaum et al., 2008, Studies 1-2), and is actually perceived as an indicator of racism by Black observers, at least in situations when ethnicity is clearly relevant (Apfelbaum et al., 2008, Study 3; cf. Norton et al., 2006).

1.6. The heart of diversity beliefs: valuing diversity

So far, diversity beliefs have been discussed primarily as answers to the question whether diversity is to explicitly recognised and acknowledged. The evidence on that is

clear – doing so improves intergroup relations. Thus, the question arises how this can be promoted. Evidently, one way is to directly communicate the benefits to intergroup relation, which should be of use when it comes to information policymakers and other authorities. Beyond, however, it appears promising to look for additional reasons why individuals might value (and thus willingly recognise and acknowledge) diversity. Such reasons have been identified at an individual and at an intergroup level. Individually, engaging with diverse others might be seen as opening up the door to novel and enjoyable experiences (Miville et al., 1999), which is particularly attractive to individuals high in endeavour-oriented personality traits (Stürmer et al., 2013). Collectively, diversity might be seen as having instrumental benefits to groups, in terms of greater creativity and productivity (Kauff et al., 2020). Either set of beliefs – or a combination thereof – entail an explicit valuing of diversity, which is the key concern of this dissertation.

1.7. Promoting the valuing of diversity

From the evidence, it appears clear that the explicit valuing of diversity, often in the form of multiculturalism, is the diversity ideology most strongly associated with positive intergroup relations. However, promoting multiculturalism explicitly is not always successful; some studies have found that multiculturalism messages can trigger threat reactions and backfire. For instance, among White Americans who identify strongly with their ethnicity, multiculturalism primes led to a greater endorsement of social dominance, and to greater prejudice against ethnic outgroups (Morrison et al., 2010). Similarly, multiculturalism primes have been shown to lead to greater endorsement of conservative political beliefs in general, and of Donald Trump in particular (Osborn et al., 2020). Therefore, different and likely less explicit ways of promoting the valuing of diversity need to be developed.

Such interventions can potentially be related to the best-supported strategy

towards prejudice reduction: intergroup contact. In the next sections, I will introduce the evidence for the role of intergroup contact in promoting social integration, before proposing an integration of that research with research into diversity beliefs.

1.8. Intergroup contact as a pathway towards social integration

The notion that contact between members of distinct social groups can reduce prejudice and improve intergroup relations has been one of the most enduring and successful paradigms in social psychology (Pettigrew & Tropp, 2011). First articulated by Williams (1947), the *intergroup contact hypothesis* was popularized by Allport (1954) in *The Nature of Prejudice*. Since then, research across many populations has confirmed the effect of contact on prejudice, explored necessary and facilitating conditions for the effect, and identified mediators. More recent research has come to consider prosocial behaviour, and behavioural tendencies directly relevant with regard to social integration. These areas will be considered in turn.

1.8.1. Intergroup contact reduces prejudice

Over the six decades since its articulation, the intergroup contact hypothesis has inspired a broad range of research, covering a wide range of groups and methods. Research has considered the effect of contact between people of different ethnic groups (van Dyk, 1990), nationalities (Brown et al., 1999), religions (M. R. Islam & Hewstone, 1993), sexual orientations (Dessel et al., 2017), (dis)abilities (Maras & Brown, 2010) and ages (Harwood et al., 2005), including various methodologies such as surveys (Dessel et al., 2017), quasi-experiments (Laar et al., 2005), field studies (M. R. Islam & Hewstone, 1993; Mousa, 2018) and laboratory-based experiments (Brown et al., 1999). 515 of these studies, with a total of 250,089 participants from 38 countries, were included in the most systematic and comprehensive meta-analysis to date (Pettigrew & Tropp, 2006), which found strong evidence in support of the contact hypothesis, in both published and unpublished materials, across all target groups, geographies and

methodologies. Depending on data correction choices, their mean rs range from -.205 to -.214, with 94% of samples showing a negative relationship between contact and prejudice.

Recent meta-analyses provide further support for the efficacy of intergroup contact. An aggregation of 148 tests of intervention studies with more than 11,000 participants that were conducted outside the laboratory confirmed that contact can reduce ethnic prejudice in a wide range of settings (Lemmer & Wagner, 2015), with a similar mean effect size as in Pettigrew and Tropp (2006) and a smaller effect for high-quality studies such as randomized controlled trials ($\hat{\mu} = .28$, equivalent to $r = .14$). The most recent meta-analysis, which focused exclusively on studies with random assignment and delayed outcome measures found an average effect size that was in between the earlier ones ($d = .39$, which corresponds to $r = .19$), but found that this was reduced to $r = .12$ ($d = .25$) when only studies concerned with prejudice towards ethnic minorities and immigrants are considered (Paluck et al., 2019). The latter effect was found to be much higher in Pettigrew and Tropp, with $r = .21$ for racial and ethnic outgroups, but their dataset included very few experiments. Overall, it appears clear that intergroup contact has a consistent, though typically rather modest, effect on prejudice.

However, compared to other established predictors of prejudice, contact is powerful. Pettigrew and Tropp (2011) use data from Heitmeyer (2004) to show that positive contact is the third-strongest predictor of prejudice after social dominance orientation and a sense of collective threat, and thus more influential than education, group deprivation or authoritarianism. It also affects implicit prejudice (Tam et al., 2006), which has proven to be more stable and less susceptible to interventions than expressed prejudice (Greenwald et al., 2009).

1.8.2. Negative contact can do harm

Not all contact reduces prejudice; clearly negative contact experiences have the potential to make things worse. However, research into intergroup contact has often equated 'contact' with positive contact such as friendships (Barlow et al., 2012). Recent research has repeatedly included calls to move beyond this narrow focus. Initially, Barlow and colleagues (2012) suggested a *negative-positive contact asymmetry,* according to which negative contact was more influential than positive contact. Later work has shown that this asymmetry is not universal but might depend on the status of the specific outgroup (Paolini & McIntyre, 2019) and that there may be an interaction between positive and negative contact (Árnadóttir et al., 2018) in that positive contact can *buffer* against the effects of negative contact, while negative contact can *facilitate* the effect of positive contact. Nevertheless, too much negative contact can increase prejudice and run counter the aim of social integration.

Fortunately, positive contact is more frequent than negative contact in most intergroup settings, and typically features more prominently in participants' recollections when asked about contact in general (Graf et al., 2014; Hayward et al., 2017).

1.8.3. Intergroup contact promotes pro-social behaviour

Improved attitudes towards outgroups should arguably be seen as a proximate aim of contact; from the beginning, the aim of this research field was to find ways to actually improve intergroup relations (Dixon et al., 2005). This, arguably, again only serves as a step towards a broader goal of social justice and equality of opportunity. This distinction matters, because there might be trade-offs between improved affect and motivation to act to address intergroup inequities, and because it suggests that a broader range of outcomes of intergroup contact needs to be considered.

1.8.4. Prejudice reduction is not enough

Initially, it might seem self-evident that reduced prejudice changes behaviour; however, research has revealed that the link between attitudes and behaviours is often weaker than might be intuitively expected. The most comprehensive meta-analysis to date suggests that the link between prejudice and behaviour is rather weak, especially when considering cognitive aspects (beliefs: $r = .08$; stereotypes: $r = .26$). When considering affective prejudice in terms of emotions towards the outgroup, the relationship reaches medium strength: $r = .36$ (Talaska et al., 2008, cf. Schütz & Six, 1996). In further analyses, Talaska et al. find that stereotypes and beliefs are significantly linked to self-reported behaviour but not to observed behaviour, echoing LaPiere's (1934) landmark study.

While this highlights the need for a careful consideration of behavioural outcomes in addition to prejudice measures, the specific claims, particularly regarding the distinction between cognitive and affective variables, need to be read with caution. Despite a comprehensive literature search and broad inclusion criteria, Talaska and colleagues only included 57 studies with the median publication year of 1975. Additionally, their sample included very few studies that related cognitive prejudice measures to clearly cognitive outcomes, such as civic participation. Recent studies have linked ethnic stereotypes with support for specific policies, such as amnesties for undocumented migrants (Abramyan & Alexander, 2020), which is at least a necessary precursor to civic action. Therefore, the notion that cognitive measures are of lesser importance does not appear to rest on very strong evidence . This is a topic I will return to at various points in the thesis.

1.8.5. Contact promotes future approach of outgroup members

Social integration requires a willingness on the part of everyone to enter into contact with members of different groups. This is inhibited by prejudice (Herek &

Capitanio, 1996), and can thus be increased through positive contact.[2] Additionally, self-efficacy with regard to engaging in successful contact with someone who is different can be lacking, which can be increased through contact (Stathi et al., 2011; Turner & Cameron, 2016).

However, even with openness to contact, the fact that members of different groups often live somewhat "parallel lives" (Dixon et al., 2020; Social Integration Commission, 2014a) is a barrier to contact. For instance, cross-ethnic friends in school rarely meet elsewhere (Jugert et al., 2012), which constrains the strength of the friendship and reduces the likelihood that it will be maintained across transitions. Apart from spatial limitations, intergroup relationships can be hard to sustain in the absence of shared cultural references and the presence of conflicting norms (Pica-Smith, 2009). Furthermore, the impact of contact on the interest to make new cross-group friends is moderated by the perceived norm of ingroup peers (Tropp et al., 2014). This matters, because direct contact experiences, especially in the form of interventions, typically involve a small sub-section of a peer group; once participants return to their original context, accountability tends to increase in-group bias again, especially if their norms now conflict with peers' (Abrams et al., 2007). Therefore, the effect of contact can be greatly reduced by countervailing norms, so that it is important to study how contact might actively motivate approach rather than just increase openness.

A cautionary point regarding the link between intergroup contact and openness to future contact is that even fairly inclusive preferences can lead to segregationary dynamics as Schelling (1971) showed through mathematical modelling.[3] For example,

---

[2] Negative contact, conversely, has been shown to reduce approach intentions towards both the group encountered and even towards other outgroups. (Meleady & Forder, 2019)

[3] The *Parable of Polygons* provides a, accessible and instructive online simulation of these dynamics, which highlights that integrated neighbourhoods are difficult to sustain if there are even modest preferences for having ingroup-members around, unless there also is a distinct preference for diversity (https://ncase.me/polygons/).

if majority-status groups desire to live in neighbourhoods that are at least 70% majority, and minority-status groups prefer neighbourhoods that are at least one-third minority, the only stable configuration would be total segregation. However, even much more inclusive preferences can lead towards segregation if people take their expectations of future dynamics into account, as they might want to take pre-emptive action when they see a tipping-point approaching. Obviously, such models are gross simplifications, but they highlight that preferences for contact might need to be rather strong for integration to occur. Further research regarding neighbourhood segregation and integration will be discussed in Chapter 4, which presents a study that focuses on such preferences.

In spite of these potential constraints on the links between prejudice, approach intentions and social integration, experimental studies suggest that contact increases cross-group interactions in daily life, at least for a certain period (Page-Gould et al., 2008). Similarly, it has been shown that cross-group friendships in school predict the share of cross-group friendships in college (Ramiah et al., 2013; Stearns et al., 2009), and that such effects are even sustained into adulthood (Emerson et al., 2002), again suggesting that effects can be long-lasting and contribute to social integration.

1.8.6. Contact promotes bystander intervention and helping behaviours

Beyond the formation and maintenance of cross-group relationships, contact has been expected to increase assertive bystander behaviour in situations of exclusion and aggression. This has been confirmed in a range of recent research: contact has been shown to be associated with adolescents' judgement of race-based exclusions as wrong (Crystal et al., 2008) and with their intentions to intervene in a name-calling scenario (Abbott & Cameron, 2014). Similarly, contact has been shown to increase the willingness of straight college students to challenge the exclusion of lesbian, gay and bisexual fellow students (Dessel et al., 2017). Conversely, it has been found that a more homogeneous friend group is associated with a reduced likelihood of intervention

against racist incidents (Palmer & Cameron, 2010).

To date, there appear to be few studies that associate contact with actual interventions, rather than measuring intentions. This might be an issue because bystanders appear to systematically over-estimate how likely they are to intervene (Mulvey et al., 2016); however, there is no specific reason to expect that intergroup contact should moderate the size of that over-estimation, so that the associations are unlikely to be affected.

Regarding other helping behaviours, Johnston and Glasford (2018) found that positive and cooperative contact predicts both intentions and commitment to help a variety of ethnic outgroups, mostly mediated through empathy. Similarly, contact with mentally ill participants was associated with increased donations to a mental health aid fund (Corrigan et al., 2002), while contact with lesbian, gay and bisexual peers was associated with greater donations to an advocacy group (Reimer et al., 2017).

1.8.7. Contact can promote civic engagement

"Amicable relations among racial and ethnic groups can exist alongside grossly unjust inequalities of opportunities and outcomes. Ceteris paribus, harmonious race relations and unprejudiced attitudes might be worthy goals— but only if other things are equal, or nearly so'' (McConahay, 1978, p. 77).

Intergroup conflicts are not just driven by prejudice but also by social arrangements that afford different opportunities to members of different groups, as highlighted at the start of the chapter. Therefore, social integration requires political action. In addition to understanding inequality per se, participants need to recognise that individual concerns are political, and that different arrangements are possible. Such social imagination is a skill that needs to be developed. It involves the capacity of "looking at the world as if it could be otherwise" and thus identifying possible changes

and avenue for action (Ayers, 1995 p. 322). My previous research in German schools suggests that its lack is one of the strongest inhibitors towards civic engagement among adolescent, especially from minority groups (Wallrich, West, et al., 2021). By providing opportunities for perspective-taking and joint reflection, an intergroup contact context may provide a beneficial environment for developing social imagination, since insights into different lived realities might highlight that things could be otherwise. However, it is less clear how contact affects the perceived need and motivation for action.

1.8.7.1. Contact typically motivates the majority group to engage civically

At present, contact appears to motivate majority group members to participate in civic action against discrimination and aggression. For instance, positive contact with lesbian, gay and bisexual students predicts activism on their behalf by heterosexual peers (Reimer et al., 2017). Contact between Catholics and Protestants in Northern Ireland predicts civic engagement for peace (McKeown & Taylor, 2017), while contact with African-Americans predicted White Americans' willingness to take part in collective action in support of the Black Lives Matter movement (Selvanathan et al., 2017). However, few studies have investigated links between contact and civic action that addresses structural inequalities; therefore, it is not clear whether that is an extension of the "protection" behaviours studied, or whether it is associated with contact in different ways.

A specific pathway by which contact might spur majority-status members to civic action is the reduction of the "ultimate attribution error" (Pettigrew, 1979), i.e. the tendency to attribute failures of the outgroup to personal flaws and those of the ingroup to circumstance, while doing the opposite with regard to successes. Contact certainly affects this; for example, in segregated Northern Irish schools, students gave more negative explanations for outgroup unemployment and more positive explanations for

ingroup unemployment while this bias was reduced in an integrated college sample (Joseph et al., 1997; see also Vollhardt, 2010, who found a similar effect of contact across a wider range of situations).

However, contact can also prompt negative reactions in the majority status group, especially when the contact makes the disadvantage and resource need of a marginalised outgroup strikingly visible. Such a perception of pressing need can trigger egotistic distress rather than altruistic responses (Batson & Shaw, 1991), especially when the moral demands that arise appear overwhelming (Seider, 2009b, 2009a). Relatedly, knowledge about a long history of discrimination can backfire when it leads to existential guilt that is self-focused and stands in the way of action for the disadvantaged (Walker & Smith, 2002). However, empirical research has more frequently associated feelings of guilt resulting from contact with positive attitudes towards reparations (Brown et al., 2008), affirmative action (Swim & Miller, 1999), and the Black Lives Matter movement (Selvanathan et al., 2017). Nevertheless, there is various research on the 'principle-implementation gap' that cautions that general support for justice might often not translate into support for specific policies that promote equality (Dixon et al., 2007), so that it is not clear, for instance, how far general support for Black Lives Matter actually reaches when demands for restitution are considered.[4]

1.8.7.2. Contact often demobilises minority group members civically

In contrast to the generally positive impact on the civic engagement of majority-status people, research has increasingly revealed "the darker side of 'we'" for minority-group members (Dovidio et al., 2016). While positive contact often increases the sense

---

[4] Opinion polls on this specific issue highlight the discrepancy. For instance, in 2020, 63% of Americans supported the Black Lives Matter movement, but only 40% supported its primary demand of defunding the police (Langer, 2020). Even fewer (26%) supported the demand for reparations for the victims of slavery (Karson, 2020).

of injustice for majority-status participants, it can diffuse it among low-status group members, be it among Black people in South Africa (Dixon et al., 2010), Black and Latinx people in the United States (Tropp et al., 2012), Maori in New Zealand (Sengupta & Sibley, 2013) or Palestinians in Israel (Saguy et al., 2009). Relatedly, the diffusion of clear narratives can lead to inequality being seen as more legitimate (Sengupta & Sibley, 2013). As a result, positive contact with majority group members typically either has no impact on collective action against exclusion (Reimer et al., 2017), or – more problematically – reduces minority participants' readiness to engage in collective action to confront injustice (Hayward et al., 2018). Negative contact, on the other hand, appears capable of motivating minority group members to engage in civic action (Hayward et al., 2018; Reimer et al., 2017). Furthermore, contact that maintains the salience of intergroup differences and thereby limits the emergence of a common identity does not result in this demobilisation (Saguy et al., 2009).

1.8.8. *Contact effects can generalize beyond the outgroup encountered*

Given the "diversification of diversity" (Vertovec, 2007), intergroup contact would be a limited strategy if contact needed to occur with members of each outgroup. Fortunately, there is evidence that contact experiences generalize, so that contact with a specific group can shift broader attitudes. For example, Tausch and colleagues (2010) used a longitudinal design in Northern Ireland to show that contact between Catholics and Protestants also led to decreased racial prejudice, while Pettigrew and Tropp (2011) used German panel data to show that contact with foreigners predicted reduced prejudice against both gay people and the homeless.

Moving beyond prejudice, Meleady and Forder (2019) showed that negative contact with "Muslim immigrants" was not only associated with reduced interest in future contact with Muslims but with various other groups of immigrants as well. Furthermore, Flores (2015) showed that contact with lesbians and gay men predicted

greater support for transgender rights, controlling for contact with trans individuals. Beyond these findings, there has been limited research to date on the generalization of behaviours or behavioural intentions.

1.9. How does contact work?

Since the overall efficacy of contact has been established, particularly through the meta-analyses cited, much research has focused on boundary conditions, moderators and mediators.

1.9.1. Boundary conditions and moderators for contact effects

Originally, Allport (1954) had specified four conditions that he deemed necessary for contact to reduce prejudice: participants should have equal status in the contact situation, have common goals, work cooperatively towards them, and there should be authority support for the contact. The study that challenged the necessity of these conditions most directly took place in South Africa during the apartheid era (van Dyk, 1990). It showed that white housewives who had close contacts with their Black maids held more positive views of Black people in general. Accordingly, Pettigrew and Tropp (2006) find that contact is typically successful even in the absence of Allport's conditions, so that they do not serve as boundary conditions. However, contact programmes that emphasise equal status, interdependent collaboration and authority support produced significantly larger effect sizes; while they are not necessary, Allport's conditions moderate and facilitate the effects of contact. Some further moderators are considered in turn.

1.9.1.1. Group salience and typicality

There has been a recurring concern that while contact might improve attitudes towards the individuals encountered, it might not shift changes towards the outgroup as a whole (Amichai-Hamburger & McKenna, 2006; Amir, 1969). In the vast majority of

cases, however, this generalisation succeeds (Lemmer & Wagner, 2015; Pettigrew & Tropp, 2006). For that, a certain level of group salience is required for the path from contact to (group-level) prejudice to work (Brown & Hewstone, 2005), since otherwise the generalisation from positive contact with an individual cannot occur. Accordingly, various studies found that the effect of contact on prejudice is moderated by group salience (Harwood et al., 2005; Voci & Hewstone, 2003). Generalisation to the outgroup can also be prevented by subtyping, i.e. when pleasant interactions with members of an outgroup are taken to show that these outgroup members are exceptions to the general rules, i.e. exceptions to the stereotypes (Wilder, 1984). Accordingly, it has been found that contact effects are stronger when the outgroup members encountered are perceived as typical of their group (Binder et al., 2009; Wilder, 1984).

1.9.1.2. Closeness and intimacy

Additionally, intimacy matters. Merely increasing physical proximity without the occurrence of real positive contact can increase prejudice (Sherif & Sherif, 1953). Furthermore, contact quantity alone also does not appear to predict a shift in attitudes, when controlling for contact quality. Instead, friendships matter most (Tropp & Pettigrew, 2005a), and even among friends, the time spent with them and the level of self-disclosure predicts contact effects more than the number of outgroup friends or the time spent with them (Davies et al., 2011).

However, when looking beyond prejudice towards greater social integration, 'weak ties' also matter as they are more likely to provide access to new opportunities than close friends are (Granovetter, 2005). Similarly, weak outgroup ties appear to predict generalised social trust and outgroup trust more closely than outgroup friendships do (Stolle et al., 2013). This might be because social trust conceptually relates to how one views strangers rather than friends.

1.9.1.3. Initial level of prejudice

The initial level of prejudice appears to moderate the effect of contact, in that the strongest reductions are seen among those who are least tolerant to start with (Adesokan et al., 2011; Hodson, 2011), for example by being the most religious (Thomsen & Rafiqi, 2017).[5] While this might sometimes be due to a floor effect among less prejudiced participants, it also seems that a desire to act without prejudice, which is often highest among those lower in prejudice(Devine et al., 2002), can stand in the way of contact. Using a speed-dating format, Vorauer (2008) showed that among low-prejudice participants, contact tends to be more about evaluations of the ingroup, which reduces the willingness to disclose personal information to the outgroup member encountered. In line with the findings regarding the mediating role of self-disclosure, cited below, this was associated with the absence of a generalization of the contact experience in a way that would affect prejudice. Relatedly, Finchilescu (2010) showed that meta-stereotypes in the form of worries about how the outgroup evaluates the ingroup, for instance with regard to being prejudiced, heighten anxiety more than actual prejudice and are thereby a strong reason for the avoidance of contact.

1.9.1.4. Group status

Contact research has generally focused on the prejudice that majority-status participants hold towards minority-status participants, partly because prejudice in that direction tends to have the most harmful social consequences. Nevertheless, there is a growing consensus in the literature that a greater focus on the effect of contact on minorities is needed. With regard to prejudice reduction, contact appears to have a stronger effect on majority-status than minority-status group members (Lemmer &

---

[5] These effects might not last, as more prejudiced participants can be expected to return to comparatively intolerant peer groups; accountability can then increase in-group bias again (Abrams et al., 2007). Unfortunately, few if any studies with a long-term follow-up consider whether initial prejudice lastingly moderates the effect of intergroup contact.

Wagner, 2015; Tropp & Pettigrew, 2005b); group status also moderates the link between contact and behaviours with regard to addressing intergroup inequality as has been discussed above. Nevertheless, as inequality and exclusion are primarily maintained by those high in social status, a focus on high-status groups appears a good place to start when it comes to testing novel variables. Therefore, this dissertation will mostly focus on majority-status groups when it comes to the main effects, and only consider moderation effects by group status in the assessment of interventions (Chapters 5 and 6).

1.9.1.5. Quantity of contact above threshold

While intergroup contact has been shown to be good for intergroup relations, as outlined so far, there is a parallel body of literature that shows that single intergroup interactions often have negative effects, particularly because they typically lead to high levels of stress (Trawalter et al., 2009). It has been suggested that a threshold effect best explains this paradox, in that more experience with contact eventually leads to a point where the negative effects of single interactions are superseded by the positive effects of contact considered so far (C. C. MacInnis & Page-Gould, 2015). If this is the case, the relationship between contact and attitudes would not be linear, but rather have an inflection point at this threshold.[6] Given the fit of linear models, it might be the case that the threshold is reached in practice in many settings, but this possibility needs to be kept in mind, particularly for situations where contact is rare.

1.9.2. Mediation

Mediators, i.e., variables that explain the link between intergroup contact and the

---

[6] MacInnis and Page-Gould (2015) propose a complex curvilinear relationship which approximates the well-established linear relationship at high levels of contact. Taking their notion of a threshold more literally, a segmented regression model would be a more parsimonious way to think about the relationship as it would allow for the (linear) slope of contact effects to vary above and below the threshold. Such a segmented regression model could then simply be expressed as a model that includes an interaction term between quantity and a binary variable that indicates whether the threshold has been reached.

various outcomes, have been widely researched. In a meta-analytic study of the mediation of the link between contact and prejudice, Pettigrew and Tropp (2008) propose empathy, anxiety and increased knowledge as three possible pathways. They considered a set of 54 studies that tested at least one of these mediators, with anxiety being by far the most common (45 studies). In line with expectations, contact did indeed increase empathy and knowledge and decrease anxiety, and all were related with decreased prejudice. The effects for empathy and anxiety were more pronounced than those for knowledge, leading the authors to suggest that contact mostly works in affective rather than cognitive ways. However, this might also be a bias introduced by a dominance of affective measures of both contact (e.g., friendships) and prejudice (e.g., feeling thermometers) in the literature that are unsurprisingly more strongly related to affective than cognitive mediators; this will be further discussed in Chapter 3, which considers mediation.

Extending the pathways, Turner, Hewstone and Voci (2007) found in four studies with independent high school student samples that self-disclosure mediates the effect of contact on prejudice, primarily by strengthening the path from contact to empathy and trust (likely a reverse of anxiety).

A different mechanism that is well established but rarely considered in contact research is the 'mere exposure' effect (Zajonc, 1968) that suggests that stimuli become more likeable merely through repeated exposure. This is also true for other-race faces, even if they are just subliminally seen (Zebrowitz et al., 2008), and might thus explain the effect of contact that does not include closeness and self-disclosure, which still has an effect on intergroup emotions (Tropp & Pettigrew, 2005a). Relatedly, Turner et al. (2007) suggest that "mere exposure" to the outgroup is the strongest predictor of implicit prejudice.

When considering behavioural outcomes rather than prejudice, findings on mediation differ. For example, the effect of contact on bystander intervention intentions has been shown to be mediated by cultural openness (Abbott & Cameron, 2014). With regard to civic action, majority group members' intentions to join in collective action against discrimination of Black people and homosexuals were predicted by perspective taking in one study (Mallett et al., 2008). Considering White's support for the Black Lives Matter movement, Selvanathan et al. (2017) found empathy and anger over injustice to be sequential mediators of the effect of contact on collective action intentions.

In this thesis, valuing diversity will be considered as a mediator of the effects of intergroup contact on both attitudes and behavioural intentions, with the expectation that such beliefs will be particularly beneficial in explaining the cognitive effects of intergroup contact.

1.10. Critiques of the intergroup contact research

Overall, the effects of intergroup contact have been extensively researched, so that there is strong evidence regarding many key questions. Some limitations have already been pointed out. This section will focus on three further weaknesses of the current evidence, which have informed the research presented in this dissertation.

1.10.1. Lack of dynamic models of intergroup contact

Many relationships that are considered in intergroup contact research are likely bi-directional and recursive, yet typically only one direction is considered. At the most basic level, one can expect that prejudice predicts contact and contact predicts prejudice. This has led some to question the validity of the interpretation of all cross-sectional

studies, which is problematic given their predominance in the field.[7] For the relationship between contact and prejudice, specifically, it has been shown longitudinally that the influence runs in both directions, with similar effect sizes, at least among high school students, and for majority-status participants (Binder et al., 2009). However, generally such research is lacking.

Most intergroup contact research has assumed sequential models of intergroup contact, in which certain predictors make contact more likely (e.g., contact readiness, (Turner & Cameron, 2016), resulting in contact that then changes other variables (e.g., intergroup empathy, Johnston & Glasford, 2017), which mediate its effect on specific outcomes. However, many mediators are likely to also support future contact, so that recent reviews have called for the adoption of *dynamic* perspectives that consider the potential for virtuous cycles (or their opposite) to emerge (Paolini et al., 2018). A recent example of such an approach showed that intercultural competence is *both* an *antecedent* and a *consequence* of intergroup contact (Meleady et al., 2020). Similarly, this dissertation will start with a longitudinal study that tests the directionality of the (hypothesized) association between intergroup contact and diversity beliefs.

1.10.2. Rare consideration of beliefs, norms and values

Fifteen years ago, it was suggested that contact has a stronger effect on affective rather than cognitive measures of prejudice (Tropp & Pettigrew, 2005a). Since then, it appears that intergroup contact researchers have mostly considered affective outcomes and affective meditators such as empathy and anxiety, which supports the understanding of how intergroup affect can be improved, yet limits our understanding of how contact can contribute to social change (c.f. Dixon et al., 2005)

---

[7] For instance, in the largest meta-analysis to date (Pettigrew & Tropp, 2006), 71% of the studies considered were cross-sectional survey studies.

A key set of beliefs that has not been studied sufficiently in relation to intergroup contact are the diversity beliefs discussed above. It is likely that these both shape contact and are shaped by contact – these relationships will be discussed next and form the key research matter for this thesis.

1.11. Links between valuing diversity and intergroup contact

"Habitual contact in itself is no guarantor of cultural exchange" – (Amin, 2016, p. 969)

The links between intergroup contact and diversity beliefs have rarely been explored to date, even though these fields of research share similar aims. This is because the focus of diversity research has been its effect on group dynamics and productivity, while work linking diversity beliefs to intergroup relations has emerged more recently (Kauff et al., 2020). However, some studies show associations between intergroup contact and the valuing of diversity; these will be reviewed next.

1.11.1. Contact shapes the valuing of diversity

Positive contact with outgroup members could be expected to lead to positive beliefs about diversity. Indeed, one study has shown that US college students' openness to diverse perspectives increased over the course of two years in college if they had frequent interactions with members of other ethnic groups (Harper & Yeung, 2013). Unfortunately, that study does not control for intergroup contact at the baseline, so that it is unclear whether contact predicted changes in diversity beliefs over time, or simply increased alongside. This limitation was addressed by a study over a longer timeframe that found that support for a colour-blind ideology (here defined as one that minimizes or denies the existence of racism) decreased over the course of four college years if White students had close Black friends (Neville et al., 2014). Crucially, they estimated the cross-sectional effect of having Black friends, and the interaction with time. Only

the latter was significant, which suggests that it was indeed contact that influenced diversity beliefs. However, they did not test whether the reverse path was also present.

Given that a valuing of diversity is associated with improved intergroup attitudes, it appears plausible that this increase in the valuing of diversity might explain how intergroup contact reduces prejudice. Such a mediation hypothesis has only been tested once to date: Asbrock et al. (2012) showed that in a German probability sample about one third of the association between the number of foreign friends and the level of prejudice could be explained by increases in the valuing of diversity.

A separate, indirect path by which contact might increase the valuing goes through perspective-taking. Positive contact usually involves perspective-taking (Aberson & Haag, 2007). Taking the perspective of an outgroup member has been shown to increase positivity towards multiculturalism and decrease positivity towards colour-blindness. However, the link is reciprocal, with positivity towards multiculturalism facilitating perspective-taking (Todd & Galinsky, 2012). Given that perspective-taking increases desire to engage in intergroup contact (Wang et al., 2014), this shows that diversity beliefs are also likely to motivate contact.

1.11.2. Valuing diversity leads to contact

Conversely to contact shaping beliefs regarding the value of diversity, it appears plausible that a valuing of diversity would foster contact. This has found support in a questionnaire study (Tropp & Bianchi, 2006), which suggested that a valuing of diversity uniquely predicted interest in intergroup contact, at least for majority-status group members. Similarly, Wolf and Van Dick (2008, cited in Kauff et al., 2020).showed that both German's interest in seeking out contact with foreigners as well as their number of foreign friends correlated with their (instrumental) valuing of diversity) Furthermore, fields studies found that a valuing of diversity was associated

with greater diversity in 552 naturally occurring relationship pairs (Bahns, 2017), which again suggests that it is associated with greater willingness to engage in intergroup contact. Here it worth noting, however, that the presence of a 'diverse' partner might enhance diversity beliefs in that situation compared to the presence of a similar partner.

Recent work on *xenophilia* (e.g., Stürmer et al., 2013) conceptualises a desire to experience diversity differently. It shows that xenophilia is empirically distinct from low xenophobia and suggests that it might have similarly deep evolutionary roots, given that learning from outgroups is as foundational a need as protection from outgroups.

1.11.3. Diversity beliefs shape contact

Associations between diversity beliefs and the nature of intergroup interactions have already been described above. These are clearly also of interest to intergroup contact research; two lines of enquiry are of particular relevance here. Firstly, experimental work by Vorauer (Vorauer, 2008; Vorauer et al., 2009) shows that valuing diversity shifts the focus on an outgroup interaction partner rather than on the self, and thereby leads to more positive encounters, particularly as experienced by minority-status participants. Secondly, a high degree of self-expansion motivation, i.e. the desire to experience personal growth through the encounter has been shown to lead to more positive outgroup interactions, as well as a higher quantity of contact (Dys-Steenbergen et al., 2016).

However, in this context it is important to note that the quality of contact depends on the content of the diversity beliefs that triggered it. Ely and Thomas (2001) identified three motivations managers hold to increase diversity in the workplace: ensuring access and legitimacy, addressing discrimination and fairness, or promoting integration and learning. All could motivate diverse hiring, but only the last supported a sustained engagement with and benefit from diversity.

Diversity beliefs are also likely to affect the focus of contact, i.e., the relative share of conversations about similarities and differences. This has been shown to determine the association between contact and civic action on the part of majority status participants – an exclusive focus on a common identity, as implied by colour-blind beliefs, fails to engage majority group members civically. Vezzali et al. (2017) were interested in the impact of contact that explicitly considers differences and found that frequent contact that explicitly addresses differences was most associated with a motivation to act for change, while the frequency of contact that focuses on commonalities was unrelated. As might be expected, at low levels of contact, a focus on differences was associated with worse outgroup attitudes, yet when the contact was frequent, both a focus on commonalities and differences resulted in more positive attitudes.

1.11.4. Engaging with diversity in a contact situation – the role of group salience

Category salience has already been introduced as a moderator required for contact effects to generalize, but there has been a long debate regarding the role of differences in intergroup contact. These align well with the debates around diversity beliefs. In line with colour-blindness, some researchers proposed to emphasize a common identity that transcends differences. Based on that approach, recategorization and colour-blindness in the contact situation were seen as key to guaranteeing a positive experience that could then have positive impacts (Gaertner et al., 1993). This initially appears plausible, given that the negative consequences of categorisation into even minimal groups have been well established (Tajfel, 1969). However, it has been shown that this can backfire for participants who strongly identify with the separate groups, who then feel threatened by the attempt at recategorization and emerge with greater bias (Crisp et al., 2006). In addition, subsequent research has shown that a salience of the

outgroup enhances contact effects (Oudenhoven et al., 1996). This has been the basis of the so-called dual-identity approach that aimed to keep both common and distinct identities salient in the contact situation (Hewstone & Brown, 1986).

More recently, even former advocates of colour-blind approaches to intergroup contact have expressed worries that this in fact enshrines the privilege of high-status participants by hiding group-based inequity (Dovidio et al., 2016). Therefore, they now advocate for contact that includes conversations about group-based differences. Valuing diversity might provide an avenue into such conversations that does not immediately lead towards guilt and defensiveness.

1.12. Overview over the present research

It has been proposed that contemporary challenges require the establishment of a diversity science (Plaut, 2010), and a 'diversity hypothesis' has been proposed to complement and update the 'contact hypothesis' (Jones et al., 2000). This dissertation contributes to that agenda by considering the dynamic relationship between intergroup contact, primarily with other ethnic groups,[8] and positive diversity beliefs, i.e., the valuing of diversity. The following hypotheses are tested in turn:

H1.   The relationship between intergroup contact and valuing diversity is *dynamic*. Specifically, longitudinally, intergroup contact leads to greater valuing of diversity *and* valuing diversity longitudinally leads to greater interest in intergroup contact (*Chapter 2*).

---

[8] Conceptually, this work is interested in different types of diversity. However, in the studies that are to come, I largely focus on ethnicity as an indicator for diversity. Despite an increasing focus on other dimensions, ethnicity is still the focus of the official diversity discourse (cf. the 'diversity index' published by the UK Office for National Statistics that measures ethnic neighbourhood integration, Large & Ghosh, 2006). This matches popular discourse: at least in the United States, while people define diversity in broad terms, most of their concrete references concern interactions with members of other ethnic groups (Bell & Hartmann, 2007) so that this is likely the most salient dimension when people are asked about diversity *per se.*

H2. Valuing diversity mediates the link between intergroup contact and its primary desired outcomes (reduced prejudice and increased willingness to counter exclusion), with particular influence when it comes to cognitive outcomes (*Chapter 3*).

H3. Valuing diversity predicts support for inclusive choices that counter social segregation and support approach (*Chapter 4*).

H4. Contact interventions increase the valuing of diversity, particularly when participants engage in conversations about differences and enter the intervention with high self-expansion orientation (*Chapter 5*).

H5. Dedicated activities that promote the value of diversity can enhance the effect of intergroup contact on valuing diversity and openness to future contact (*Chapter 6*).

Given that the relationship between intergroup contact and valuing diversity, and the role of valuing diversity as a mediator of contact effects are the key concerns of this dissertation, *Chapter 7* presents an internal meta-analysis that aggregates all datasets used in this dissertation to test the proposed model.

CHAPTER 2:

Contact experiences predict valuing diversity longitudinally

In order to begin the empirical exploration of the relationships between intergroup contact and valuing diversity, this chapter will explore their association over time. To date, several studies have associated valuing diversity with intergroup contact, for instance through the observation of naturally occurring dyads (i.e. pairs of people in contact), where members of diverse dyads held more favourable views of diversity (Bahns, 2017; Bahns et al., 2015). Directionally, it has been shown that valuing diversity predicts greater *interest* in contact, at least for majority group members (Tropp & Bianchi, 2006; Yogeeswaran et al., 2020). Similarly, in an experimental study among primary school children, exposure to a storybook promoting multiculturalism led to a (short-term) reduction in ethnic self-segregation during lunchtime, though the effect disappeared within 48 hours (McKeown et al., 2017). Similarly, among adults valuing diversity has been shown to be associated with lower intentions to avoid exposure to and interactions with outgroup members (Kauff & Wagner, 2012, Study 1). However, there is little research to date that links diversity beliefs to the incidence of intergroup contact.

As far as I am aware, only one study has tested the link between intergroup contact and diversity beliefs over time. In their sample of US college students, Harper and Yeung (2013) found that for White students, the frequency of social interactions with students of a different race or ethnicity over the course of two years predicted a greater valuing of diversity in their junior year. However, they only controlled for valuing diversity at baseline and did not include an autoregressive path for contact. This weakens the longitudinal approach substantially as it is not clear whether an earlier level of contact predicted valuing diversity, or whether the frequency of contact changed

alongside changes in valuing diversity. Furthermore, they did not consider the valence of contact, implicitly following the majority of contact research in conflating contact with positive contact.

There is also a range of literature that shows that past intergroup contact predicts future intergroup contact (Binder et al., 2009; Braddock & McPartland, 1989; Emerson et al., 2002). It has been suggested that this effect is mediated through increased valuing of diversity (Bahns, 2017). This would induce a virtuous cycle in which contact and valuing diversity reinforce each other, but this is not yet supported by empirical evidence. Indeed, the need for longitudinal studies to further our understanding of the links between intergroup contact and valuing diversity has been highlighted by a recent review (Kauff et al., 2020).

This study aimed to test the dynamic association between intergroup contact and valuing diversity. It extended earlier studies that had established an association (e.g., Harper & Yeung, 2013; Tropp & Bianchi, 2006) by employing a cross-lagged panel model to test the directionality of the effect longitudinally. The specific hypotheses were:

H1: Experiences of positive and negative contact predict positive and negative changes in the value placed on diversity over time.

H2: Valuing diversity predicts the experience of more positive contact over time.

2.1. Methods

2.1.1. Power calculations

The study required the estimation of a path model, which I initally wanted to base on observed variables. In that case, it would have contained 15 paths. Given the frequent recommendation in the literature that the sample should contain at least 10

cases per parameter to be estimated (e.g., Jackson, 2003; Kline, 2015), I aimed for a sample size of 150 participants. After a reviewer pointed out that this approach would fail to consider measurement invariance, I now present a structural equation model with latent variabels, for which the sample size is rather small. However, the original path model is shown in Appendix 2.1, and yielded nearly identical results.

2.1.2. Participants

Data for this study were collected in two surveys of psychology students in two English universities carried out at the beginning of Semester 1 (T1, October) and the beginning of Semester 2 (T2, February) in their first academic year. University 1 was based in London and had 51% of Black and minority-ethnic students, while University 2 was located in Norwich, with 15% of Black and minority-ethnic students. Only the responses from respondents who identified as White were considered; 211 such students responded to the survey at T1, out of which 151 returned at T2 (72%). Eighty-one percent were female; the mean age was 19.4 years at T1 ($SD = 2.9$ years, range: 17 to 43 years). The two timepoints were approximately 110 days apart, which is in line with other recent studies concerning the longitudinal effects of intergroup contact (e.g., Meleady et al., 2020; Reimer et al., 2017).

The research adhered to all BPS ethical guidelines and was approved by the Goldsmiths' Psychology Departmental Ethics Committee. Participants could receive partial course credits for their participation

2.1.3. Measures[9]

In the surveys, I measured White students' contact with Black students. Black students were selected as the focal outgroup as they are the largest ethnic minority in the

---

[9] The questionnaire included some additional measures to support an ongoing project regarding the interaction between positive and negative contact. The full questionnaire and dataset are available on GitHub and the OSF.

United Kingdom (3.9% of 18-19 year-olds in England and Wales, GOV.UK, 2019) that can plausibly be treated as a single group. Though Asian British make up a greater share of the UK population, attitudes towards different Asian British groups (e.g., Indian, Bangladeshi and Pakistani), strongly diverge, rendering the overarching category less ideal for joint analysis (Ehsan, 2018).

2.1.3.1. Contact.

To measure positive and negative contact, the measure developed by Reimer et al. (2017) was used. For negative contact, participants were asked how often they had been "ridiculed", "verbally abused", "made to feel unwanted", "intimated" or "threatened" by Black British people, from 1 = *Never* to 7 = *Very often*, with Cronbach's α = .75 at T1 and α = .86 at T2. Likewise, positive contact was measured by asking how often participants had been "complimented", "befriended", "made to feel welcome", "supported" or "helped", from 1 = *Never* to 7 = *Very often*, with Cronbach's α = .86 at T1 and α = .88 at T2.

2.1.3.2. Valuing diversity.

The value participants place on diversity was measured as the mean of three items adapted from Tropp and Bianchi (2006), asking to what extent they value it in British society, in their university and in their group of friends, from 1 = *Not at all* to 7 = *Very much,* with Cronbach's α = .89 at T1 and α = .91 at T2.

2.1.4. Attrition and missing data

As indicated above, 60 participants (28%) who had responded to the survey at T1 did not return at T2. In order to test whether this missingness could be assumed to be completely at random, following Ridout (1991), I estimated a logistic regression with drop-out as the dependent variable and contact, valuing diversity, gender, age and university as predictors. This showed that women were less likely to drop out (OR =

0.32, p = .004), and students at University 2 were marginally more likely to drop out

(OR = 2.1, p = .08). The substantive variables were not related to drop-out, ORs <= 1.2,

ps > .2. Therefore, while missingness was not completely at random, it is unlikely to

introduce substantial bias. Full-information maximum likelihood estimation was used,

as it has been shown to avoid bias and maintain both power and proper rejection rates in

longitudinal designs, even in cases of small samples with non-normal distributions

(Shin et al., 2017).

Apart from drop-out between the waves, there were a small number of missing

responses due to items being ommitted on the questionnaires. This was the case for no

more than five respondents per variable (2.4% of the sample). Given this low rate of

missingness, pairwise or casewise deletion could be used without substantial risk of

introducing bias. However, given that full-information maximum-likelihood approaches

achieve greater power without introducing bias, these cases were retained in the data.

2.1.5. *Analytical approach*

Analyses were conducted using the *lavaan* 0.6-6 package (Rosseel, 2012) in R

4.0.1 (R Core Team, 2020), using the FIML estimator. After testing for measurement

invariance, following Mackinnon et al. (2020), I estimated a cross-lagged panel model.

2.2. Results

Descriptive statistics and correlations between the variables are shown in Table

2.1. These were generally in line with expectations. It is worth noting that positive and

negative contact were only weakly correlated at T1, r = .16, p = .016, and unrelated at

T2, r = -.05, p = .58, confirming that they indeed represent distinct dimensions.

Negative contact was generally more *weakly* linked with valuing diversity than positive

contact was, even though this difference was only significant at T1, p = .024.

Table 2.1. *Descriptive statistics and correlations between the variables*

Variable	M (SD)	1.1	1.2	2.1	2.2	3.1
1.1 Positive contact T1	4.41 (1.43)					
1.2 Positive contact T2	4.51 (1.35)	.54 *** [.43, .66]				
2.1 Negative contact T1	1.76 (0.95)	.16 * [.03, .30]	.08 [-.08, .24]			
2.2 Negative contact T2	1.87 (1.01)	.05 [-.11, .20]	-.05 [-.21, .11]	.55 *** [.44, .67]		
3.1 Valuing diversity T1	6.14 (1.02)	.22 ** [.09, .35]	.22 ** [.07, .36]	-.02 [-.16, .12]	-.08 [-.23, .07]	
3.2 Valuing diversity T2	5.75 (1.11)	.22 ** [.08, .37]	.33 *** [.19, .47]	-.17 * [-.33, -.01]	-.28 *** [-.43, -.13]	.46 *** [.34, .58]

Notes. M and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation.

T1: Measured at timepoint 1   T2: Measured at timepoint 2

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

2.2.1. Longitudinal association between contact and valuing diversity

To test Hypotheses 1 and 2, I estimated a cross-lagged panel model linking positive and negative contact experiences and the valuing of diversity. The first step was to test whether the longitudinal measurement model fit the data and whether measurements were invariant over time (Little et al., 2007). For that, I estimated the measurement model, consisting of latent factors and the attendant items for positive and negative contact as well as the valuing of diversity at both time points, with freely estimated parameters. The variance of each latent variable was set to unity to scale the factor and the residual errors of repeated measurements of the same item were allowed to correlate, in order to reflect systematic error over time (Mackinnon et al., 2020). The goodness-of-fit was evaluated by considering the χ^2-statistic, the comparative fit index (CFI), the root-mean-square error of approximation (RMSEA) and the standardised root-mean-square residual (SRMR). CFIs greater than .90, RMSEAs and SRMRs below

.08, as well as χ2/df ratios below 3 were taken as evidence for satisfactory model fit

(Kline, 2015). The measurement model showed satisfactory fit, with $\chi^2(271) = 518.0$, p

$< .001$, χ2/df = 1.91, CFI = .91, RMSEA = .066 and SRMR = .070. Next, in order to test

for measurement invariance, factor loadings of the same items were constrained to be

equal at the two time-points, and the fit of the restricted model was compared to that of

the unrestricted model. The change in BIC was taken as the leading criterion for model

selection (Lin et al., 2017; Mackinnon et al., 2020). This suggested that the restricted

model was preferable, ΔBIC = -37.8, so that metric measurement invariance was

confirmed.[10] I could thus move on to testing the cross-lagged panel model.

For that, I first estimated a multi-group model with separate parameters for the

two universities and compared it to a model in which all parameters were constrained to

be equal across the two universities. The more parsimonious constrained model did not

show worse fit, $\chi^2(29) = 8.56$, $p > .99$, so that a pooling of the samples was warranted.

The resulting model is shown in Figure 2.1; paths not shown had p-values above .1 and

the coefficients are fully standardised.

---

[10] As an alternative to relying on ΔBIC, Mackinnon et al. (2020) propose considering the changes in AIC, CFI and BIC as well as a χ^2 significance test and using the majority verdict to determine model selection. Here, ΔAIC was 5.7, below the cut-off typically taken to indicate a difference between the models (Raftery, 1995). Similarly, ΔCFI was below -.01, which again indicates that the null-hypothesis of invariance should not be rejected (G. W. Cheung & Rensvold, 2002). Only the χ^2 test pointed in a different direction, with $\chi^2(13) = 31.71$, p = .003. Overall, the results thus lead one to retain the null hypothesis of measurement invariance.

Figure 2.1. *Cross-lagged panel model connecting contact and valuing diversity*
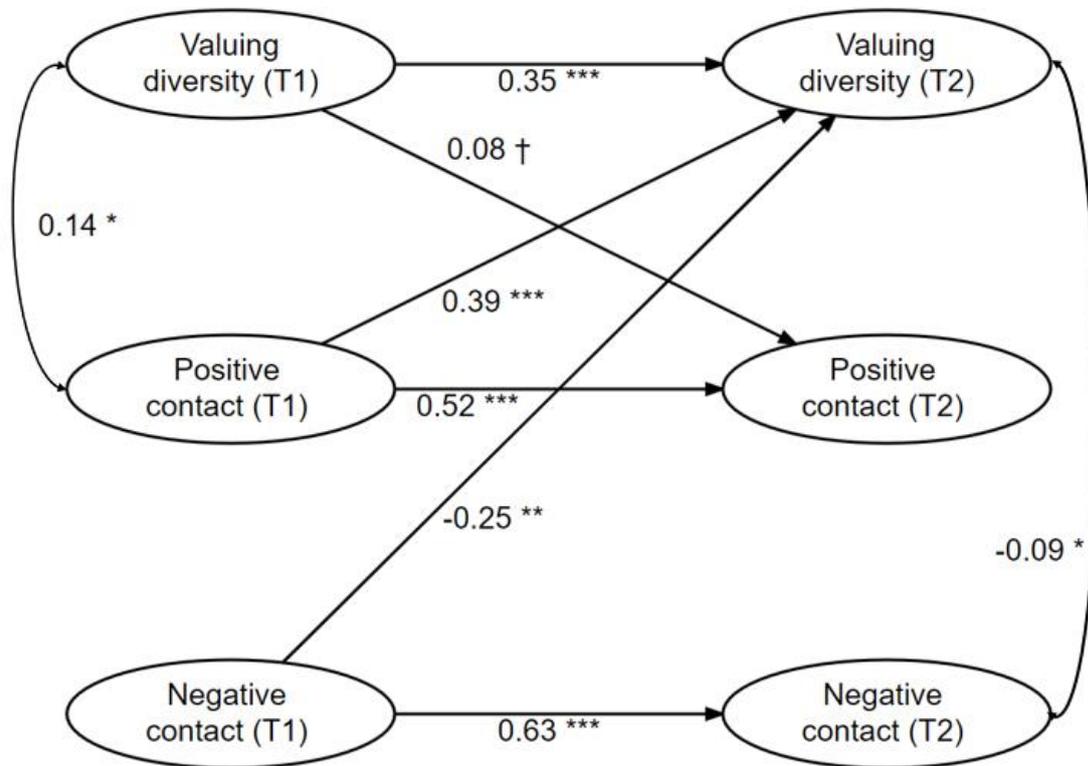
Notes: Standardized coefficients; paths with *p*-values above .1 are not shown for simplicity.
      * p < .05, ** p < .01, *** p < .001

All auto-regression paths were substantial and significant. Negative contact at T1 predicted lower valuing of diversity at T2, with $\beta = -0.25$, $p = .006$, 95% CI [-0.42, -0.07], while more frequent positive contact had the opposite effect, with $\beta = 0.39$, $p < .001$, 95% CI [0.19, 0.60]. In the opposite direction, valuing diversity at T1 did not significantly predict positive contact at T2, with $\beta = 0.08$, $p = .08$, 95% CI [-0.01, 0.17] or negative contact at T2, with $\beta = -0.03$, $p = .62$, 95% CI [-0.16, 0.09]. Comparing the strength of reciprocal paths, the path from positive contact to valuing diversity was significantly stronger than the path from valuing diversity to positive contact, $z = 2.73$, $p = .006$. For negative contact, the difference approached significance, with $z = 1.94$, $p = .052$. Full results are reported in Table 2.2.

Table 2.2. *Results of the Cross-Lagged Panel Model Connecting Contact and Valuing Diversity*

Predictor	Valuing diversity (T2) β [95% CI]	p	Positive contact (T2) β [95% CI]	p	Negative contact (T2) β [95% CI]	p
Valuing diversity (T1)	0.35 [0.21, 0.49]	< .001	0.08 [-0.01, 0.17]	.078	-0.03 [-0.16, 0.09]	.615
Positive contact (T1)	0.39 [0.19, 0.60]	< .001	0.52 [0.33, 0.71]	< .001	-0.09 [-0.27, 0.09]	.342
Negative contact (T1)	-0.25 [-0.42, -0.07]	.006	0.10 [-0.01, 0.22]	.072	0.63 [0.42, 0.83]	< .001

Notes: Standardised coefficients estimated with full-information maximum-likelihood.
T1: Measured at timepoint 1   T2: Measured at timepoint 2

Therefore, Hypothesis 1 was supported; both positive and negative contact predicted changes in valuing diversity over time. Hypothesis 2 was not supported; over the period considered here, valuing diversity did not significantly predict the frequency of positive or negative intergroup contact, and at least for positive contact, the path from contact to valuing diversity was stronger than the path from valuing diversity to contact.

2.2.2. *Changes in valuing diversity over time*

For many students, the transition into university represents a time when they encounter a more diverse range of peers than in school, and when the topic of diversity gains increased salience. Therefore, it was interesting to see whether the value placed on diversity would systematically differ between the timepoints. Comparing the descriptive statistics in Table 2.1, there appears to have been a decline in the valuing of diversity. The mean of valuing diversity at T2 was significantly and substantially lower than at T1, $z = 4.48$, $p < .001$, $d = .34$.

2.3. Discussion

The results suggest that both positive and negative contact experiences shape the value students come to place on diversity over time. Therefore, it appears that positive

intergroup contact is a promising avenue towards promoting the valuing of diversity, which has been associated with a variety of positive intergroup outcomes. Regarding the suggestion from earlier cross-sectional research that valuing diversity would also lead to more intergroup contact and thus fuel a virtuous cycle between contact and valuing diversity, the results are inconclusive: while the estimated paths from valuing diversity to future contact point in the expected direction, the coefficients are small and not statistically significant. While three months might have been too short to reveal an effect, the context of students' transition into their university life should have provided a favourable context for such an effect to emerge, given that most social relationships are in flux at such a time. Also, there were opportunities for increased contact with Black students at both universities; one had 5.5% Black students, while the other had 13.2% Black students, yet there was no evidence for different relationships between the variables in the two universities. Therefore, it appears warranted to focus on the effect of contact on valuing diversity and to explore the consequences of an increase in valuing diversity.

Concerningly, the results also indicated that students' valuing of diversity decreased on average between the timepoints. This might be a symptom of a 'diversity shock' in which the transition to a more diverse context is initially challenging for majority-status group members. Such an effect has been found earlier in the transition from homogenous primary schools to more diverse secondary schools (Birtel et al., 2020). It suggests that such transitions might be worthwhile contexts for interventions and highlights an opportunity for further research.

2.3.1. Limitations

Longitudinal models should ideally include random intercepts to reliably separate the within-person process from stable between-person differences (Hamaker et al., 2015). However, this was impossible to implement under the circumstances as it

requires at least three waves; two waves do not allow one to distinguish stability due to autoregression from stability due to trait differences. Therefore, future research with more timepoints should be conducted to validate the findings. This would also enable one to test whether the relationship between the variables differs depending on the duration under consideration and, for instance, whether valuing diversity might come to shape contact over a longer timeframe.

2.4. Conclusion and connection

The results presented here strongly suggest that positive experiences of intergroup contact increase the valuing of diversity, while negative experiences decrease it. Given that valuing diversity has been shown to contribute to positive intergroup relations in various ways, this suggests that it might serve as a mediator that can help explain how contact effects come about. The following chapters test that possibility.

CHAPTER 3:

Diversity beliefs mediate the impact of intergroup contact

on attitudes and inclusive behaviours

Having established in the previous chapter that intergroup contact shapes the valuing of diversity, I now explore to what extent this might explain how established outcomes of intergroup contact come about. Below, these outcomes will be introduced, before I turn to a discussion of mediation.

3.1. Outcomes of intergroup contact

Over many decades of research, intergroup contact has been established as one of the strongest precursors to prejudice reduction (Pettigrew & Tropp, 2006). More recently, there has been a greater focus on behavioural outcomes, as social transformation requires changes in behaviour alongside changes in attitudes. Through that work, intergroup contact has been shown to contribute to prosocial behaviour in both the civic and the personal sphere, at least among members of majority-status groups (Abbott & Cameron, 2014; Reimer et al., 2017; Selvanathan, Techakesari, Tropp, & Barlow, 2017). This is partly because contact leads to improved attitudes, and because attitudes tend to predict behaviour. For instance, studies have found that the level of racist beliefs predicts the signing of petitions (Brannon et al., 1973), the outcome of juridical judgments (Dovidio et al., 1997, Study 2) and interview invitations in a hiring process (Brief et al., 2000). However, the prejudice-behaviour link is not particularly strong, as suggested by general research on the attitude-behaviour gap and specific research on prejudice. With regard to the former, a meta-analysis of 88 studies found a mean r of .38 between attitudes and behaviours (Kraus, 1995), while a meta-analysis of 60 studies that specifically focused on links between prejudice and discriminatory behaviours estimated the population correlation to be .29 (Schütz & Six,

1996). Thus, both types of outcomes should be considered separately in research on the effects of intergroup contact.

3.1.1. Established mediators of intergroup contact

So far, research has considered a wide range of mediators of the link between contact and its potential outcomes. In this chapter, I consider valuing diversity alongside the most established mediators to see whether it can contribute to explaining the links and whether the mediation patterns differ depending on the outcome under consideration.

Most mediation studies to date have focused on explaining the effect of intergroup contact on prejudice. Knowledge about the outgroup, intergroup empathy and anxiety have been considered most frequently; Pettigrew and Tropp (2008) conducted a meta-analysis of 54 papers, including 91 samples, that included one or more of these mediators. They concluded that there is evidence for mediation through each of the three candidate mediators. However, while the paths from contact to each of them had similar strength, knowledge was a much weaker predictor of prejudice than empathy and anxiety, so that its contribution to explaining the association of contact with prejudice was much lower.

In line with that finding, their meta-analysis concludes that "affective factors, such as anxiety reduction and empathy, are clearly major mediators relative to the more cognitively oriented mediator of knowledge" (p. 929). However, the evidence regarding the low(er) importance of cognitive factors is weak. Initially, it needs to be noted that the correlations between knowledge and prejudice were highly heterogeneous across the 17 samples that included measures of knowledge, with rs ranging from +.13 to -.38, and that the way the construct was operationalized differed hugely. Some studies did not measure actual knowledge but focused on other variables that can be considered weak

proxies for knowledge. For instance, Eller and Abrams (2004) measured the frequency of engagement with the outgroup culture while Goto and Chan (2005) measured the perceived familiarity with the outgroup culture, which was only minimally related to actual knowledge for the subset of the sample where they also measured knowledge ($r = .10$, $p = .06$). Some studies employed very unreliable measures, culminating in a Facts on Aging Quiz with a coefficient α of .07 (Meshel, 1997), while only few measured knowledge in the way one might expect, with questions focused on the outgroup under consideration and a reliable instrument (e.g., Holmes et al., 1999). Furthermore, one might expect that the relative importance of affective and cognitive mediators differs depending on whether the prejudice measure employed is affective or cognitive. Even though they had shown before that intergroup contact affects affective and cognitive prejudice differently (Tropp & Pettigrew, 2005a), Pettigrew and Tropp did not consider differences between the outcome measures in their meta-analysis of mediation, probably because of the small number of studies. Thus, while it appears established that empathy and anxiety are important mediators, the role of cognitive factors deserves further study, especially with regard to different possible outcome measures, as previously highlighted.

Furthermore, most research into the effects of intergroup contact has focused on changes in prejudice, rather than changes in behaviour, which are arguably more important when it comes to the required social changes that motivate most research into intergroup relationships. However, there has been some work into the mediation of the links between contact and behaviour/behavioural intentions. The effect of contact on *bystander intervention intentions* was mediated through empathy, in-group bias and cultural openness, but not through anxiety (Abbott & Cameron, 2014). The effect on *approach intentions* was mediated through increased trust, improved attitudes and – in some analyses – reduced anxiety (Turner et al., 2013). The effect on *helping intentions*

was mediated through anxiety (Hutchison & Rosenthal, 2011). Finally, the effect on *homophobic behaviours* was mediated through reduced anxiety and reduced prejudice (Mereish & Poteat, 2015). Overall, anxiety and intergroup attitudes emerge as the mediators with the strongest evidence base; both empathy and cognitive mediators have rarely been tested in this context. Additionally, most research has focused on interpersonal behaviours, rather than higher level activities such as support for policies that address intergroup inequalities. This will be tested as a potential contact outcome here.

Overall, further study regarding the role of cognitive factors in the mediation of contact effects in relation to the valence of contact and the nature of the outcome is needed. Here, this will take place with a focus on valuing diversity. To date, only one study has tested whether the valuing of diversity can help explain the association between contact and prejudice. Asbrock et al. (2012) showed that in a German probability sample about one third of the association between the number of foreign friends and the level of prejudice could be explained by increases in the valuing of diversity. However, they did not consider negative contact and restricted themselves to a single outcome. Thus, their findings will be replicated and extended here.

3.1.2. Hypotheses

The study tests four models of parallel mediation, each aiming to disentangle the association between (positive and negative) contact and one of four possible outcomes of contact: two types of prejudice (cognitive and affective) and two types of behavioural intentions (policy support and bystander intervention). The following hypotheses are tested:

H1: There will be significant indirect paths through empathy, anxiety and valuing diversity in each model, consistent with a hypothesis of parallel mediation.

H2: The paths through valuing diversity will be stronger for the cognitive outcomes (cognitive prejudice and policy support) compared to the models for affective outcomes.

H3: Conversely, the paths through empathy will be stronger in the affective rather than the cognitive models (i.e., for affective prejudice and bystander intervention intentions).

H4: The paths through anxiety will be similar across the outcomes.

In order to focus the construct of valuing diversity on cognitive processes, this chapter employs a measure that focuses on beliefs regarding the value of diversity for British society rather than the valuing of diverse relationships in one's personal life.

3.2. Methods

3.2.1. Power analysis

Power for complex mediation models cannot be assessed analytically; instead, Monte Carlo simulations are recommended (Schoemann et al., 2017; Z. Zhang, 2014). However, to date such Monte Carlo simulations are only easily accessible for some basic model types, for instance in the web application that accompanied the paper by Schoemann et al. (https://schoemanna.shinyapps.io/mc_power_med). As the source code for that app is freely available, I extended it to provide support for a model with three parallel mediators, in line with what I wanted to test here. (https://github.com/LukasWallrich/mc_power_med).

With that application, I conducted a power analysis with 5,000 replications, consisting of 20,000 Monte Carlo draws each, as recommended by Schoemann et al. (2017). Based on a review of previous research, I assumed that contact, the mediators and outcomes would be correlated with medium strength ($r = .3$), while I assumed the correlations between the mediators to be weaker ($r = .2$), given that I would not expect

direct relationships between them. This analysis suggested that a sample size of at least 213 participants would be required to obtain 80% power to detect the indirect effects.

3.2.2. *Participants*

Two hundred and seventeen undergraduate psychology students who identified as White participated in this study (89% female). All were either British citizens (95%) or immigrants who intended to stay in the United Kingdom (5%); international students who were only in the UK to attend university were excluded from the sample.[11] The participants' ages ranged from 18 to 45 years (M = 20.5, SD = 4.3).

Participants were recruited in five universities in London and South East England to ensure socio-economic and ethnic diversity in the participants' life experiences; across these universities, the share of minority-ethnic students among incoming undergraduates ranged from 15.4% to 60.6%, while the share of Black students, specifically, ranged from 4.2% to 35.4%, compared to a UK average of 23.4% for minority-ethnic and 7.7% for Black students (data for the 2016/17 academic year, retrieved from UCAS and HESA statistics).

3.2.3. *Measures*

The questionnaire included measures of intergroup contact, the proposed mediators (empathy, anxiety, valuing diversity) and the proposed outcomes (prejudice, bystander intervention intentions and support for policies that address inequality and promote cultural inclusion).[12]

---

[11] As some data collection took place in lectures where no participant screening was possible, 66 Britons from a minority ethnic background and 37 international students participated. They were excluded from the analyses.

[12] As data were collected in collaboration with academics in each university, the questionnaire included a wide range of additional measures, only some of which have been analysed fully to date. The full questionnaire and dataset are available on GitHub and the OSF.

3.2.3.1. Contact.

Participants were asked to report their experience of positive and negative contact with Black British people. This was measured by asking about the frequency of five positive (Cronbach's α = .90) and five negative (α = .91) experiences, taken from Reimer and colleagues (2017), e.g., "being complimented," "being ridiculed" (1 = *never,* 7 = *very often*). Black British people were selected as the outgroup as they are the largest ethnic minority in the United Kingdom (3.9% of 18-19 year-olds in England and Wales, GOV.UK, 2019) that can plausibly be treated as a single group. Though British Asians make up a greater share of the population, attitudes towards Indians on the one hand, and Bangladeshi and Pakistani on the other, strongly diverge, rendering the overarching category unsuitable for joint analysis (Ehsan, 2018).

3.2.3.1. Mediators

3.2.3.2. Valuing diversity

Valuing diversity was measured with five items drawn from Adesokan et al. (2011), anchored to 1 = *strongly disagree*, 7 = *strongly agree*, Cronbach's α = .87:

- British society generally benefits from the involvement of people from different cultural backgrounds.
- Too many people from different cultural backgrounds can be a recipe for trouble. (*reversed)*
- It is easier to solve problems in Britain (politics, economy) if there is input from people who are different from each other.
- Being a multi-ethnic nation is an advantage for achieving progress in the UK.
- Different ethnic/cultural groups are enriching to British culture.

3.2.3.3. Empathy

Intergroup empathy was measured with two items taken from Swart et al. (2011), namely "If I heard that a Black British person was upset, and suffering in some

way, I would also feel upset" and "If I saw a Black British person being treated unfairly, I think I would feel angry at the way they were being treated", anchored to 1 = *strongly disagree*, 7 = *strongly agree,* Spearman-Brown reliability = .84.

3.2.3.4. Anxiety

Intergroup anxiety was measured in line with work by Turner, Hewstone and Voci (2007) by asking participants to imagine "being put into a university hall of residence where you are living only with African and Caribbean British students" and reporting how they would feel on three five-point semantic differentials: happy-unhappy, worried-pleased and comfortable-tense, α = .87.

3.2.3.5. Prejudice measures

For *affective prejudice*, participants rated how they generally feel towards Black British people on two five-point adjective scales: cold-warm and positive-negative (first reverse-coded, Spearman-Brown reliability = .95). *Cognitive prejudice* was measured with five statements of belief that covered attitudes (Katz & Hass, 1988) and symbolic racism (Henry & Sears, 2002), anchored to 1 = *strongly disagree*, 7 = *strongly agree*, last two reversed, α = .77:

- Discrimination against Blacks is no longer a problem in the UK.
- Irish, Jewish, Indians and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favours.
- One of the biggest problems for a lot of Blacks is their lack of self-respect.
- Black people do not have the same employment opportunities in the UK that Whites do. *(reversed)*
- Over the past few years, Blacks have gotten less than they deserve. *(reversed)*

3.2.3.6. Behavioural intentions

Support for inclusive policies was measured with an original scale (loosely

based on the issues identified by Saucier & Miller, 2003) that asked for agreement with eight policies, anchored to 1 = *strongly disagree,* 7 = *strongly agree,* α = .89:

- Government should use regulation to help minorities get better housing.

- Government should provide additional funding to improve education for minorities.

- Universities should provide additional scholarships for applicants from disadvantaged groups

- Curricula for schools and universities should be required to include more Black and ethnic minority voices (e.g., authors and scholars).

- The next person to appear on a new bank note should be from an ethnic minority.

- Government should punish discrimination in the labour market more strictly.

- Government should focus more on addressing hate speech against minorities.

- Government should do more to protect minorities from police violence and abuse.

Bystander intervention intentions were measured with the following vignette that participants were asked to imagine:

> *You are at a house party hosted by someone in your course*
> *whom you don't know very well. You are just hanging out with a*
> *friend when there is a knock on the door. Someone you think you*
> *recognize but have not spoken to sticks his head out of the window*
> *and says loudly: "Don't let them in, it's just a load of Black kids."*

They were then asked how likely they were to show six possible reactions, adapted from Dickter, Kittel, & Gyurovski (2012): "Tell your conversation partner that you are not happy about that comment," "Tell the guy that his comment may be understood as offensive and that he might want to be more careful," "Tell the guy that you don't want to hear such racist comments and that he should better apologize," "Go to open the door and welcome the new arrivals to the party," "Go to open the door and

apologize for the comment / try to make them feel better," "Tell the host of the party about the incident if she/he hasn't overheard it." (1 = *not at all likely, 7 = very likely,* α = .83).

3.2.4. Procedure

Participants were invited to take part in a study on student life, contact experiences and social attitudes. When asked for their informed consent, they were told that the study specifically concerned behaviour and attitudes towards people from a different background. They then responded individually to a questionnaire, after which a debrief was provided. In two universities, the data were collected with paper questionnaires that students filled in during a lecture, while three universities administered it online as part of their research participation scheme. In four of the five universities, participants could receive partial course credits for their participation. The research adhered to all BPS ethical guidelines and was approved by the Goldsmiths' Psychology Departmental Ethics Committee.

3.2.5. Analytical approach

Analyses were conducted using the *lavaan* package (Rosseel, 2012) in R (R Core Team, 2020). Full-information maximum likelihood estimation was used, as this allows for cases with missing responses to be included, which prevents the loss of power and potential for bias associated with listwise deletion (Shin et al., 2017). Confidence intervals and standard errors for significance tests were bootstrapped with 5,000 resamples to ensure robustness to deviations from multivariate normality. For indirect effects, *p*-values were obtained by using the relationship between significance tests and confidence intervals: I created Monte Carlo confidence intervals of widths varying by 0.01%, and then subtracted the width of the widest confidence interval that did not include zero from 1 (Jorgensen et al., 2019). The significance of differences in indirect effects was assessed by calculating the difference of indirect effects in each

bootstrap sample and constructing a confidence interval on that basis.

3.3. Results

3.3.1. *Missing data and normality*

None of the variables had more than 5% missing data. To test whether this could be assumed to be Missing Completely At Random, the method proposed by Jamshidian and Jalal (2010) and implemented in the *MissMech*-package (Jamshidian et al., 2014) was used. It showed that the data deviated from multivariate normality (Hawkins test, $p < .001$) and that the appropriate non-parametric test of homoscedasticity did not enable us to reject the hypothesis that the data is missing completely at random, $p = .12$. Nevertheless, listwise deletion would lead to a substantial loss in statistical power, so that the use of a full-information maximum likelihood estimator was indicated (Enders & Bandalos, 2001). The deviations from multivariate normality mostly represented substantial skew towards little contact or little prejudice, respectively. Therefore, bootstrapping with 5,000 resamples was used to obtain all confidence intervals and significance tests.

3.3.2. *Descriptive statistics and correlations*

Table 3.1 shows descriptive statistics and correlations. The reliability is reported with Cronbach's alpha for all scales with more than two items; for affective prejudice and intergroup empathy, the Spearman-Brown coefficient is shown instead. All scales exhibited a good level of internal consistency. Apart from the relationship between positive and negative contact, all correlations were significant and pointed in the expected direction, so that the hypothesized mediation models could be tested. It is worth noting that the correlation coefficients tended to be larger for positive than negative contact in almost all cases (though none of these differences was individually statistically significant), and that the frequencies of positive and negative contact were

unrelated.

Table 3.1. *Descriptive statistics and correlations*

Variable	M (SD)	Distributions	Reliability	1	2	3	4	5	6	7	8
1. Pos. contact	4.25 (1.35)		.90								
2. Neg. contact	2.15 (1.21)		.91	-.02 [-0.14, 0.11]							
3. Valuing diversity	5.50 (1.11)		.87	.44 *** [0.31, 0.57]	-.32 *** [-0.47, -0.17]						
4. Empathy	5.82 (1.18)		.84	.34 *** [0.20, 0.49]	-.20 ** [-0.35, -0.05]	.50 *** [0.36, 0.66]					
5. Anxiety	2.45 (0.97)		.87	-.35 *** [-0.49, -0.21]	.30 *** [0.15, 0.47]	-.38 *** [-0.51, -0.25]	-.23 *** [-0.36, -0.10]				
6. Cognitive prejudice	2.78 (1.01)		.77	-.22 *** [-0.35, -0.09]	.22 ** [0.08, 0.38]	-.58 *** [-0.72, -0.45]	-.35 *** [-0.48, -0.22]	.28 *** [0.14, 0.43]			
7. Affective prejudice	1.77 (0.81)		.95	-.38 *** [-0.51, -0.26]	.19 * [0.04, 0.34]	-.46 *** [-0.60, -0.32]	-.50 *** [-0.65, -0.36]	.46 *** [0.33, 0.60]	.35 *** [0.21, 0.50]		
8. Bystander intervention	5.01 (1.22)		.83	.27 *** [0.12, 0.42]	-.17 * [-0.31, -0.04]	.39 *** [0.24, 0.55]	.40 *** [0.26, 0.55]	-.32 *** [-0.46, -0.18]	-.34 *** [-0.47, -0.21]	-.39 *** [-0.53, -0.26]	
9. Policy support	5.32 (1.03)		.89	.36 *** [0.23, 0.48]	-.25 *** [-0.41, -0.11]	.69 *** [0.54, 0.87]	.46 *** [0.34, 0.58]	-.36 *** [-0.50, -0.22]	-.61 *** [-0.77, -0.48]	-.46 *** [-0.62, -0.33]	.42 *** [0.28, 0.56]

Notes. M and SD are used to represent mean and standard deviation, respectively. Reliability is shown with Cronbach's α. Values in square brackets indicate the 95% confidence interval for each correlation. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

3.3.3. Tests of mediation

Figures 3.1 to 3.4 show the mediation models for the four outcomes, while the direct and indirect effects are summarized in Table 3.2. The results were generally in line with expectations: valuing diversity was a significant mediator in three of the four models; it mediated the link between contact and all outcomes *except for* affective prejudice. In line with Hypothesis 2, the indirect effects were particularly strong when it came to the cognitive outcomes, i.e., cognitive prejudice and policy support. Here, the paths through valuing diversity explained at least 71% of the total effect, and a comparison of the confidence intervals shows that these paths were significantly stronger than those through empathy or anxiety. Moreover, the indirect paths through valuing diversity were stronger for the cognitive outcomes than for the affective outcomes.

Empathy, conversely, was also significant in three of the four models; it did *not* mediate the link between contact and cognitive prejudice. Anxiety, finally, was a significant mediator in only two of the models; it did *not* mediate the links between contact and both cognitive outcomes (cognitive prejudice and policy support).

Collectively, the three mediators could explain most of the relationships between contact and each outcome; none of the direct effects were significant and many estimates were very close to zero. The pattern of results was similar for positive and negative contact, however, in most models the coefficients for negative contact were slightly smaller.

Figure 3.1. *Mediation model for **affective prejudice***
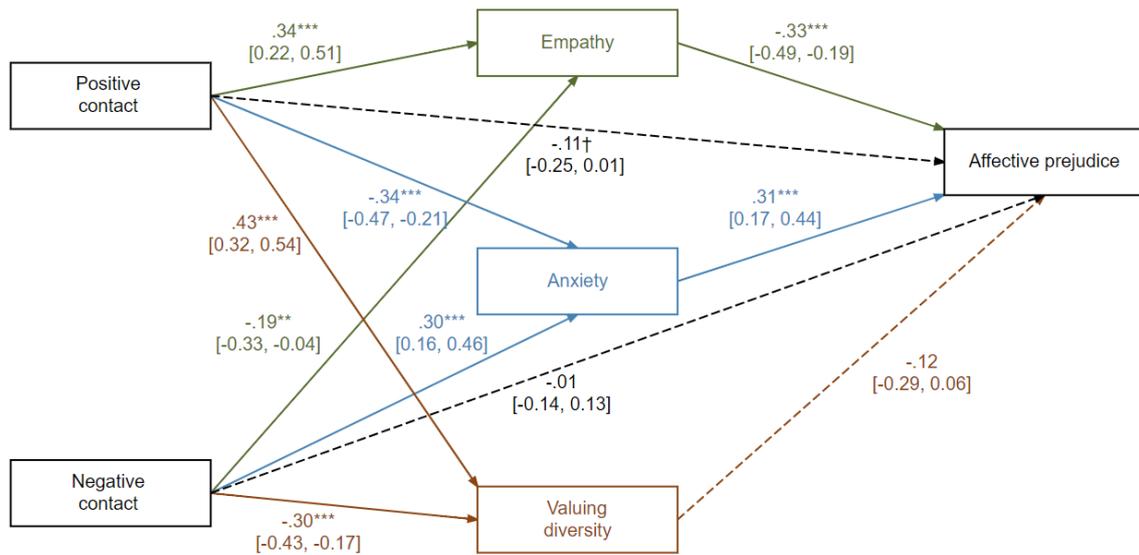
Notes: Standardized coefficients, estimated with maximum likelihood. Confidence intervals based on 5,000 bootstrap samples. Paths with $p > .05$ are dashed. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Figure 3.2. *Mediation model for **cognitive prejudice***
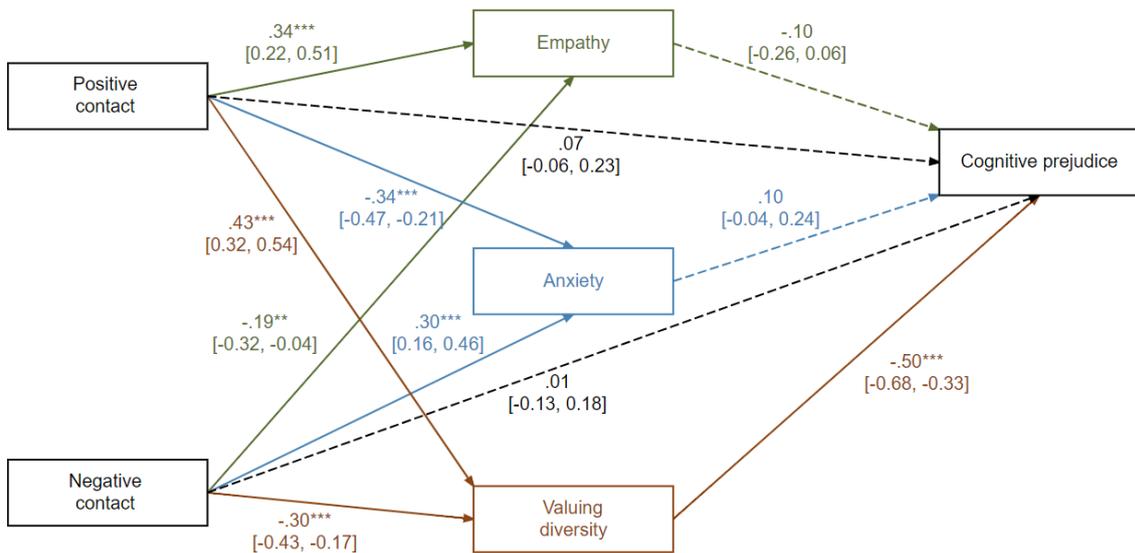
Notes: Standardized coefficients, estimated with maximum likelihood. Confidence intervals based on 5,000 bootstrap samples. Paths with $p > .05$ are dashed. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Figure 3.3. *Mediation model for **bystander intervention***
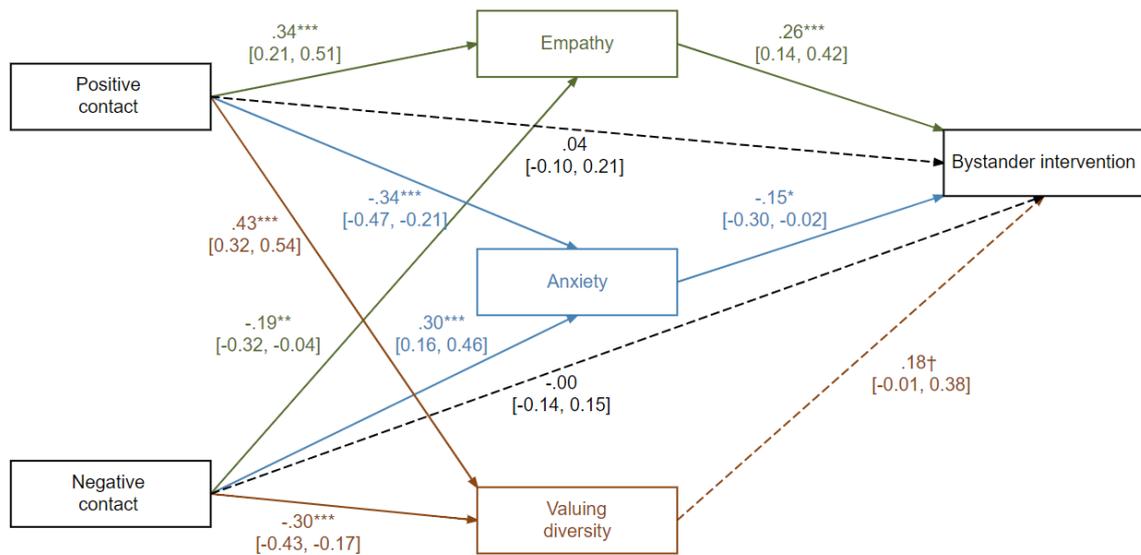
Notes: Standardized coefficients, estimated with maximum likelihood. Confidence intervals based on 5,000 bootstrap samples. Paths with $p > .05$ are dashed. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Figure 3.4. *Mediation model for **policy support***
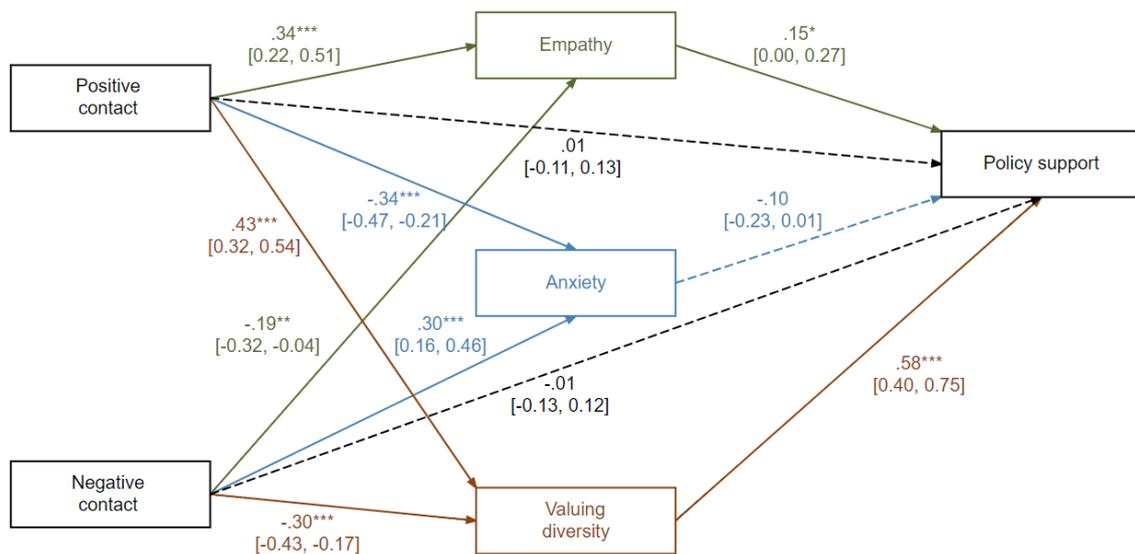
Notes: Standardized coefficients, estimated with maximum likelihood. Confidence intervals based on 5,000 bootstrap samples. Paths with $p > .05$ are dashed. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3.2. *Direct and indirect paths from contact to various potential outcomes*

| To | From | Direct | Indirect through … | | | Total |
			Empathy	*Anxiety*	*Val. diversity*	
Affective prejudice	Pos. contact	-.11 † [-0.25, 0.01]	**-.11** [-0.20, -0.06]	**-.11** [-0.18, -0.05]	-.05 [-0.13, 0.03]	-.38 *** [-0.52, -0.26]
	Neg. contact	-.01 [-0.14, 0.13]	**.06** [0.02, 0.13]	**.09** [0.04, 0.18]	.04 [-0.01, 0.10]	.18 * [0.04, 0.34]
Cognitive prejudice	Pos. contact	.07 [-0.06, 0.23]	-.03 [-0.09, 0.01]	-.03 [-0.09, 0.01]	**-.22** [-0.32, -0.14]	-.21 ** [-0.33, -0.08]
	Neg. contact	.01 [-0.13, 0.18]	.02 [-0.01, 0.06]	.03 [-0.01, 0.09]	**.15** [0.07, 0.25]	.21 ** [0.07, 0.37]
Bystander intervention	Pos. contact	.04 [-0.10, 0.21]	**.09** [0.05, 0.17]	**.05** [0.01, 0.11]	.08 [-0.00, 0.17]	.26 *** [0.13, 0.43]
	Neg. contact	-.00 [-0.14, 0.15]	**-.05** [-0.11, -0.02]	**-.05** [-0.11, -0.01]	**-.05** [-0.14, -0.00]	-.15 * [-0.29, -0.02]
Policy support	Pos. contact	.01 [-0.11, 0.13]	**.05** [0.01, 0.10]	.03 [-0.00, 0.09]	**.25** [0.17, 0.35]	.35 *** [0.23, 0.47]
	Neg. contact	-.01 [-0.13, 0.12]	**-.03** [-0.07, -0.00]	**-.03** [-0.08, -0.00]	**-.17** [-0.29, -0.09]	-.24 *** [-0.38, -0.11]

Notes. Standardized coefficients estimated with full-information maximum likelihood. 95% confidence intervals for indirect effects based on bias-corrected bootstraps with 5,000 resamples, indirect effects with confidence intervals that do not include 0 are **bolded**.

Val. diversity = valuing diversity.

† $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$.

3.4. Discussion

The results are consistent with the hypotheses and suggest that valuing diversity should be considered alongside empathy and anxiety when it comes to the mediation of the effect of intergroup contact. It is of particular importance with regard to cognitive outcomes, i.e., prejudicial beliefs and the support for policies that address inequalities and promote inclusion. Empathy and anxiety, on the other hand, are stronger mediators with regard to affective outcomes, such as affective prejudice and bystander intervention intentions. This stands in contrast with earlier research that concluded that

cognitive mediators are of lower importance compared to affective mediators (Pettigrew & Tropp, 2008). Such research generally did not distinguish between various outcomes of intergroup contact.

3.4.1. *Cognitive and affective prejudice – distinct patterns of mediation*

Prejudice has many dimensions; Tropp and Pettigrew (2005a) suggested a distinction into affective factors (primarily emotions) and cognitive factors (stereotypes and beliefs). Their meta-analysis found that (positive) contact affects both dimensions, even though the relationship between contact and affective outcome measures were stronger than that between contact and cognitive outcomes. Here, I found the same pattern for positive contact, which was more strongly associated with affective than cognitive prejudice. However, this was not the case for negative contact, which had a similarly sized relationship with both types of prejudice.

The mediation analyses highlighted the importance of considering different dimensions of prejudice. While the association of contact with affective prejudice could be adequately explained through empathy and anxiety, valuing diversity became an important mediator once the cognitive dimension of prejudice was considered, and neither of the other mediators reached statistical significance. The same held true for behavioural intentions: while the relationship of contact with bystander intervention intentions was mostly mediated by empathy and anxiety (with a smaller but significant contribution from valuing diversity), valuing diversity became the strongest mediator when the links between contact and the support for inclusive policies was considered. Given that attitude and behaviour changes in both the cognitive and affective domain are needed to bring about a more integrated and equitable society, the results presented here suggest that valuing diversity is a mediator worthy of further study. Conversely, the absence of significant mediation through anxiety in all four models with cognitive outcomes suggests that the reach of this well-established mediator might be more

restricted than is often assumed.

The relatively weak contribution of anxiety stands in contrast to the plethora of previous research reviewed above that has identified it as an important predictor. While the confidence intervals around the effects are large, so that the estimated size of indirect effects should not be over-interpreted, this pattern deserves further attention. Anxiety was relatively closely associated with valuing diversity ($r = -.38$). Conceptually, that is somewhat unexpected, given that anxiety is an affective construct focused on concrete interactions, while valuing diversity is a cognitive measure focused more at the group level. However, it might hint at the importance of considering intergroup threat when it comes to mediation. Anxiety would appear to be closely related to the perception of realistic threat, while valuing diversity should be closely related to the absence of symbolic threat. While these dimensions are theoretically distinct (Stephan et al., 1998), they tend to load onto a single factor in empirical research (e.g., Guerra et al., 2020; Kanas et al., 2017). Therefore, future research might explore the relative contributions of anxiety, threat and valuing diversity to further clarify these findings.

3.4.2. *Negative and positive contact – weak asymmetries*

Though of limited relevance to the argument of this dissertation, the relative strength of effects of positive and negative contact deserves brief consideration, given the practical importance of any asymmetry and the current research interest in the topic. Previous research has found that negative contact has a stronger effect on prejudice than positive contact does, which initially led Barlow and colleagues to propose a general positive-negative contact asymmetry (Barlow et al., 2012). While this continues to be reported, the evidence is mixed, with some studies replicating the original asymmetry (Graf et al., 2014; Hayward et al., 2017; Landmann et al., 2019), while others find positive contact to be a stronger predictor than negative contact (Bagci & Turnuklu,

2018; Meleady & Forder, 2019; Visintin et al., 2017; Wölfer et al., 2017). In that context, it is noteworthy that in these results, with entirely unrelated measures of positive and negative contact ($r = -.02$), the strength of the relationships between both types of contact, the mediators and the possible outcomes were similar, with no significant differences between any of the pairs of regression or correlation coefficients. However, regarding bystander intervention and affective prejudice, the estimated relationships between positive contact and the other variables were consistently larger than those for negative contact.

That negative contact has a relatively stronger relationship with cognitive prejudice might be explained by the finding that it serves to make intergroup boundaries salient (Paolini et al., 2014). While group salience has generally been suggested to enhance contact effects (e.g., Voci & Hewstone, 2003), as a cognitive moderator, it would likely have a particularly strong influence when it comes to cognitive outcomes. The relatively stronger relationship between positive contact and affective outcomes, conversely, might (in part) be due to the reciprocal relationship between contact and attitudes (Binder et al., 2009). If people are led by their positive attitudes to seek out contact, this would likely primarily lead to an increase in positive contact. While I am not aware of research that has contrasted affective and cognitive prejudice as predictors of approach intentions, it appears plausible that they would be more immediately shaped by affect towards individuals rather than beliefs regarding groups. These speculations could serve as avenues towards future research.

3.4.3. Limitations and future directions

Evidently, mediation models based on cross-sectional data can never establish causality. Nevertheless, if the indirect paths were not significant and substantial, the question of causality would become moot. Here, the model is consistent with a hypothesis according to which positive contact increases intergroup empathy, increases

the valuing of diversity, and decreases intergroup anxiety, which in turn leads to reduced prejudice and greater intentions to engage in inclusive behaviours, while negative contact could have the opposite effects.

Apart from the question of causality, the conclusions are limited by the fact that I did not consider behaviours but rather behavioural intentions, mostly because ecologically valid behaviours cannot be measured with paper-and-pencil questionnaires. While intentions and behaviours are clearly not the same, the link between them has been shown to be nearly twice as strong as the link between attitudes and discriminatory behaviours (Schütz & Six, 1996). Thus, I expect that the findings regarding behavioural intentions would hold regarding behaviours, yet this is also in need of future validation.

3.5. Conclusions and connections

This study indicates that valuing diversity might mediate the link between positive and negative contact and a range of cognitive and affective outcomes. In that, it adds nuance to earlier work that suggested a dominance of affective factors and highlights the importance of considering different potential outcomes of contact separately. Empathy and anxiety, the most established mediators, appear to matter most when it comes to affective outcomes, while valuing diversity presents a potentially stronger route from contact to cognitive outcomes.

This chapter builds on the previous chapter, which established that there is a path from contact to valuing diversity, which is likely to be stronger than the reverse path (which was not significant in that dataset). It shows that the effect of contact on valuing diversity is capable of explaining other effects of intergroup contact. However, so far, results are limited by their reliance on convenience student samples in a single intergroup context. Therefore, the next chapter will provide a replication and extension of the mediation analysis in a different intergroup context, using a large-scale

probability sample.

CHAPTER 4:

Further evidence for the role of diversity beliefs
as a mediator of intergroup contact

The previous studies in this dissertation showed that intergroup contact shapes the valuing of diversity, and that, consequently, changes in valuing diversity might explain some of the effects of intergroup contact. This study uses a random population sample from the German General Social Survey (ALLBUS) 2016 (GESIS, 2017) to test whether the mediation results replicate in a different context and in a larger, representative sample. ALLBUS again uses a different measure of valuing diversity, which allowed me to test whether the results are robust to different operationalizations of the construct. In addition, the dataset included a unique outcome measure – the choice of potential neighbourhoods to live in – which offered a chance to consider an outcome critical to social integration, and one where reverse causation is less likely than it might have been in the previous chapter.

4.1.1. *Neighbourhood segregation – the importance of preferences*

Across countries, neighbourhoods are often segregated along ethnic lines, as part of the general tendency for 'birds of a feather' to flock together (McPherson et al., 2001). This can occur in the absence of strong sentiments against the outgroup, as Schelling (1971) showed in his classic simulation paper (cf. J. Zhang, 2004, for a recent discussion and extension). Small preferences to live among a substantial share of ingroup members can lead to a neighbourhood 'tipping' into a high level of segregation. In terms of Schelling's model, incompatible preferences can most obviously lead to a dynamic towards segregation, for instance when Whites want to live in neighbourhoods that are at least three quarters White and minorities in neighbourhoods that are at least a third minority. In these cases, preferences can only be satisfied in homogeneous

neighbourhoods. However, even compatible preferences can imply a narrow range of acceptable compositions. Even if those are achieved temporarily, such preferences can lead to strong segregation dynamics given that a perceived development *towards* a neighbourhood becoming undesirable in terms of its ethnic make-up might lead to moves that accelerate such development past the tipping point, for instance due to a fear that house prices will decline once the neighbourhood has become undesirable.

Of course, such basic models that ignore individual differences in preferences cannot hope to predict real-world dynamics. Nevertheless, they highlight that individual preferences matter and that it is important that they are broad, in the sense that individuals should be open to a range of neighbourhood compositions, as narrow preferences highlight the risk that small changes can lead to the onset of 'White flight' that entrenches segregation. Such transitions have been well-documented in in the United States (Boustan, 2010; Crowder & South, 2008; Pais et al., 2009; Wurdock, 1981) and in South Africa (Durrheim, 2005); most research has found that this has been motivated by a combination of racial attitudes and pull factors that made suburbia more attractive (Krysan, 2002). Recent research has highlighted that while White people's willingness to live in diversity has increased in the United States over the past decades, diversity has increased faster, so that White's preferences continue to contribute to high levels of segregation (Howell & Emerson, 2018)

4.1.2. Neighbourhood segregation in Europe

The US and South Africa are special cases in that spatial segregation was encouraged or even enforced by law until relatively recently. This has not been the case in Europe. Nevertheless, in the UK, almost 90% of minority-ethnic respondents report to live in majority-minority wards, while 80% of Whites estimated that the majority of people living in walking distance from them are also White (Uslaner, 2010). In the Netherlands, natives or Western immigrants were more likely to move out of diverse

neighbourhoods than minority residents, and were more likely to move to 'White' areas (Bolt et al., 2008). In Denmark, Sweden, Belgium and the Netherlands, patterns of segregation are similar, with 50% of the population living in micro-neighbourhoods where the share of non-European immigrants is less than half of what would be expected under an even distribution (E. K. Andersson et al., 2018). In Germany, lastly, neighbourhood ethnic segregation is comparatively low, with very few minority-majority neighbourhoods (Schönwälder & Söhn, 2009); however, when considering the level of micro-neighbourhoods, for instance blocks of flats, it has been shown that immigrants groups are also much more likely to live among their compatriots than what would be expected by chance alone, even when controlling for a wide range of socio-economic factors (Sager, 2012).

Neighbourhood segregation matters because it has been shown to predict a decline in social trust (Uslaner, 2010) and makes it unlikely that the beneficial effects of intergroup contact can be realised (Pettigrew et al., 2010). On a broader level, it also increases political polarisation and entrenches inequality as it makes it less likely that intergroup coalitions can form (Massey & Denton, 1993). Previous studies based on the US General Social Survey have shown that racial prejudice predicts neighbourhood choices. Using data from 1990, Cutler et al. (1999) found that Whites who lived in segregated areas reported a lower willingness to live in a neighbourhood where half of their neighbours would be black. This was replicated ten years later, with data from the US General Social Survey 2000, when Uslaner (2010) showed that White's desired degree of ethnic diversity in their neighbourhood was related to negative stereotypes about African Americans. This study set out to test a broader predictor of neighbourhood choice that might be less entrenched: the general valuing of (cultural) diversity.

Additionally, earlier research has suggested that personal contact with Black

people increases White's willingness to live in diverse neighbourhoods (Ihlanfeldt & Scafidi, 2004). This study sets out to replicate the relationship and to test whether it is mediated by diversity beliefs.

4.1.3. Hypotheses

This study primarily aimed to test three hypotheses:

H1. Valuing diversity predicts participants' intergroup approach intentions (operationalized by the choice of diverse neighbourhoods as potential places to live).

H2. Valuing diversity mediates the relationship between both positive and negative contact and intergroup approach intentions.

H3. This mediation occurs in parallel with that through attitudes towards foreigners, which will also mediate some of the relationship between contact and approach intentions.

As an exploratory addition, I also tested the following hypothesis:

H4. Choosing a diverse neighbourhood is more strongly predicted by the level of positive beliefs about diversity, while avoiding a diverse neighbourhood is more strongly predicted by the level of negative beliefs.

4.2. Methods

4.2.1. Dataset

This study is based on data from the German General Social Survey (ALLBUS) 2016 (GESIS, 2017). ALLBUS employs a random cluster sampling approach of residents of Germany above the age of 18 that covered 162 sample points and a total of 3,490 respondents in 2016, with a purposive oversampling of respondents from East Germany. The data was collected through computer assisted personal interviewing

(CAPI) during the summer of 2016.

Only German citizens were asked about their intergroup contact experiences, thus only they were included in the analyses here, which led to the exclusion of 250 participants. Additionally, as a proxy for ethnicity (which is typically not collected in German surveys), only participants who reported that both their parents had been born in Germany (or in formerly German territories in Eastern Europe) were included. If participants reported the place of birth for only one of their parents, only that parent was considered. This led to the exclusion of another 409 participants. Finally, the contact measures I used asked for the frequency of positive and negative contact, which was not requested from participants who reported that they had not had any contact with foreigners, nor could it be meaningfully imputed for them. Therefore, those respondents were excluded, which led to a final 212 cases being dropped. For all analyses in this paper, the responses were weighted in line with the guidance in the variable report. This yielded a (weighted) sample of 2,618 respondents (M_{Age} = 50.8 years, SD = 17.4 years, 49.3% female, 17.7% from East Germany).

4.2.2. Measures

4.2.2.1. Intergroup contact.

Contact was measured with two items that asked how often people had made positive/negative experiences with foreigners (recoded to 1 = *never*, 5 = *very often*). In line with the findings of Studies 1 and 2, these items were only weakly correlated, r = -.27, p < .001, so that they could be treated as independent predictors.

4.2.2.2. Mediator: Valuing diversity.

Beliefs about the value of diversity for society were measured with two items: "A society with a high degree of cultural diversity is better able to tackle new problems" and "It is better for a country if all people belong to a common culture" (reversed). The

Spearman-Brown reliability for a scale consisting of these two items was .59, which is acceptable for a two-item-scale, particularly if it includes reverse-coding (Loewenthal, 2001). However, separate analyses with only one of the items are provided as a robustness check.

4.2.2.3. Control mediator: Attitudes towards foreigners.

In order to ensure that diversity beliefs are not just a proxy for attitudes towards foreigners, I added these as a parallel mediator. They were measured with seven items, including "Foreigners are taking jobs away from Germans", "Foreigners help secure pensions" (reversed), and "The many foreign children in the schools prevent a good education for German children", all measured on a seven-point scale (1 = *do not agree at all* to 7 = *agree entirely*). These items formed a consistent scale, with Cronbach's α of .80.

4.2.2.4. Dependent variable: Approach intentions.

Participants were shown illustrations of 13 neighbourhoods, each made up of 49 house pictograms, that only differed in the share of white and black houses (the neighbourhoods contained between 0 and 48 black houses; cf. Figure 4.1). They were informed that light houses represented Germans, while grey houses represented foreigners. Participants were then asked to select all neighbourhoods that they would like to live in. The share of 'minority' houses in the most diverse neigbourhood selected was taken as measuring *approach* intentions.[13]

---

[13] Neighbourhood *avoidance* was also measured: Participants were presented with the remaining neighbourhoods and asked which of those they would *not* like to live in. *Approach* was the preferred outcome variable as it is less likely to be affected by social desirability, especially given the question ordering. However, in a supplementary analysis to test for robustness, the least diverse neighbourhood avoided above the most diverse neighbourhood approached was taken to express *avoidance* intentions; the findings were substantially the same (see Appendix 4.1).

Figure 4.1: *Illustrations used to measure approach intentions*
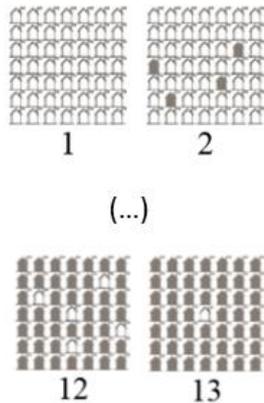
4.2.2.5. Covariates.

Age, gender and participants' level of education were included as demographic covariates, the latter coded following the ISCED 97 classification, from *1 = basic education* to *6 = tertiary education, second stage*. Due to their established association with diversity beliefs, participants' placement on a political ideology scale (1 = *strongly on the left* to 10 = *strongly on the right*) and their region of origin (West Germany or East Germany)[14] were controlled for.

Finally, the share of foreigners in the current neighbourhood was controlled for. Here, ALLBUS conducted a survey experiment, asking half of the respondents to report the percentage freely while providing four categories to the other half (1 = *(Almost) no foreigners* to 4 = *Mostly foreigners*). To combine these variables, the percentage responses were ranked and split into four categories in line with the proportions of answers given on the categorical question.

---

[14] East Germany, the formerly socialist part, has a much more homogeneous population than West Germany, with 4.4% of the population holding a foreign citizenship, compared to 12.6% in the West. In line with that, anti-foreigner attitudes and support for right-wing populism is more widespread in the East (Wallrich et al., 2020a).

4.2.3. Missing data

7.5% of cases had missing data on at least one of the variables under consideration. Political orientation was missing most frequently, at 3.4%, all other variables were provided by at least 98.2% of participants. Nevertheless, listwise deletion would lead to a substantial loss of power and risk introducing bias. As full-information maximum-likelihood estimation is not supported by the *lavaan* package when survey weights are used, multiple imputation was used to retain all available information. This procedure creates multiple datasets in which missing values are differently imputed based on the observed data, thereby accounting for the uncertainty introduced by missing data (Rubin, 2004). Each dataset is then analysed separately, before the results are aggregated to obtain the final parameter estimates and significance tests. The imputation was carried out using the *mice* package (Buuren & Groothuis-Oudshoorn, 2010) in R (R Core Team, 2020), which uses a predictive mean modelling approach that is robust to possible non-normality. In line with guidance by White et al. (2011), eight imputations were used.

After multiple imputation, the analyses were performed with the *lavaan.survey* R package (Oberski, 2014). Bootstrapping with survey weights is complex (Stapleton, 2008) and not yet implemented in *lavaan.survey*, while Monte-Carlo simulations have been shown to result in reliable confidence intervals for indirect effects in such situations (D. P. MacKinnon et al., 2004; Preacher & Selig, 2012). Therefore, tests of indirect effects in this study are based on 20,000 Monte Carlo simulations, using the *semTools* R package (Jorgensen et al., 2019).

To consider the robustness of the findings with regard to choices about the treatment of missing data, the results of the initial regression model are presented both with multiply imputed data and with case-wise deletion; all other analyses are reported only based on the multiply imputed data.

4.3. Results

Descriptive statistics and correlations between the continous variables are shown in Table 4.1. Overall, neighbourhood choices were related to valuing diversity and to all possible covariates with medium to large effect sizes (Cohen, 1988); only the links with negative contact and education were comparatively weak. The measures of contact were associated both with neighbourhood choices and valuing diversity, so that the initial conditions for later tests of mediation were fulfilled.

Table 4.2 shows the distribution of the categorical variables and their association with neighbourhood approach. This indicates that neighbourhood choices were unrelated to gender, but that respondents in West Germany and those who had at least some foreigners in their current neighbourhood selected more diverse neighbourhoods. The measures of contact were associated both with neighbourhood choices and valuing diversity, so that the initial conditions for later tests of mediation were fulfilled. Regarding potential asymmetries between positive and negative contact, it should be noted that the correlations between positive contact and its potential outcomes (approach, valuing diversity and attitudes towards foreigners) were larger than those for negative contact. Given that the confidence intervals in Table 4.1 do not overlap, these differences between the correlation coefficients were statistically significant.

.

Table 4.1. *Descriptive statistics and correlations*

Variable	M (SD)	1	2	3	4	5	6	7
1. Approach intentions	0.37 (0.22)							
2. Valuing diversity	2.95 (0.72)	.39 *** [0.36, 0.43]						
3. Attitude towards foreigners	4.36 (1.20)	.53 *** [0.50, 0.55]	.52 *** [0.49, 0.55]					
4. Political Orientation	5.89 (1.69)	.29 *** [0.25, 0.33]	.28 *** [0.24, 0.32]	.33 *** [0.28, 0.37]				
5. Age	50.82 (17.43)	-.35 *** [-0.39, -0.31]	-.07 *** [-0.11, -0.03]	-.12 *** [-0.16, -0.08]	-.05 * [-0.10, 0.01]			
6. Education	3.82 (1.08)	.17 *** [0.13, 0.21]	.14 *** [0.10, 0.18]	.19 *** [0.15, 0.22]	.09 *** [0.04, 0.14]	.06 ** [0.02, 0.09]		
7. Positive contact	3.95 (0.97)	.37 *** [0.33, 0.40]	.34 *** [0.30, 0.37]	.42 *** [0.38, 0.45]	.21 *** [0.17, 0.25]	-.11 *** [-0.15, -0.07]	.19 *** [0.15, 0.23]	
8. Negative contact	2.16 (0.99)	-.11 *** [-0.15, -0.06]	-.17 *** [-0.21, -0.13]	-.27 *** [-0.31, -0.24]	-.17 *** [-0.23, -0.11]	-.18 *** [-0.22, -0.14]	-.05 * [-0.09, -0.01]	-.27 *** [-0.31, -0.23]

*Note*s: *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. Correlations are based on multiple imputation of missing data, with 8 imputations. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 4.2. *Descriptive statistics and associations for categorical variables*

	N	Share	*M (SD)* approach intentions
Gender			
female	1291	49.3%	0.38 (0.21) [a]
male	1327	50.7%	0.37 (0.22) [a]
Region			
East Germany	463	17.7%	0.30 (0.21) [b]
West Germany	2155	82.3%	0.39 (0.22) [a]
Neighbourhood diversity			
(Almost) no foreigners	813	31.0%	0.33 (0.20) [b]
Some foreigners	1466	56.0%	0.39 (0.22) [a]
Many foreigners	302	11.5%	0.41 (0.24) [a]
Mostly foreigners	37	1.4%	0.35 (0.26) [ab]

*Note*s: *M* and *SD* are used to represent mean and standard deviation for approach intentions for that group, respectively. Within each variable, the means of groups with different superscripts differ with $p < .05$ (*p*-values were adjusted using the Holm-method.)

4.3.1. Regression models

Next, I used multiple OLS regression to assess the unique contribution of valuing diversity and thereby test Hypothesis 1. The results are shown in Table 4.3, both based on the multiply imputed dataset and based on case deletion. There are no substantial differences between any of the coefficients, which indicates that the results of this study do not depend on choices regarding the treatment of missing data. Valuing diversity explained a significant unique part of the variance in neighbourhood choices after controlling for a wide range of covariates, with an effect size similar to that of political orientation. Attitudes towards foreigners were more strongly associated with neighbourhood choice.

Table 4.3. *OLS regression models to predict neighbourhood approach intentions*

	Multiple imputation		Listwise deletion	
	B (SE)	*β [95% CI]*	*B (SE)*	*β [95% CI]*
(Intercept)	0.00 (0.02)	0.08 [0.03, 0.12]	0.00 (0.02)	0.08 [0.04, 0.13]
Valuing diversity	0.04 (0.01)***	0.14 [0.11, 0.18]	0.04 (0.01)***	0.15 [0.11, 0.18]
Attitude towards foreigners	0.06 (0.00)***	0.36 [0.32, 0.39]	0.06 (0.00)***	0.36 [0.32, 0.40]
Political orientation (right-wing)	0.02 (0.00)***	0.12 [0.09, 0.16]	0.02 (0.00)***	0.12 [0.09, 0.15]
Age	-0.00 (0.00)***	-0.29 [-0.32, -0.26]	-0.00 (0.00)***	-0.29 [-0.32, -0.26]
Gender (female)	0.00 (0.01)	0.01 [-0.05, 0.07]	0.00 (0.01)	0.01 [-0.05, 0.07]
Region (East)	-0.06 (0.01)***	-0.26 [-0.34, -0.18]	-0.06 (0.01)***	-0.27 [-0.35, -0.19]
Education	0.02 (0.00)***	0.09 [0.06, 0.12]	0.02 (0.00)***	0.09 [0.06, 0.12]
N	2615		2472	
R^2	.41		.41	
F-tests	$F(7, 2607) = 257.48, p < .001$		$F(7, 2464) = 249.79, p < .001$	

Note: Given that dummy variables lose their interpretability when standardised (Fox, 2015), *β* for region and gender are semi-standardised, indicating the impact of that level on the standardized outcome variable.
† $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

4.3.2. Test of the mediation

I estimated a mediation model to test whether there were indirect paths from the contact measures through valuing diversity to approach intentions, in addition to the established path through attitudes towards foreigners, controlling for demographic covariates and political orientation in both the mediators and outcome variables. The resulting model is shown in Figure 4.2, while the coefficients for the direct and indirect paths are shown in Table 4.4. The model indicated that positive and negative contact had significant indirect effects on neighbourhood choice through both mediators, while the direct effect was only significant for positive contact. Overall, positive contact had a much stronger effect on approach intentions than negative contact did. Likewise, the mediation through valuing diversity was stronger for positive than negative contact, both in terms of the size of the indirect effect and its share of the total effect, because positive contact was more closely associated with valuing diversity than negative contact was.

Figure 4.2. *Mediation of the effect of intergroup contact on neighborhood choice*
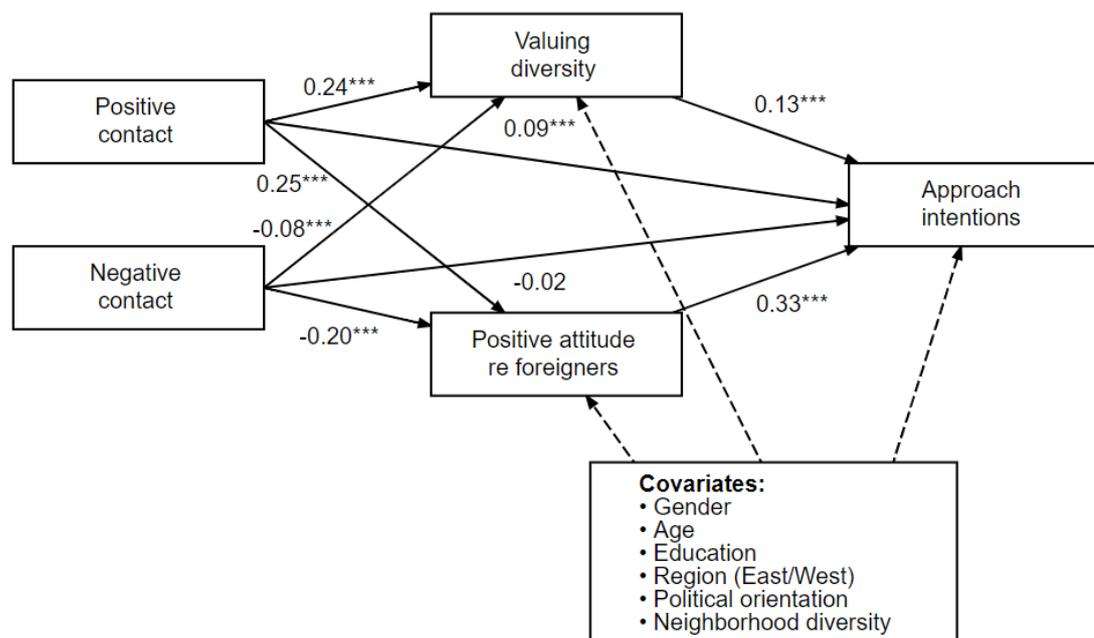
Note: Standardized coefficients, estimated with maximum likelihood on multiply-imputed data.
  † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 4.4. *Mediation of the effect of intergroup contact on neighborhood choice*

Predictor	**Paths** (standardized coefficients)			
	Direct	Indirect through		Total
		Foreigner attitudes	*Valuing diversity*	
Positive contact	0.09 *** [0.06, 0.13]	**0.08** [0.07, 0.10]	**0.03** [0.02, 0.04]	0.21 *** [0.17, 0.25]
Negative contact	-0.02 [-0.05, 0.02]	**-0.07** [-0.08, -0.05]	**-0.01** [-0.02, -0.00]	-0.09 *** [-0.13, -0.05]

Note: Values in square brackets indicate the 95% confidence intervals based on 20,000 Monte Carlo simulations. Indirect effects with confidence intervals that do not cross 0 are **bolded**.
    † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

4.3.3. *Exploratory analyses*

In the survey, participants were asked to select neighbourhoods to live in and then given the opportunity to explicitly reject any of the remaining ones. They were also asked about positive and negative diversity beliefs, i.e., whether a diverse country would be better equipped to tackle new problems and whether it would be better for a country if all people belong to a common culture. As an exploratory hypothesis (H4), I tested whether positive beliefs primarily predict selection, while negative beliefs predict rejection. Additionally, to check robustness, I considered whether rejection, as a related alternative outcome measure, is also predicted by valuing diversity. Table 4.5 shows OLS regression models for selection and rejection with the two separate predictors. Overall, the exploratory hypothesis is not supported – both items are associated with rejection and selection with very similar standardised coefficients. However, the additional model gives reason to be confident in the main findings: while some demographic variables are less strongly associated with rejection, valuing diversity remains a key predictor when selection is replaced by rejection as the outcome variable.

Table 4.5. *OLS regression models for neighbourhood selection and rejection*

	Neighbourhood selection		Neighbourhood rejection	
	B (SE)	*β [95% CI]*	*B (SE)*	*β [95% CI]*
(Intercept)	0.09 (0.03)***	0.08 [0.03, 0.12]	0.63 (0.04)***	-0.01 [-0.06, 0.04]
Valuing diversity *(negative)*	0.02 (0.00)***	0.09 [0.05, 0.12]	-0.02 (0.01)***	-0.08 [-0.12, -0.04]
Valuing diversity *(positive)*	-0.02 (0.00)***	-0.08 [-0.11, -0.04]	0.02 (0.01)***	0.07 [0.03, 0.11]
Pos. attitude towards foreigners	0.06 (0.00)***	0.36 [0.32, 0.39]	-0.06 (0.00)***	-0.29 [-0.33, -0.25]
Political orientation *(right-wing)*	0.02 (0.00)***	0.12 [0.09, 0.16]	-0.01 (0.00)***	-0.07 [-0.11, -0.04]
Age	-0.00 (0.00)***	-0.29 [-0.32, -0.26]	0.00 (0.00)***	0.09 [0.06, 0.13]
Gender *(female)*	0.00 (0.01)	0.02 [-0.04, 0.08]	-0.00 (0.01)	-0.01 [-0.08, 0.06]
Region *(East)*	-0.06 (0.01)***	-0.26 [-0.34, -0.18]	0.01 (0.01)	0.04 [-0.05, 0.13]
Education	0.02 (0.00)***	0.09 [0.06, 0.12]	-0.01 (0.00)†	-0.03 [-0.07, 0.00]
N	2615		2615	
R²	.41		.19	
F-tests	$F(8, 2606) = 223.63, p < .001$		$F(8, 2606) = 72.52, p < .001$	

Note: Given that dummy variables lose their interpretability when standardised (Fox, 2015), *β* for region and gender are semi-standardised, indicating the impact of that level on the standardized outcome variable.
† $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

4.3.4. Robustness check: logistic regression

The choice of neighbourhood might be construed as an ordinal rather than continuous variable, thereby violating a key assumption of linear models. Therefore, a proportional odds logistic regression model was run on the multiply imputed data using the *MASS* package in R (Venables & Ripley, 2002); the pseudo R^2 value was obtained using the *pscl* package (Jackman, 2020).

There were no substantial differences in the results, as can be seen in Appendix 4.1. Again, all predictors, except for gender, were significant; the mean log-likelihood pseudo-R^2 across the 21 imputations was .43. For each standard

deviation increase in valuing diversity, respondents were 44% more likely to choose the next more diverse neighbourhood than those who valued diversity less.

4.4. Discussion

In line with my hypotheses, valuing diversity was associated with the choice of more diverse neighbourhoods as potential places to live, and helped explain the association of positive and negative contact with these approach intentions. The observed data was consistent with a model in which valuing diversity mediated the relationships of positive contact and negative contact with approach intentions, in parallel with the mediation through attitudes towards foreigners and after controlling for a wide range of covariates. It is worth noting that the estimated indirect effects are small. However, small effects can have large social consequences when they shape many decisions that affect many people (Greenwald et al., 2015), as is arguably the case with decisions regarding everyday segregation.

4.4.1. Do attitudes predict actual neighbourhood choices?

There is some evidence that attitudes towards outgroups – while related to approach intention – do not substantially predict the likelihood of leaving one's neighbourhood. For instance, Kaufmann and Harris (2015), using longitudinal data over a period of 20 years, did not find evidence of 'white flight' from diverse areas. However, in line with the focus of this study, they found some evidence for White's avoidance of diverse areas when moving. Similarly, Andersen (2017) showed that in Denmark, White avoidance of diverse neighbourhoods made a greater contribution to residential segregation than White flight. Similar effects of White avoidance of diverse neighbourhoods have also been found in Sweden (R. Andersson, 2013; Bråmå, 2006). Therefore, the results presented here suggest that positive intergroup contact – through valuing diversity – might contribute to White people's choice to

move into diverse neighbourhoods, thereby contributing to social intregation, while negative contact might have the reverse effect.

4.4.2. Strengths and limitations

The study is based on a large representative sample, which shows that the results are not restricted to liberal student samples. It is also based on a long multi-topic questionnaire, with items used here asked at different points, which reduces the risk that findings are due to illusory correlations and/or acquiescence.

Also, in the stimuli, diversity was presented as binary; the houses were either black or white. This is a important as it has been shown that diversity is more likely to be stable when multiple minorities are present, while White flight has been predicted particularly strongly be the presence of a single minority (Reibel & Regelson, 2011). Thus the situation presented represents a less palatable version of diversity, which makes the results particularly important while potentially placing a limit on their generalisability.

As in the previous chapter, the dataset used is cross-sectional, which does not allow for the drawing of the causal conclusions that are implied by a mediation model. Nevertheless, the fact that the data are consistent with such a (theoretically plausible) model is an important finding in the research process.

4.5. Conclusions and connections

Using a large representative sample and an outcome measure highly relevant to social integration, I have shown that valuing diversity is associated with a tendency to make inclusive (neighbourhood) choices. In addition, this chapter offers a conceptual replication of the previous chapter, showing that valuing diversity mediates an effect of intergroup contact, using a random population sample with different measures and including a wide range of control variables.

Moreover, it links with Chapter 2 in finding that valuing diversity predicts a measure of intergroup approach intentions. This adds some further credence to the idea that the relationship between contact and valuing diversity might be dynamic and mutually reinforcing.

Given the apparent benefits of valuing diversity in terms of a wide range of outcomes associated with intergroup contact, the following chapters test contact interventions and explore how the link between (positive) contact and valuing diversity can be strengthened, with the expectation that this will in turn contribute to prejudice reduction and behaviour change.

CHAPTER 5:

Contact interventions can increase valuing diversity

Chapter 2 suggested that contact experiences increase the value placed on diversity over time, based on an analysis of longitudinal data. Chapter 3 and 4 proposed that this explains the beneficial effects of positive intergroup contact on prejudice and behavioural intentions. In moving towards an exploration of ways to further enhance the impact of contact on diversity beliefs, I now assess the effect of an intense contact intervention on the valuing of diversity through a comparison of pre- and post-measures. The contact intervention under consideration was the English National Citizen Service (NCS) – a four-week-long government-funded summer programme for 16-year-olds, who are brought together in diverse cohorts and work together on various projects.

5.1.1. *National Citizen Service (NCS)*

The NCS was introduced in 2011 to "promote a more responsive, cohesive and engaged society" (Mycock & Tonge, 2011) and the British government has committed to offer an NCS place to any young person who wants to participate (Wilson, 2016). The programme is run by providers who recruit broadly, yielding a participant population that is roughly representative of English 16-year olds, with an over-representation of ethnic minorities and economically disadvantaged participants (National Audit Office, 2017). Participants are primarily recruited with the promise of adventure and personal growth and sign up with that expectation (Mills & Waite, 2017), which reduces the risk of self-selection based on the attitudes studied here.

The Challenge, this project's main research partner was the largest provider at the time of research, serving more than 50,000 participants each summer.

The exact structure of the NCS varies between providers. In the case of *The Challenge,* it was a four-week journey: the first week focused on outdoor activities that aimed to strengthen confidence and to build the group, the second week emphasised trust, teamwork and leadership, and in the final phase participants designed and implemented a social action project in cooperation with a local charity. Participants in each group were drawn from a specific local area, but then allocated to groups that were as diverse as possible with regard to both their ethnic and socio-economic background.

Evaluations of the NCS have shown moderate effects in increasing social trust, understanding and confidence in one's ability to interact with members of different groups and a commitment to voluntary action, alongside contributions to positive youth development (Cameron et al., 2017). *The Challenge's* own reporting has indicated that their programme allowed participants to develop an appreciation of diversity (The Challenge, 2019), yet these findings were based on participants agreement with a question on their programme experience, where the socially desirable answer was not difficult to guess. This study assesses the impact of the intervention with more rigour.

5.1.2. *Two facets of valuing diversity*

In the studies presented so far, valuing diversity was measured in multiple ways with largely consistent results. However, conceptually, two distinct components were considered: a preference for experiencing diversity, the focus of Chapter 2, and a belief that diversity has instrumental value, which was considered in Chapters 3 and 4. Here, both conceptions are included, so that results can be

contrasted.

5.1.3. *Potential moderators of the relationship between contact and diversity beliefs*

The present study tested whether participation in a contact intervention was associated with an increase in the valuing of diversity. Additionally, it tested two potential moderators that have been proposed to enhance the effects of intergroup contact on related outcomes: conversations about differences and self-expansion orientation. Conversations about differences have been shown to predict greater contact effects on social change motivation (Vezzali et al., 2017) and appear to be an obvious indicator of the salience of distinct group memberships, which has been repeatedly shown to enhance contact effects (Brown & Hewstone, 2005). Therefore, I expected frequent conversations about differences to predict greater increases in valuing diversity over the course of the programme. Self-expansion orientation, the desire for personal growth in the context of social relationships, has been shown to predict both more intimate contact, and greater intergroup approach behaviours (Dys-Steenbergen et al., 2016). Therefore, I expected participants who entered the contact intervention with greater self-expansion orientation to experience greater gains in valuing diversity over its course.

5.1.4. *The present research*

The study was a secondary data analysis of *The Challenge's* impact survey for the summer of 2018, which contained pre- and post-measures on the outcomes of interest. Given that participants self-selected into the programme, it was not feasible to recruit a comparable control group. Nevertheless, it appears very likely that changes over the course of four weeks would be driven by experiences during those four weeks rather than developmental processes of maturation or other external

influences.[15] Four specific hypotheses were tested:

H1: Participation in the NCS will be followed by higher preferences for working in diverse teams.

H2: Participation in the NCS will be followed by higher conviction that diverse teams perform better and thus have instrumental value.

In addition, I tested two hypotheses regarding the moderation of the intervention's effect:

H3: A higher *frequency of conversations with other participants about differences* will moderate the intervention effect, in terms of predicting greater effects on both facets of valuing diversity.

H4: Participants' initial *self-expansion orientation* will moderate the intervention effect, in terms of predicting greater effects on both facets of diversity.

5.2. Methods

5.2.1. Participants and procedure

Our partner organisation sampled participants to complete their impact assessment survey, which was the source for my data. 2,324 participants responded to the start-of-programme survey, while 1,631 participants responded to the end-of-programme survey. Unfortunately, due to issues during data collection, only 705 pairs of responses could be matched for analyses of changes during the programme.

Due to a split survey design unrelated to this research, information on ethnicity was only collected from a randomly selected subset of respondents. Among

---

[15] Of course, changes could have been driven by external events such as news stories that could have influenced the participants. However, given that the participants were recruited from across England, such events would likely have needed to make national news. I could not identify any events that could provide a reasonable alternative explanation from a search of news archives or from conversations with The Challenge programme managers.

those (*N* = 431), 140 were White, 45 were from a mixed and 246 from a minority-ethnic background. This is likely to be representative of the full sample, so that more than half of the participants were from minority-ethnic backgrounds (mostly South Asian, African and Carribbean British).

5.2.2. Measures

Preference for diversity. Participants were asked for their agreement with the following statement: "I prefer being on a diverse team to one where everyone is from the same background as me (regardless of whether I accomplish the team goals or not)" from 1 = *strongly disagree* to 5 = *strongly agree.*

Instrumental value of diversity. Participants were told to "Imagine a new team that needs to deliver a project together at school or college." and asked for their agreement with: "In most cases, diverse teams will produce better results than one where everyone is from the same background", from 1 = *strongly disagree* to 5 = *strongly agree.*

Conversations about difference. Participants were asked, "During NCS, how often did you discuss with other participants the differences in your lives and experiences?", from 1 = *never* to 5 = *very often.*

Self-expansion orientation. Self-expansion orientation was measured with four items derived from Dys-Steenbergen et al. (2016), asking participants to what extent they were looking forward to "meet different people", "challenge your worldview", "revise future goals" and "have new experiences", from 1 = *not at all* to 5 = *very much*. The mean of the items yielded a scale with acceptable internal consistency, Cronbach's α = .79.

5.2.3. Analytic approach

Single-item Likert-type measures with only five response options cannot be treated as interval data, so that linear models are not appropriate. For instance, Bauer and Sterba (2011) conducted a simulation study that suggested that linear models will typically yield biased results in this case unless there are at least seven response categories. Therefore, the responses should either be dichotomized and analysed with logistic regression or maintained as they are and analysed with ordinal models. Since dichotomizing responses results in a loss of power, ordinal models, specifically the proportional-odds logistic regression model, should be preferred (Hedeker, 2015). Therefore, the analyses were conducted by running ordinal mixed-effects models for each outcome, with interaction terms to test the hypothesized effects of conversations about difference and self-expansion orientation. Random intercepts were included to account for repeated measures within participants and for the nesting of participants in the cohorts. Modelling was done using the using the *ordinal*-package (Christensen, 2019) in R (R Core Team, 2020).

5.3. Results

Table 5.1 shows the distribution of responses and their associations for all variables under consideration. Given that they are based on single-item measures using five-point scales, Pearson's *r* is not an appropriate description of their association, so that non-parametric correlations (Spearman's ρ) are shown instead. It is worth noting that the association between preferring diversity and perceiving its instrumental value was stronger at the end than at the start of the programme.

Table 5.1. *Distribution of variables and Spearman correlations*

Variable	M (SD)	Distribution	1	2	3	4	5
1. Diversity value *Start*	3.63 (0.94)						
2. Diversity value *End*	3.63 (0.89)		.33 *** [.25, .40]				
3. Diversity preference *Start*	3.52 (1.02)		.31 *** [.23, .37]	.24 *** [.16, .32]			
4. Diversity preference *End*	3.79 (0.98)		.21 *** [.13, .28]	.50 *** [.43, .55]	.21 *** [.13, .29]		
5. Discussed differences	3.60 (0.92)		.10 * [.02, .18]	.23 *** [.15, .30]	.06 [-.02, .14]	.17 *** [.09, .25]	
6. Self-expansion orientation	3.92 (0.66)		.24 *** [.16, .31]	.13 ** [.04, .21]	.21 *** [.13, .28]	.14 *** [.06, .23]	.15 *** [.07, .24]
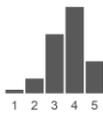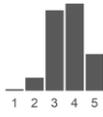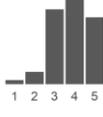
Notes. Values in square brackets indicate the 95% confidence interval for each correlation.
† $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

5.3.1. Changes in diversity beliefs

Proportional-odds mixed-effects models were run to test whether the likelihood of higher levels of agreement with the diversity belief items changed during the programme. Regarding participants' preference to work in diverse teams, there was a significant increase during the programme, $z = 5.06$, $p < .001$. Overall, 36.7% reported a greater preference to work in diverse teams at the end of the programme compared to the start, while 21.3% reported a lower preference. However, their conviction that diverse teams perform better did not change during the programme, $z = -0.64$, $p = .52$, with 25.6% reporting a higher and 30.4% of participants a lower rating at the end.

It should be noted, however, that a comparison of repeated measures can be distorted by regression to the mean. In the present data, changes in both outcome variables were related to the baseline response, indicating regression to the mean, Wilcoxon signed-rank tests for preference $W = 361588$, $p < .001$, for instrumental value $W = 388661$, $p < .001$. Permutation tests can be used to create a null distribution that considers this observed regression to the mean, and thus offer a more reliable test of intervention effects (Furrow, 2019). However, they cannot (straight-forwardly) account for the nesting of participants into cohorts. In the present case, the results from the hierarchical models and the permutation tests concur in showing an intervention effect on participants' preference for diversity, but not their belief in its instrumental value. The results of the permutation test are shown in Appendix 5.1.

5.3.2. *Influence of conversations about differences and self-expansion orientation*

To test whether the observed changes were moderated by measures describing the participants' experience during the programme, two approaches were considered. A common approach is to model this as an interaction between the time of measurement and the variable of interest as proposed by Mangiafico (2016). However, even with the inclusion of random intercepts for each participant, this does not account for potential for regression to the mean. Therefore, if there is evidence for regression to the mean, the baseline response should instead be entered as a covariate in the model that predicts the final response (Barnett et al., 2005). Given that this was the case here, I ran mixed-effects proportional-odds models with diversity beliefs at the end of the programme as the dependent variable, predicted by diversity beliefs at its start and the potential moderator of interest.

5.3.2.1. Conversation about differences

The frequency of conversations with fellow participants about differences in their lives predicted increases in both measures of valuing diversity, i.e., it increased the preference for working in diverse teams, LR $\chi^2(4) = 18.92$, $p < .001$, and the conviction that such teams perform better, LR $\chi^2(4) = 25.55$, $p < .001$. Figure 5.1 shows the association for both measures. Given that conversations about differences had an effect on the post-test score after controlling for the pre-test score for both outcomes, Hypothesis 3 is confirmed: they are a significant moderator of the intervention effect, with more frequent conversations about differences predicting a greater intervention effect.

Figure 5.1.

Relationship between conversations about differences and valuing diversity
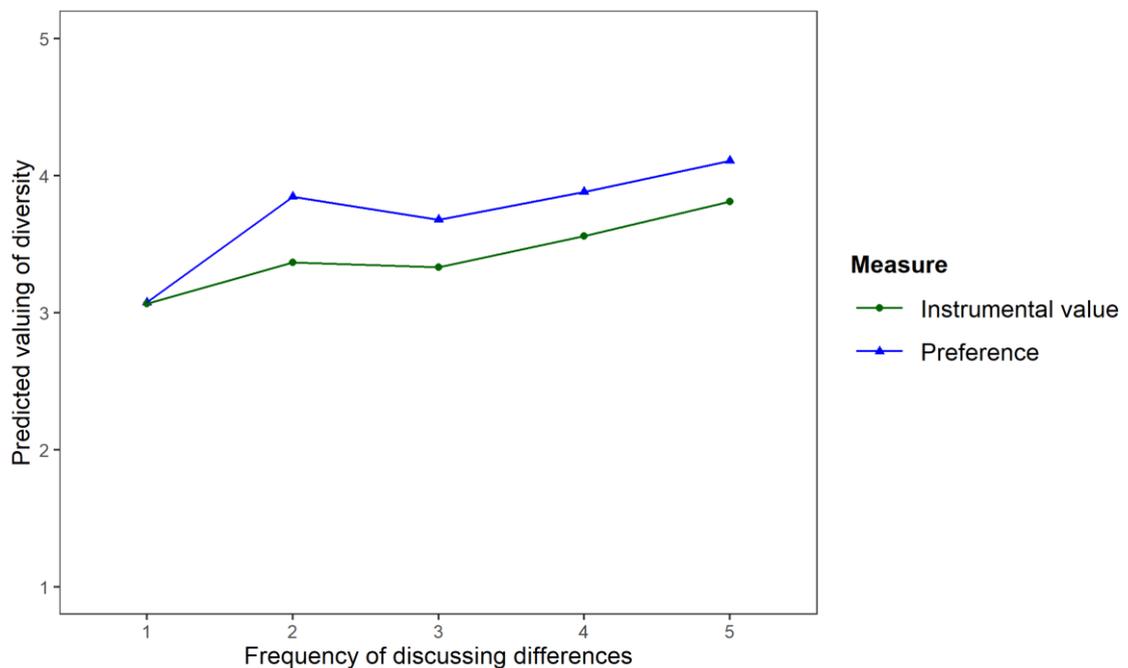
5.3.2.2. Self-expansion orientation

Self-expansion orientation predicted a greater preference for working in diverse teams at the end of the programme, controlling for that preference at the start of the programme, LR $\chi^2(1) = 4.13$, $p = .042$. However, it did not significantly

moderate the effect of the programme on the belief that diverse teams have instrumental value, LR $\chi^2(1) = 1.33$, $p = .25$. Both relationships are shown in Figure 2. Thus, Hypothesis 4 is only partly confirmed; self-expansion orientation only moderates the effect of the programme on a preference for diverse teams, but not on the belief in their instrumental value.

Figure 5.2.

Relationship between self-expansion orientation and diversity beliefs
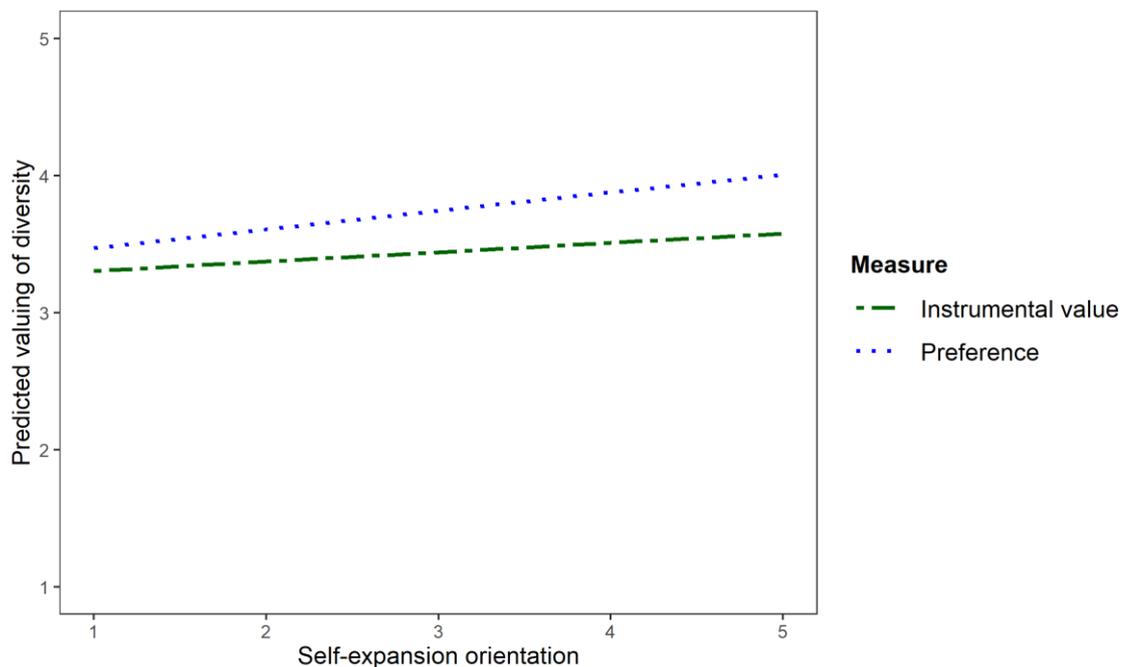
5.3.3. *Differences between majority- and minority-status participants*

Given that this study was the first that included minority-status participants, I wanted to test whether the impact of the intervention on participants' valuing of diversity was moderated by their ethnicity. This was not the case for either outcome measure, preferring diverse teams: LR $\chi^2(1) = 0.57$, $p = .45$, nor recognising their instrumental value: LR $\chi^2(1) = 0.01$, $p = .91$. Furthermore, participants from White and minority-ethnic background did not differ in their valuing of diversity on either measure at the start of the programme, instrumental value: $\Delta M = -0.13$, 95% CI

[-0.34, 0.07], $t(241.5) = 1.26$, $p = .208$; preference: $\Delta M = -0.04$, 95% CI [-0.25,

0.18], $t(269.8) = 0.33$, $p = .741$, equal variances not assumed.

5.4. Discussion

In line with expectations, participating in the NCS was followed by increases

in participants' valuing of diversity, but only if this was operationalised as their

personal preference to work in diverse teams (*H1 supported*). The instrumental belief

that diverse teams produce better results did not change during the programme; in

fact, the observed means were identical at the start and end *(H2 not supported).*

Regarding moderators of the effect, conversations about differences that made

diversity salient predicted greater increases in both aspects of valuing diversity (*H3*

supported). Self-expansion orientation only predicted changes in the preference for

diverse teams but did not significantly predict increases in the belief in their

instrumental value *(H4 partially supported).* Overall, these results support the earlier

finding that intergroup contact experiences can shape the valuing of diversity, but

suggest that the strength of the effect and the mechanisms that bring it about may

differ depending on the facet of valuing diversity under consideration.

5.4.1. Different facets of valuing diversity

Preferences for working in diverse teams appear more malleable and appear

to be influenced by a wider range of contact experiences, while beliefs in the

instrumental value of diversity were only influenced for participants who engaged in

conversations about differences. This might be explained by hypothesising that a

preference for working in diverse teams might be more closely linked to a general

desire to engage in intergroup contact, which has been shown to be an outcome of

contact in previous work (Binder et al., 2009; Emerson et al., 2002). Deeming

diversity to have instrumental value might require a greater cognitive engagement

with the topic of diversity, which only comes about when diversity is explicitly considered in a contact situation. Given that some of the positive effects of valuing diversity established earlier, such as its relationship with policy support and cognitive prejudice, might depend on a recognition of its instrumental value, it is important to design contact interventions in ways that are likely to influence it.

5.4.2. *Moderators of the intervention effect*

Frequent conversations about differences predicted greater increases in both types of valuing diversity. Given that conversations about differences almost necessarily lead to greater category salience, this accords with more general findings regarding the importance of category salience in moderating contact effects (Brown & Hewstone, 2005). Earlier research has also shown that conversations about differences were associated with greater intentions to tackle intergroup inequality (Vezzali et al., 2017). This result might be explained by shifts in valuing diversity, even though that was not measured in the original study. Despite this power of conversations about differences, it should be noted that they do not always benefit participants. Vezzali et al. (2017) found that their positive effect was restricted to adolescents who engaged in repeated positive contact that also covered commonalities between the groups. Similarly, an experiment in the context of the NCS found that scheduled group discussions about differences did not increase social trust on average, compared to a passive control, while scheduled discussions about commonalities increased social trust. However, the conversations about differences increased trust among participants who started at a relatively high baseline (Sanders et al., 2017), which again suggests that conversations about difference might work if they are embedded into positive intergroup experiences.

Self-expansion orientation, by contrast, only predicted greater changes on one

facet of valuing diversity (personal preferences) and, even there, appeared to have less influence than conversations about differences did. That there was an effect on one of the outcomes strengthens earlier findings that self-expansion orientation can positively influence the experience of intergroup contact and strengthen its outcomes (Dys-Steenbergen et al., 2016), so that interventions that boost self-expansion orientation should be researched further. However, for the purposes of this dissertation, I decided to focus on explicit conversations about differences, given that they appear potent, but also given that there are clear questions around how they can be implemented.

5.4.3. *Minority- and majority-status participants*

This study was the first in this thesis that also included minority-status participants. They entered the intervention with the same diversity belief as White participants did and were equally affected by the intervention. While the sample is too small and the estimates too imprecise to definitely reject the notion that participants' status moderates the link between contact and the valuing of diversity, it suggests that the main findings so far are unlikely to be restricted to majority-status participants.

5.4.4. *Strengths and limitations*

This study relied on single-item measures with only five response options, which increases measurement error and might bias linear models. Therefore, proportional-odds models had to be used, which make estimations of effect sizes more difficult and not comparable with other studies using linear models. However, it replicates the effect shown so far primarily in university student samples and cross-sectional studies in a large adolescent sample that allowed a consideration of changes over time. The setting can be considered as quasi-experimental, so that changes in

attitudes can very likely be attributed to the intergroup contact experience rather than any underlying confounders. It also introduces a consideration of two distinct facets of valuing diversity and suggests that these are likely to be differentially affected by contact experiences.

5.5. Conclusions and connections

The study reported in the present chapter showed that participants' preference to work in diverse groups increased across the board over the course of an intense intergroup contact intervention. Furthermore, participants' recognition that diverse teams can produce better results increased among those who had conversations about differences. The use of pre- and post-measures that framed an intervention resembles an experimental design. While the lack of randomisation and a control group precludes the drawing of firm causal conclusions, this adds confidence to survey-based findings regarding the links between intergroup contact and diversity beliefs. Furthermore, the results concerning conversations about differences and self-expansion orientation suggest design features of contact interventions that might be tweaked to maximise the impact of contact on valuing diversity. The next chapter will report an experiment that aimed to assess the impact of scheduled conversations about differences and their value.

CHAPTER 6: Increasing the potency of contact

by talking about differences – a field experiment

The previous chapter showed that participation in the NCS summer programme, a contact intervention for 16-year olds in England, was associated with increases in valuing diversity. However, there was substantial heterogeneity: the increase in preferences for working in diverse teams was stronger among participants who engaged in frequent conversations about differences, while the recognition that diverse teams have instrumental value only increased among that subset of participants. Therefore, the present chapter presents a randomised controlled trial in the field in which participants were allocated to different conditions that triggered conversations about differences and diversity, with the aim to increase the effect of the NCS programme on participants' valuing of diversity.

6.1. Motivation for the intervention design

Conversations about differences are associated with positive contact outcomes (Vezzali et al., 2017 and previous chapter here). However, simply triggering such conversations does not always lead to the desired outcomes. In an earlier field experiment in the context of the NCS programme, scheduled conversations about differences between participants had no effects on social trust for the majority of participants, while similar conversations about commonalities increased trust (Sanders et al., 2017). The conversations about differences only had a positive effect on participants already high in social trust. Earlier research suggested that what might matter most is an explicit appreciation of the benefits the differences bring to participants' lives (Nagda, 2006).

Therefore, the intervention aimed to trigger conversations about differences

in identities and life experiences with a framing that attempted to prompt a positive appraisal of diversity. Rather than explicitly setting injunctive norms, the interventions attempted to shift the perception of descriptive norms (i.e. of what the most commonly held beliefs are), which has been shown to be more effective than exhortations or most forms of enforcement (Cialdini, 2007). Therefore, the conversations did not address unfair differences in status and opportunities, but rather differences worthy of appreciation and capable of adding value to everyone's life experiences. This general approach is in line with a recent intervention on university campuses which communicated that a majority of peers held pro-diversity attitudes and engaged in inclusive behaviours, which led to more positive attitudes on the part of majority-status students and a greater sense of inclusion on the part of minority-status students (Murrar et al., 2020).

6.2. Interpersonal versus intergroup differences

The previous chapter, and the majority of past research, did not distinguish between different types of differences contact partners might talk about. Conceptually, however, conversations might focus on the group-level, and then be about cultural practices or status-inequalities, or about the individual-level, with a much broader range of potential topics, including cultural differences but not limited to them. Diversity can be valued at both levels, yet their discussion might result in distinct outcomes. Therefore, two interventions were tested here, each targeting one of these levels.

6.3. Context for the present study

Like the impact assessment presented in the previous chapter, the current research was conducted in partnership with a UK volunteering program, the National Citizens Service (NCS). The NCS was introduced in 2011 to "promote a more

responsive, cohesive and engaged society" (Mycock & Tonge, 2011) and now reaches more than 100,000 16-year-olds per year (NCS Trust, 2020).They join a four-week programme with a diverse group of youth from their local area that focuses on personal development, soft skills, and social action. Throughout the programme, they spend much of their time in teams of around 10 participants in which they engage in activities, reflect on their experiences and collaborate in designing and executing a social action project. In this context, I asked teams to take part in one of three group discussions, led by a facilitator. Two were interventions that intended to promote the valuing of diversity by means of triggering conversations about differences, while the third served as an active control.

This study used an active control condition, designed to control for non-specific intervention effects. In the circumstances, a passive control would not have been feasible, given that the groups would have spent the time together in any case. Moreover, the use of active controls that are close to interventions in all aspects but the ones under study has been widely suggested as a way to increase internal validity (Greenberg & Harris, 2012; Karlsson & Bergmark, 2015; Redick et al., 2015). Here, specifically, any effects of the interventions compared to a passive control could have been attributed to positive structured interaction in a diverse group. Since that was not the aspect of the intervention of interest, an active control was needed that also included such positive interaction, so that this could be ruled out as the explanation of any differences between the groups.

6.3.1. Hypotheses and questions

The primary focus of this experiment was to test whether explicit pro-diversity interventions work in an intergroup context with adolescents. For this, the hypotheses were that:

H1: A conversation focused on how individuals benefit from diversity in society, taking place in an intergroup context, will increase their valuing of diversity, particularly the value placed on diversity in society.

H2: A conversation focused on how individuals benefit from diversity in teams, taking place in an intergroup context, will increase their valuing of diversity, particularly the value placed on diversity in teams.

Given the potential for valuing diversity and diversity interventions to affect the nature of intergroup contact, and the evidence that they have various relationships with group status, I considered three questions without offering specific hypotheses:

Q1: Does the effect of the interventions differ between ethnic majority- and minority-status group participants?

Q2: Do the pro-diversity interventions affect the valence of subsequent contact experiences during the NCS programme and the frequency of conversations about differences?

Q3: Do the pro-diversity interventions affect participants' intentions to engage in future intergroup contact and to address diversity pro-actively?

Furthermore, I did not have specific hypotheses about which condition would be more effective. Thus a comparison of effect sizes was the final question of interest.

6.4. Methods

6.4.1. Procedure

Groups of approximately 10 participants within NCS cohorts of approximately 60 participants were randomly allocated to one of three conditions (Table 6.1). Each group met for 35-minutes on the first evening of the programme.

Full instructions for facilitators are available with the analysis code on GitHub.

Outcome measures were collected at the end of the NCS programme, approximately

three weeks after the discussions took place, so that any effects obtained are likely to

be robust over time. Participants and facilitators were blinded with regard to the

specific hypotheses.

As a manipulation check, facilitators were asked how closely they were able

to follow the instructions and how engaged the participants were (due to limitations

in data collection, their responses cannot be matched to specific participants). Out of

353 respondents, 69.1% reported that they implemented the activity exactly as

described, while only 3.1% reported making big adjusments. This did not differ

between the conditions, $\chi^2(6) = 6.73$, $p = .35$. Similarly, participants' engagement

(rated from 1 = *very disengaged* to 4 = *very engaged*) was reported to be high, mean:

3.16, SD: 0.75, and not dependent on the condition, $F(349, 2) = 1.17$, $p = .31$.

6.4.2. Participants

My research partner organisation asked 1,196 participants from 187 cohorts

who had taken part in one of the sessions to complete their impact assessment

survey, which included the measures used here. 635 of them (53.1%) responded to at

least one of the two dependent variables, while 417 additional participants (34.9%)

responded to at least one of the exploratory variables, with no significant differences

between conditions. The number of participants per condition who provided data on

one of the dependent variables and on any of the variables considered is shown in

Table 6.1.

Due to a split survey design unrelated to the present experiment, all items

regarding ethnicity were only collected from a randomly selected group of

participants. Among those ($N = 435$), 150 identified as White, 37 reported a mixed,

and 248 a minority-ethnic background (mostly South Asian, African and Carribbean British). I believe this is representative of the full sample, suggesting that approximately 60% of participants were from minority-ethnic backgrounds.

Table 6.1. *Overview of the conditions*

Condition	Activity
1. Valuing diversity in teams (N = 234/363)	Participants spend time to find differences between them. They can first collect obvious things that come to mind, and then take turns talking about their live experiences, interests and preferences when it comes to living and working together. Responses are written down.
	Then they are asked: "Of course sometimes differences make it harder to understand each other and to work together well. But they also often offer opportunities. Let's talk about that. How do you think you individually or we as a team can benefit from some of these differences?" Facilitators are asked to probe how some differences might allow participants to have new experiences, and how they might make the team more productive or creative.
2. Valuing diversity in society (N = 221/392)	Participants collect examples for how diversity has enriched life in Britain. They are split into two groups and asked to write down things that they enjoy that were made / introduced to the UK by someone from a different social group / ethnicity / faith.
	When they share what they wrote, facilitators have a set of examples to contribute if participants do not mention these themselves, including food, music, sports, celebrations. The session concludes by going around the group and having everyone say the one thing from the list that is most important to them.
3. Control: Getting to know each other (N = 180/297)	Participants sit in a circle and are asked to talk about themselves, their interests, hobbies and life experiences. All categories mentioned are written down.

Note. N denotes the number of participants allocated to that condition who provided data. The first number refers to responses to at least one of the dependent variables, the second number to responses to any variable considered here.

6.4.3. Pre-registration and deviations

This study was pre-registered on AsPredicted
(https://aspredicted.org/blind.php?x=97u35i). However, the pre-registration was
based on a plan to collect data from approximately 5,000 participants and their
facilitators, which would have been required to obtain sufficiently precise estimates
of effect sizes to enable meaningful comparisons between the intervention conditions
and to have adequate power for the testing of interactions in multi-level models.
However, *The Challenge* entered a period of organisational disarray half-way
through the implementation period that resulted in them entering administration, so
that the collection of responses to their programme impact survey became
intermittent. Therefore, the analyses presented here deviate substantially from the
pre-registration. I could not use repeated measures but had to rely on post-treatment
scores, as the required matching of participants would have reduced the sample size
even further. Furthermore, I could not test for treatment effects among the
participants who had actually experienced the intervention, as the facilitator
responses could not be matched to participant groups. Therefore, the results reported
concern the intention-to-treat, i.e. the effect of an allocation of partiicpants to
intervention, which almost certainly understates intervention effects. Apart from that,
planned mediation analyses and tests of interactions are now under-powered, so that
the results presented here should be treated with caution.

6.4.4. Measures

6.4.4.1. Valuing diversity in teams

Participants were given pairs of statements and asked which they agreed with
more, on a five-point differential scale. Firstly, "I prefer working with people who
are from a similar background to me" or "I prefer working in a team of people from
different backgrounds." Secondly, they were asked to imagine a team that needs to

deliver a project together in school, and say whether "In most cases, diverse teams will produce better results than one where everyone is from the same background" or "In most cases, a team where everyone is from the same background will produce better results than a more diverse team." I pre-registered that I would treat them as one scale if their correlation were above .5; in fact, it was .503. Therefore, they were averaged and formed a consistent scale, with Spearman-Brown reliability of .67. To support comparisons with the previous chapter, proportional-odds models with the two measures as separate outcomes are presented as robustness checks.

6.4.4.2. Valuing diversity in society

Again, participants were given pairs of statements and asked which they agreed with more, on a five-point differential scale. Firstly, "I would like to live in a country where more people share the same ethnic background" or "I am happy to live in a country where there are people from many different ethnic backgrounds" and secondly, "It would be better if people from minority groups just take on English values and traditions" or "It is good if people from minority groups keep their own values and traditions alive". I had also pre-registered to treat them as a single scale if the correlation exceeded .5; it was .63. Therefore, the items were averaged to form a consistent scale, with Spearman-Brown reliability of .77.

6.4.4.3. Exploratory variables

Four exploratory variables were considered: the valence of participants' intergroup contact during the programme and the frequency of conversations with other participants about differences in their lives were considered as potential mediators of the intervention effects, while participants' interest in future intergroup contact at the end of the programme and their commitment to addressing diversity in a team setting were considered as potential secondary outcomes. Contact valence

was measured with a single item, asking participants to report how positive or negative their interactions with "participants who are quite different from [them]" had been, from 1 = *very negative* to 5 = *very positive.* The frequency of conversations about differences was measures with a single item, asking participants "During NCS, how often did you discuss the differences in your everyday lives with other participants?" from 1 = *never* to 5 = *very often.*

Participants approach intentions regarding future intergroup contact was measured with a single item, asking White participants for their level of agreement with the statement: "In the future, I will actively approach people from a Black British background to get in touch" and participants from ethnic minority backgrounds the same with regard to "people from a White British background", from 1 = *strongly disagree* to 5 = *strongly agree.* Finally, their willingness to address diversity in a team setting was measured with two items, asking participants how important it is in a team "to openly discuss race, gender, religion, disability, wealth, sexuality etc." and how comfortable they would be with having such discussions, each on a five-point scale from 1 = *not at all important/comfortable* to 5 = *very important/comfortable.* The mean resulted in a scale with Spearman-Brown reliability of .57.[16]

6.4.5. *Missing data and analytic approach*

As mentioned above, questions regarding ethnicity were only presented to approximately half of the participants. This also included the measures regarding the valuing of diversity in society and the interest in future intergroup contact, so that these analyses are based on smaller sample sizes. Specifically, the number of

---

[16] Due to this low reliability, I also ran separate analyses with the two items. These revealed the same trends as are reported below. Asking for the levels of comfort and importance was deemed to be a better measure of willingness to address diversity than a direct request for a behavioural intention, as it seemed to be less likely to be dominated by social desirability pressures.

responses for each dependent/exploratory variable was as follow: valuing diversity in teams: 538, valuing diversity in society: 280, approach intentions regarding future intergroup contact: 384, contact valence: 994, willingness to address diversity: 979. In order to use all available information, each analysis is based on all participants who responded to the relevant variables.

The analyses were conducted by running mixed-effects models for each outcome; random intercepts were included to account for shared variance between participants in the same cohort. Modelling was done using the *lme4* package in R (Bates et al., 2015). Significance testing was done using the Kenward-Roger approximation to degrees of freedom, as it has been shown to produce acceptable Type I error rates at any sample size (Luke, 2017). As pre-registered, null models with random intercepts were tested against the single-level null model to determine whether random effects should be included. The random intercepts for cohorts were significant when it came to predicting the valuing of diversity in teams, LR $\chi^2(1) =$ 4.26, $p = .039$. They were not significant for the valuing of diversity in society, or for most of the exploratory variables. However, for consistency, multi-level models are reported throughout.

6.5. Results

Table 6.2 presents the descriptive statistics for the entire sample (grand means and correlations are shown in Appendix 6.1). Across the dependent variables and the two potential secondary outcomes, there were no substantial differences between conditions, with ps > .70. As can be seen in Table 6.2, there were some differences between conditions among the potential mediators. They were not significant with regard to contact valence, $F(2, 126) = 1.39$, $p = .253$, but were significant with regard to the frequency of discussions of differences, $F(2, 118) =$

3.15, p = .046. Post-hoc tests will be reported below.

Table 6.2. *Means and standard deviations by condition*

Condition	N	Valuing diverse teams	Valuing diverse societies	Contact valence	Discussing differences	Approach intentions	Willingness to address diversity
1. Diversity in teams	363	2.88 (0.99)	2.98 (1.13)	3.88 (0.75)	3.45 (0.93)	3.37 (0.91)	3.76 (0.86)
2. Diversity in society	392	2.85 (1.05)	3.02 (1.18)	3.79 (0.82)	3.27 (0.90)	3.42 (0.81)	3.68 (0.87)
3. Control	297	2.81 (1.01)	3.02 (1.14)	3.92 (0.75)	3.39 (0.90)	3.40 (0.80)	3.74 (0.88)

Note. *N* denotes the number of participants allocated to that condition who provided data on at least one of the variables. Standard deviations are shown in brackets.

In contrast to the null-finding for all outcomes, Table 6.3 highlights that the results differed substantively depending on the participants' group status. With regard to the two primary outcome measures – the valuing of diversity – minority-ethnic participants' responses tended to shift in a positive direction after the interventions, while White participants' responses tended to shift in a negative direction. (In line with the pre-registration, the small number of participants who reported a mixed background were excluded from this analysis.)

Table 6.3. *Means and standard deviations by condition and group status*

Condition	N	Valuing diverse teams	Valuing diverse societies	Contact valence	Discussing differences	Approach intentions	Willingness to address diversity
Ethnic minority							
1. Diversity in teams	84	3.07 (0.93)	3.08 (1.18)	3.83 (0.79)	3.49 (1.06)	3.34 (0.93)	3.82 (0.92)
2. Diversity in society	84	3.13 (0.98)	3.00 (1.02)	3.75 (0.77)	3.28 (0.74)	3.27 (0.72)	3.59 (0.81)
3. Control	80	2.50 (1.00)	2.92 (1.14)	3.89 (0.79)	3.39 (0.90)	3.24 (0.71)	3.75 (0.89)
White							
1. Diversity in teams	73	2.74 (0.93)	2.82 (0.98)	3.91 (0.70)	3.42 (0.88)	3.41 (0.88)	3.89 (0.86)
2. Diversity in society	45	2.72 (1.05)	3.08 (1.37)	4.09 (0.88)	3.23 (0.99)	3.67 (0.87)	3.98 (0.78)
3. Control	32	2.92 (0.80)	3.23 (1.21)	4.13 (0.72)	3.59 (0.91)	3.77 (0.82)	3.80 (0.98)

Note. *N* denotes the number of participants allocated to that condition who provided data on at least one of the variables. Standard deviations are shown in brackets.

6.5.1. Intervention effects by group status

To test the significance of these differences between groups, I ran multi-level models considering only the participants who reported their ethnicity. Regarding the valuing of diversity in teams, there was no main effect of condition, $F(2, 59) = 0.88$, $p = .42$, but there was a significant interaction between group status and condition, $F(2, 196) = 3.97$, $p = .020$. The results can be seen in Figure 6.1. The team-focused intervention had the expected effect for minority-ethnic participants and substantially boosted their valuing of diverse teams, with Hedges' $g = 0.59$, 95% CI [0.14, 1.08], Holm-adjusted $p = .026$. However, conversations about the valuing of diversity in society had a very similar effect on them, with Hedges' $g = 0.63$, 95% CI [0.17,

1.13], Holm-adjusted p = .021. Neither intervention had a statistically significant effect on White participants' valuing of diversity in teams.

Figure 6.1. *Valuing diversity in <u>teams</u> depending on intervention and group status*
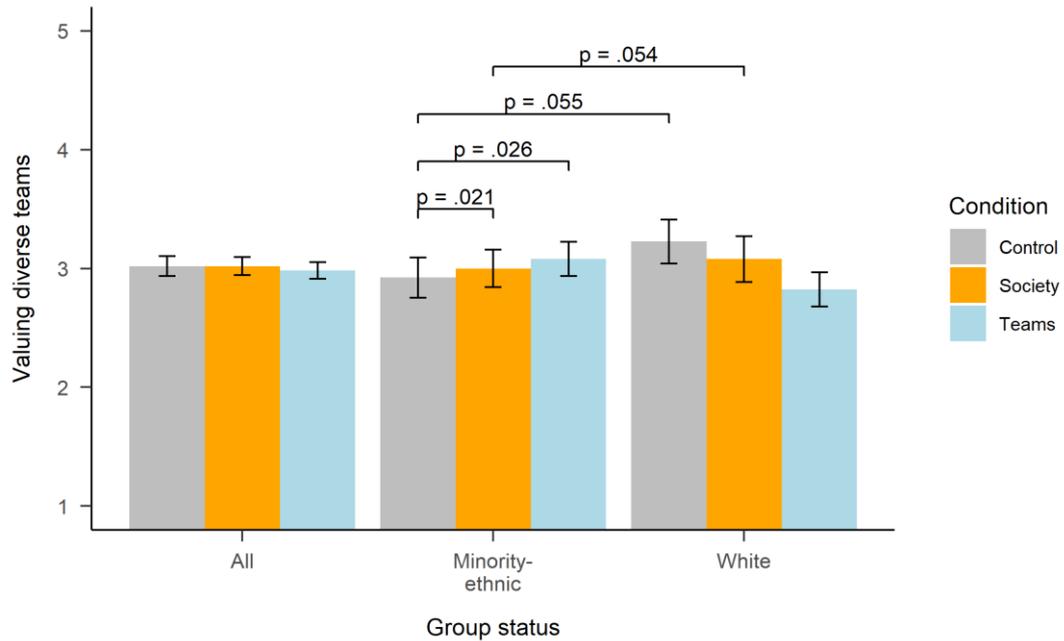
Note: Error bars show standard errors of the mean, p-values are adjusted with Holm-Bonferroni method.

Regarding the valuing of diversity in society, there was no main effect of condition, $F(2, 54) = 0.19$, $p = .83$, and no interaction, $F(2, 244) = 1.05$, $p = .35$.

6.5.2. *Intervention effects on exploratory variables*

Apart from the impact of the interventions on valuing diversity, this study included exploratory investigations of two potential mediators and two potential supplementary outcomes. The relationship between conditions and these variables, depending on participants' group status, is shown in Figure 6.3.

6.5.2.1. **Potential mediators**

As reported above, there was a main effect of intervention on the frequency with which participants discussed differences in their lives with others over the course of the programme, $F(2, 118) = 3.15$, $p = .046$. Unexpectedly, the intervention

focused on diversity in society appeared to lead to *fewer* discussions about differences over the course of the programme, compared to the team-focused intervention, with Hedges' $g = -0.20$ [-0.05, -0.35], Holm-adjusted $p = .051$, and the control, with Hedges' $g = -0.13$ [0.02, -0.29], Holm-adjusted $p = .18$. The team-focused intervention did not differ from the control, with Hedges' $g = -0.07$ [0.09, -0.23], Holm-adjusted $p = .57$. There was no interaction with group status, $F(2, 355) = 0.71$, $p = .493$. However, given that the interaction tests were underpowered, it is worth considering the effect sizes for White and minority-ethnic participants separately. Here, it appears that the decline might have been more pronounced among White, with Hedges' $g = 0.38$ [-0.08, 0.86], unadjusted $p = .087$, than among minority-ethnic participants, with Hedges' $g = 0.12$ [-0.20, 0.44], unadjusted $p = .55$.

Given that conversations about differences were positively correlated with the valuing of diversity in society, $r = .20$, $p = .001$, this potential mediator might in fact have acted as a suppressor. Therefore, a simple mediation model was tested (Figure 6.2). It was estimated using the *lavaan*-package (Rosseel, 2012) in R; confidence intervals for indirect effects were based on 5,000 bootstrap resamples.[17] The results showed that there was a negative indirect effect of the society-focused intervention (compared to both the control and the team-focused intervention) on participants' valuing of diversity in society through a decrease in the frequency of conversations about differences throughout the programme, with a standardised coefficient of -0.06, 95% CI [-0.15, -0.01]. The point estimate for the direct effect was of similar size and went in the opposite direction; however, due to greater variability, the direct effect is far from significant, with 0.09, 95% CI [-0.18, 0.36].

---

[17] Confidence intervals are percentile-based, simply ranging from the 2.5th to the 97.5th percentile of the bootstrap distribution, in line with the widely used PROCESS Macro (Hayes, 2012) and as per the recommendation of a simulation study that found that it offered the best compromise between power and Type I error rate (Hayes & Scharkow, 2013).

Figure 6.2. *Suppression model for effect of society-focused session*
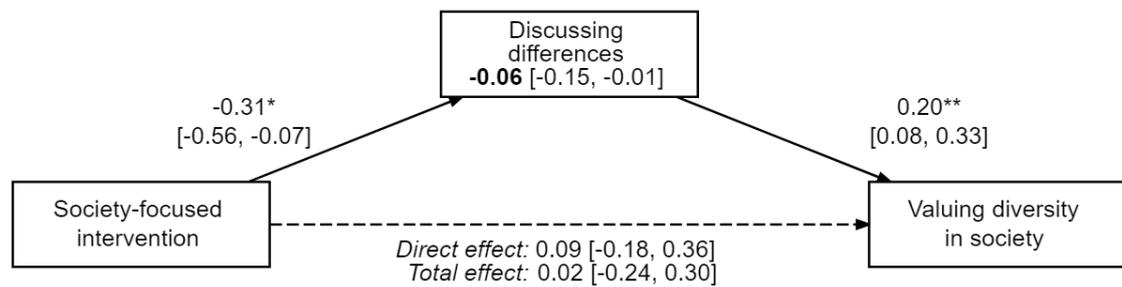
Notes: Standardized coefficients. * *p* < .05, ** *p* < .01

Regarding contact valence, there was no significant main effect, $F(2, 57) = 0.97$, $p = .387$ or interaction with group status: $F(2, 360) = 0.94$, $p = .392$.

6.5.2.2. Potential outcomes

Regarding the two potential secondary outcomes, there were again no main effects or significant interactions with ethnicity, i.e., on approach intentions regarding future intergroup contact, main effect: $F(2, 55) = 0.76$, $p = .474$, interaction: $F(2, 330) = 2.55$, $p = .079$ or on willingness to address diversity in teams, main effect: $F(2, 63) = 0.13$, $p = .879$, interaction: $F(2, 361) = 0.88$, $p = .415$.

It should be noted, however, that these tests of interactions are severely underpowered due to the small sample and the split questionnaire design that led to information on ethnicity to be available for even fewer participants. Thus, the pattern of interactions should still be considered for all exploratory variables, so that trends that might inform future research can be identified. Figure 6.3 shows the estimated marginal means by condition and group status and highlights the intervention effects that approach statistical significance.

Figure 6.3. *Exploratory variables depending on intervention and group status*
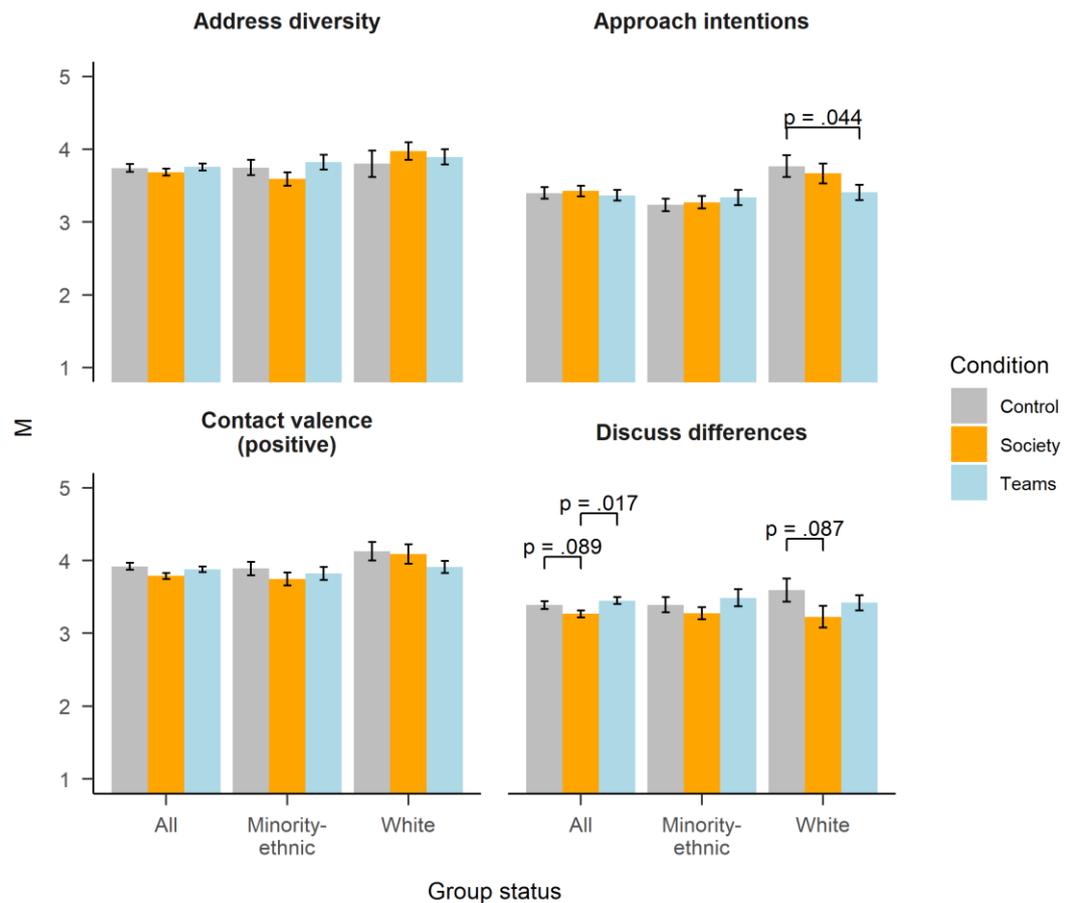
Notes: Error bars show standard errors of the mean. *p*-values are not adjusted for multiple comparisons. Top row shows potential outcomes, bottom row potential mediators.

Concerningly, the team-focused intervention led to reduced approach intentions regarding future intergroup contact among White participants, with Hedges' g = 0.42 [-0.01, 0.86], unadjusted p = .044. It needs to be emphasized that this post-hoc comparison follows a marginally significant tests of the interaction and would not be significant with appropriate controls for multiple comparisons. Nevertheless, this decline represents trends that run counter to expectations and suggest that there might be a backfiring effect among White participants, which needs to be borne in mind for future research.

6.5.3. Separate analysis of the diversity outcomes

It was shown in the previous chapter that the two items used to measure

valuing diversity might represent conceptually and empirically distinct facets of the concept. Therefore, I ran additional proportional-odds models with the two items as separate outcomes. Again, there was no main effect of the intervention on either facet of valuing diversity. The interaction with ethnicity only reached significance regarding the personal preference for working in diverse teams, $LR\ \chi^2(2) = 8.95$, $p = .01$. Figure 6.4 shows the estimated probability of each response to the question regarding the preference for diverse teams by activity and participants' group status; it shows that both the team and the society-focused activity were associated with higher probabilities to express a preference for diverse teams among minority-ethnic participants, while there was no such relationship among White participants. Regarding the belief in the instrumental value of diverse teams, the interaction did not reach significance, even though the estimates pointed in the same direction, $LR\ \chi^2(2) = 2.56$, $p = .28$.

Figure 6.4. *Impact of activities on preferences for diverse teams by ethnicity*
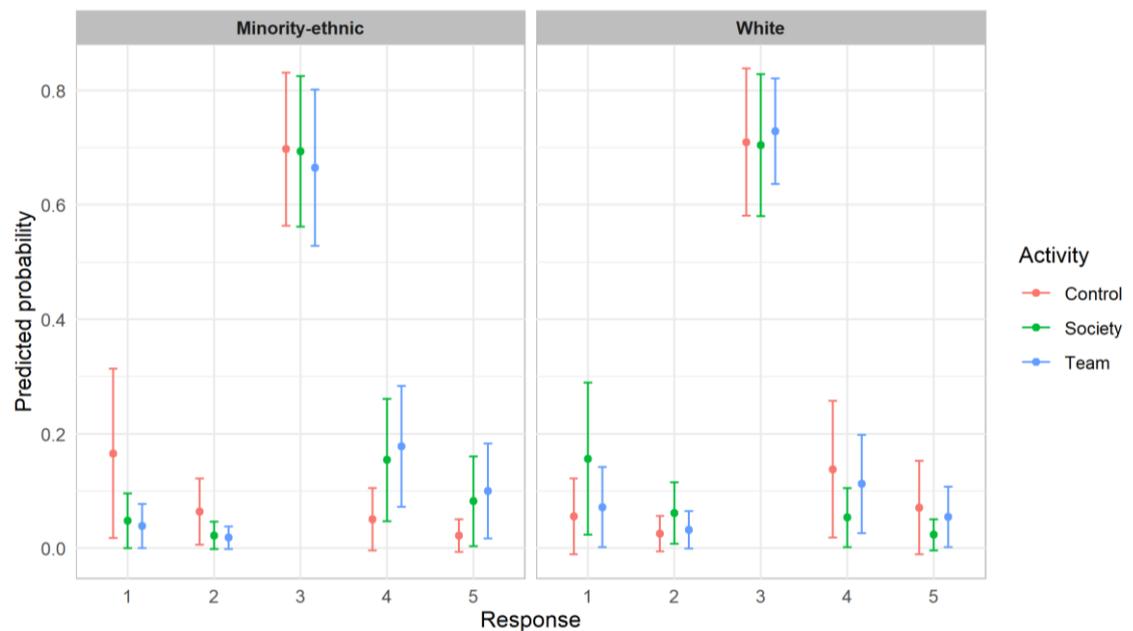
Note: Error bars show 95% confidence intervals.

6.6.   Discussion

Overall, the results show that the interventions did not have the desired effect

as they did not influence the valuing of diversity – either in teams or in broader society – at the level of the full sample. However, this null finding is qualified by an interaction, which showed that both interventions served to significantly increase the valuing of diversity in teams among minority-ethnic participants, while the estimated effect for White participants was not significant yet pointed in the opposite direction. Unexpectedly, the condition that focused on diversity in society led to a reduction in the frequency of conversations about differences over the course of the programme, which then had a significant negative indirect effect on the valuing of diversity in society. This effect appeared particularly pronounced among White participants though the sample was too small to test that interaction with confidence. Importantly, there is also some evidence that the condition focused on the value of diverse teams decreased White participants' approach intentions regarding future intergroup contact.

Given that minority-ethnic participants showed a (marginally) lower preference for working in diverse teams in the control condition, the observed effect of the interventions on them is valuable, yet the lack of impact – or potentially the negative impact – on White participants is concerning. Several potential explanations for the observed pattern can be offered.

Initially, Vorauer (2008) suggested that contact will only succeed when participants' mental focus is on the outgroup rather than on the ingroup. A pro-diversity framing, especially when introduced by someone in a position of authority, might achieve this for minority-group members, while it might have a different effect on majority-status participants. Given that White people often feel that their culture is not valued as part of multicultural celebrations (Harris, 2013), discussions of the value of other groups' contributions might raise symbolic threat. Such a threat-

induction has been shown in prior research; for instance, Morrison et al. (2010) showed that multiculturalism primes triggered threat reactions among White Americans who strongly identify with their ethnicity. It has also been shown that cultural threat impedes on cognitive function by reducing creativity (Chen et al., 2016); it thus appears likely that it can impede learning more broadly and thus dampen the impact of any intervention with a substantial cognitive component. In addition, such perceived threat might have increased anxiety in majority-status participants, which may be the reason for the (marginally significant) finding that participation in the condition that discussed diversity in society actually led to fewer conversations about differences over the course of the programme among majority-status participants. Along these lines, future research should test interventions that explicitly promote a valuing of *all* diverse contributions, including those from majority-status groups.

While this line of explanation might explain the differential impact of the condition focused on the valuing of diversity in society, it appears less pertinent for the condition that highlighted the valuing of diversity in teams. Here, a potential explanation could be that the communication of strong inclusive norms at the start of a contact intervention (in this case, on the first day of a four-week experience) might have a differential effect on different groups. A recent study highlighted the potential of such an intervention in terms of increasing the sense of belonging among minority-status students (Murrar et al., 2020), which might explain their increased desire to continue working in diverse teams in the future. For majority-status participants, on the other hand, the communication of similar norms, which might lead to a discussion of situations in which they are violated might result in a triggering of meta-stereotypes. Meta-stereotypes in intergroup contact often take the form of a belief that one is considered to be biased by one's interaction partner

(Finchilescu, 2010) and have been shown to be a potent predictor of avoidance among Whites, even when experimentally induced (C. MacInnis & Hodson, 2012). To date, there is little research into situational triggers of meta-stereotypes in natural settings; future intervention research could attempt to measure this, so that it can be tested whether such an adverse effect arises and, if so, how this can be mitigated.

Lastly, it has been suggested that majority-status participants can experience a diversity-shock when entering a more diverse environment that impedes on the expected effects of increased contact (Birtel et al., 2020). Given that a majority of participants in the sample came from minority-ethnic backgrounds, it can be supposed that many White participants experienced a level of diversity during the programme that strongly exceeded what they were used to. Emphasizing this diversity at the start of the programme might have exacerbated this 'shock', and thus counteracted the effect of the intervention. Given that the body of intergroup contact literature offers reasons to expect that such a shock will wear off with time (Dovidio et al., 2017), future research should test whether similar interventions scheduled later during a contact intervention might have a positive effect on all participants.

Conversely, the substantial positive effect of both interventions on minority-ethnic participants' valuing of diverse teams is encouraging, given that contact interventions typically only have weak effects on minority-status participants (Tropp & Pettigrew, 2005b). Given that it was not accompanied by an increase in contact valence during the programme, it most likely came about through an increase in their sense of belonging, induced by the establishment of a social norm that valued diversity. This would be in line with the findings by Murrar et al. (2020), yet future research should include a direct measure of belonging to identify the pathway(s) by which the effect of the intervention comes about.

6.6.1. Limitations and strengths

As noted above, the sample, though relatively large and diverse, was much smaller than planned, and thus did not offer sufficient power for reliable estimates of effect sizes, or tests of interaction effects. This was exacerbated by the fact that data on ethnicity was only available for a subset of participants. The fact that the interaction for one of the outcome measures was still significant highlights the size of the observed difference.

Furthermore, due to the reliance on an existing survey to which only a select few items could be added, this study relied on short scales with few answer options. This increases measurement error, which again makes effect size estimates imprecise and might have contributed to the lack of significance of some trends. Similarly, it precluded the complete exploration of pathways, so that much of the interpretation of the findings relies on conjectures. Additionally, the conditions were designed to each be meaningful experiences for the participants, which came at the expense of comparability. Specifically, the team- and society-focused sessions might not just have differed with regard to the facets of diversity under consideration, but also with regard to the intensity of participants engagement, given that the team-focused session required discussion, while the society-focused session could have been completed by merely listing ideas. Future research should aim to remove such confounds to increase the robustness the conclusions. Nevertheless, the findings suggest specific pathways for future research and highlight that simple pro-diversity interventions, though popular among practitioners, need to be carefully designed in order to achieve the desired effects.

Some strengths of the study should also be noted: its setting in the field, embedded in a widespread contact intervention, increases ecological validity and

thus the potential relevance for applied researchers and practitioners. Furthermore, the measurement of effects after a cooldown period of more than three weeks, a comparison to an active control, and the blinding of facilitators and participants to the specific hypotheses under consideration jointly result in conservative tests of the intervention effects.[18] They ensure that any effects observed are not due to situational priming, but due to somewhat lasting changes.

6.7. Conclusions and connections

The present study tested an intervention designed to enhance the effect of a contact intervention on participants' valuing of diversity. It only had the desired effect on one of two outcomes, and that only among minority-status participants, while there were some indications that the interventions might have small negative effect on majority-status participants. This highlights some complexities of intervention design and suggests that generating opportunities for positive and intimate contact without specifically promoting valuing diversity might be an effective way to increase the valuing of diversity. This will be further discussed in the concluding chapter.

---

[18] The fact that I could not include manipulation checks on the participant level, and thus estimated the intention-to-treat effect rather than the treatment effect, further adds to the conservative nature of the test.

CHAPTER 7:

Meta-analysis of the mediation effect across all samples

This dissertation proposes that intergroup contact affects the valuing of diversity, and that the valuing of diversity mediates the association of intergroup contact with outgroup attitudes. These relationships can be tested in all samples used for this dissertation, so that a single-paper meta-analysis can be used to assess the robustness of the finding and estimate the effect sizes more precisely.

Single-paper meta-analyses have been advocated as a valuable addition to multi-study research projects, as they enable a quantitative assessment of the overall significance of findings and provide better estimates of likely population effect sizes than any of of the individual studies could (McShane & Böckenholt, 2017). However, to obtain reliable results, meta-analyses – particularly those with a small sample size – need to include all available effect sizes, and can quickly be distorted if any datasets remain in the proverbial file-drawer (Vosgerau et al., 2019). Therefore, I include an additional dataset and multiple additional measures into this chapter that were not relevant to the analyses presented to far, yet can contribute effect sizes here.

7.1. Two methodological challenges

Most meta-analyses aim to estimate a single overall effect size from a set of independent tests, and to potentially test moderators of that effect. Here, however, I wanted to use all available information and thus include multiple tests from the same sample, and to test the significance of indirect effects.

7.1.1. Use of dependent effect sizes in meta-analyses

Frequently, papers report more than one relevant effect size, for instance when different measures of the same broad construct are used – for instance, the mediation analyses presented in Chapter 3 included both cognitive and affective

measures of outgroup attitudes. However, as they are based on the same sample, these effect sizes are clearly dependent on each other. Given that the independence of observations is a key assumption of traditional meta-analytic frameworks, such dependence cannot be ignored as that might lead to biased estimates and inflated Type I error rates (M. W.-L. Cheung, 2015b). Most meta-analyses deal with this problem by deriving a single effect-size from each independent sample, either by averaging the reported effects or by selecting one effect size, even though the presence of dependence or the solution are often not clearly reported (Ahn et al., 2012). In general, the averaging of effect sizes is problematic as it understates the variance of results and reduces statistical power, while the selection of the 'most relevant' effect size affords degrees of freedom to the researchers that are undesirable in the pursuit of replicable research (Lakens et al., 2016). In the field of contact research, various approaches have been used. Lemmer and Wagner (2015), in meta-analytically reviewing direct contact interventions, and Miles and Crisp (2015), in meta-analysising imagined contact, reported using systematic approaches to selecting effect sizes within samples that were informed by theory and only resorting to averaging when necessary. Pettigrew and Tropp (2006), on the other hand, averaged across the reported tests within each sample. However, none of the major meta-analyses to date employed an approach that allows for the inclusion of all reported effect sizes and thus makes full use of the information available.

Fortunately, such approaches have been developed in recent years. They are based on the notion that effect sizes as are nested in samples, so that the dependence can be accounted for by using multi-level models (M. W.-L. Cheung, 2015b). A practical approach to this has been outlined by Wilson et al. (2016), using the *metafor* (Viechtbauer, 2010) and *metaSEM* (M. W.-L. Cheung, 2015a) packages in R, which enabled me to include all available information when estimating the meta-

analytic correlations between the constructs under consideration here.

7.1.2. Meta-analyses of mediation

Most meta-analyses focus on the estimation of main effects and the testing of moderators, while a plethora of research focuses on mediation. To date, the major example of a meta-analysis of mediation in the contact field was provided by Pettigrew and Tropp (2008), who showed that intergroup empathy and anxiety can explain a substantial share of the contact-prejudice link. However, their method relied on the Sobel-test for the significance of indirect effects, which should not be applied when using multi-level models as I intend to do here (D. MacKinnon, n.d.).

Since then, two distinct approaches for the the meta-analytic testing of indirect paths have emerged (M. W.-L. Cheung, 2020). Indirect and direct effects can either be calculated for each sample and then included into a multivariate meta-analysis, or an overall correlation matrix can be estimated meta-analytically and then be included into a structural equation model (Two-Stage SEM; TSSEM). The latter allows for the inclusion of dependent effect sizes, as explained above, enables the estimation of more complex models, and reduces the influence of measurement error as multiplication only takes place with the more reliable combined correlation estimates. Therefore, it is preferred here.

7.2. Methods

7.2.1. Samples and effect sizes

I included all samples used for this dissertation, and any other samples I have collected that include some measures of positive and negative contact frequency, valuing diversity and intergroup attitudes. This included the five samples presented in the dissertation so far, and one additional sample collected in Uttar Pradesh, India

(which will be described and discussed in the results).[19] Together, they contained

6,304 participants, providing 57 effect sizes relevant to the test of the mediation, split

into between 7 and 12 effect sizes for each correlation. Table 7.1. shows the dataset.

Table 7.1. *Effect sizes included in the meta-analysis*

			Correlations (Pearson's *r*)					
			Pos ~	*Pos ~*	*Neg ~*	*Pos ~*	*Neg ~*	*Div ~*
Study		***N***	*Neg*	*Div*	*Div*	*Att*	*Att*	*Att*
Ch. 2: Longitudinal	*T1*	211	.16	.22	-.02	-.34[a]	.10 [a]	-.24 [a]
	T2		-.05	.33	-.28	-.46 [a]	.20 [a]	-.41 [a]
	T2					-.37[b]	.22 [b]	-.38 [b]
Ch. 3: Mediation UK		224	-.01	.43	-.30	-.38[c]	.18[c]	-.44[c]
						-.20[d]	.21[d]	-.55[d]
Ch. 4: Mediation DE		2,618	-.27	.34	-.17	-.42	.27	-.52
Ch. 5: NCS 2018		2,165	-.17	.12[e]	-.04[e]	-.22	.06	-.08[e]
				.09[f]	-.05[f]			-.03[f]
Ch. 6: NCS 2019		934	-.15	09[g]	-.13[g]	-.25	.15	-.07[g]
				.06[h]	.02[h]			-.01[h]
Suppl.: India		152	.00	.14[e]	-.02[e]	-.56	-.02	-.21[e]
				-.06[f]	-.25[f]			-.21[f]

Notes: In the correlations, Pos: positive contact, Neg: negative contact, Val: valuing diversity, Att: outgroup attitude. The superscript letters indicate the differences between measures in the case of multiple effect sizes per sample:

[a] Thermometer measure, [b] Prejudicial beliefs, [c] Affective Prejudice, [d] Cognitive Prejudice, [e] Preference for diversity, [f] Instrumental value of diversity [g] Valuing diversity in teams [h] Valuing diversity in society

---

[19] While most samples only included majority-status participants, the two samples from the National Citizen Service also included minority-status participants. However, the correlations did not significantly differ between these groups, as the comparison of restricted (equal covariances) and unrestricted multilevel models in *lavaan* showed (for 2018: $\chi^2(20) = 9.62$, $p = .97$, for 2019: $\chi^2(20) = 15.94$, $p = .72$). Therefore, data from all participants was included in this meta-analysis.

7.2.2. *Estimation of the meta-analytical values*

Initially, all observed effect sizes were used to estimate an aggregated correlation matrix, using a multilevel model that accounted for the nesting of effect sizes in studies and measures. This followed the procedure proposed by Wilson et al. (2016) and relied on the *metafor*-package (Viechtbauer, 2010) for the estimation. I deviated from Wilson et al. (2016) by using Restricted-Maximum Likelihood (REML) to estimate the model, as the standard maximum likelihood approach underestimates the variance, particularly in small datasets (Gonzalez-Mulé & Aguinis, 2018). Furthermore, I allowed the residual heterogeneity to differ for each correlation estimate, which is recommended when there are theoretical reasons to expect heteroscedasticity (Rubio-Aparicio et al., 2020). Here, this appeared likely, if only because some constructs were based on nearly identical measures across studies (e.g., contact), while others were based on very diverse measures (e.g., attitudes).

To estimate the correlations, a mixed-effects model was used that assumes that the effect sizes considered here do not constitute the population of effect sizes of interest, but rather a random draw from that population (Gonzalez-Mulé & Aguinis, 2018). This follows the established procedure in the field (Lemmer & Wagner, 2015; Pettigrew & Tropp, 2006), as it provides for a more conservative test of the hypotheses and enables a (tentative) generalisation of the results beyond the studies considered here.[20]

Typically, effect sizes in a meta-analysis should be weighted by their sampling variance. However, in the case of correlation coefficients, this is problematic because their variance depends on their value. Therefore, I followed

---

[20] In addition, mixed-effect models assign more even weights to the effect sizes that do not exclusively depend on the sample size but also on the estimate of heterogeneity (Gonzalez-Mulé & Aguinis, 2018). While this has occasionally been described as a limitation, I consider it a benefit here as it ensures that the results are not dominated by the studies with the largest samples.

Wilson et al. (2016) and Pettigrew and Tropp (2008) and weighted the effect sizes solely by the sample sizes.

After the estimation of the correlation matrix, the *metaSEM*-package (M. W.-L. Cheung, 2015a) was used to estimate the hypothesized mediation model, including both positive and negative contact as simultaneous predictors. As the model is just identified, no fit indices are available. The significance of indirect effects was determined by considering whether likelihood-based confidence intervals contain zero, while significance values for individual paths are based on Wald-tests.

7.3. Results

Table 7.2 shows the aggregated correlation matrix. All correlations were significant and of small to medium size. It is worth noting that positive contact was more strongly associated with outgroup attitudes than negative contact was, while both had similar links to valuing diversity. Further, the frequencies of positive and negative contact were not correlated.

In this step, it also became apparent that there was substantial residual heterogeneity between the effect sizes, $Q(\mathrm{df} = 51) = 762.68$, $p < .001$. Regarding the different correlation coefficients, the τ^2 index of residual heterogeneity was largest for the correlation between valuing diversity and outgroup attitudes (.045); looking at Table 7.1, it can be surmised that this is due to the weak correlation between valuing diversity and outgroup attitudes int the two NCS samples. There was also substantial heterogeneity regarding the association between the frequencies of positive and negative contact (.015). For all other correlations, the estimated heterogeneity between studies amounted to less than 4% of the estimated size of the correlation coefficients.

Table 7.2. *Aggregated correlation matrix*

	Positive contact	Negative contact	Valuing diversity
Negative contact	-.09 [-0.22, 0.04]		
Valuing diversity	.15 ** [0.06, 0.25]	-.11 ** [-0.19, -0.03]	
Outgroup attitudes *(negative)*	-.34 *** [-0.45, -0.24]	.17 *** [0.11, 0.23]	-.27 ** [-0.45, -0.10]

Notes: *** $p < .001$, ** $p < .01$, * $p < .05$. Values in square brackets indicate the 95% confidence interval for each correlation. Based on 57 effect sizes from 6 independent samples, total $N = 6,304$.

The aggregated correlation matrix, together with its estimated covariances, was then used to estimate the mediation model shown in Figure 7.1. All paths in the mediation model were substantial and significant. Likelihood based confidence intervals indicated that both indirect effects were significant, for positive contact with a path estimate of -0.03, 95% CI [-0.07, -0.01] and for negative contact with an estimate of 0.02, 95% CI [0.003, 0.05]. With estimated total effects of 0.33 for positive contact and 0.14 for negative contact, the indirect effect amounted to 9.3% of the total effect in the case of positive contact and 14.1% in the case of negative contact.

Figure 7.1. *Meta-analytic mediation model*
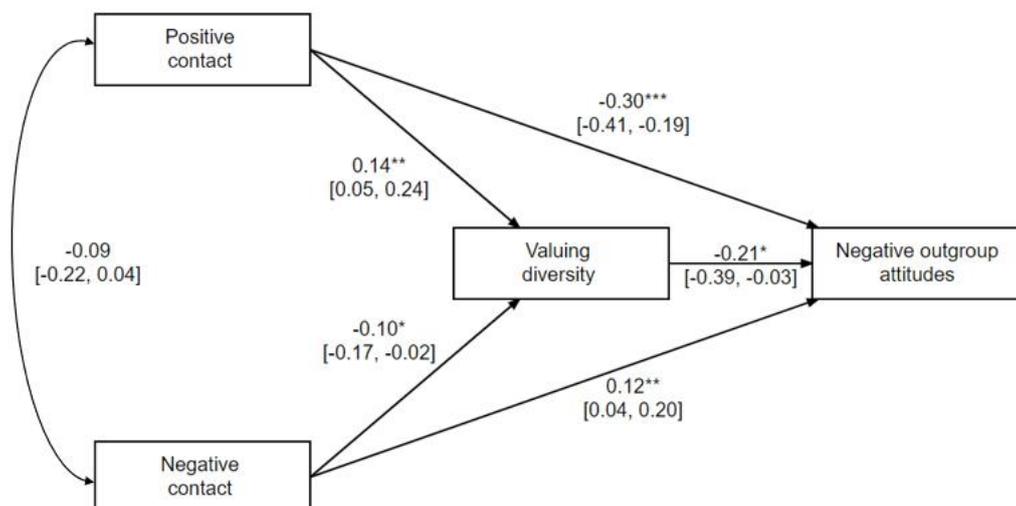
Notes: *** $p < .001$, ** $p < .01$, $p < .05$. Coefficients are standardised, 95% CI are shown in square brackets

7.3.1. The Indian sample: meta-analytical insights into a null-finding

The Indian data included here were collected for a separate study into facilitation and buffering effects for intergroup contact (Wallrich et al., under review). It focused on majority-status Hindus' contact with Muslims in Uttar Pradesh, a state that has regularly seen communal violence between Hindus and Muslims. I added a measure of valuing diversity into the survey for an exploratory cross-cultural analysis of the links between contact, diversity beliefs and outgroup attitudes. However, the observed correlations between contact and two measures of valuing diversity were relatively small and not significant in three out of four tests. To determine whether these null findings are an indication for cross-cultural differences or more likely the result of sampling error in this small study, I used the meta-analytical framework to test whether they were significantly different from earlier findings.

Therefore, I added a dummy variable distinguishing the Indian effect sizes from others as a moderator to the mixed-effects model. When including all correlations, this moderator was not significant, $F(1, 50) = 2.00$, $p = .16$. The largest deviations between the Indian effect sizes and the remainder concerned the zero-order correlations between contact and attitudes, which are of little interest to the present work.[21] When including only the correlations related to valuing diversity, the moderation effect was again not significant, $F(1, 28) = 0.32$, $p = .58$. Therefore, I refrain from interpreting the results from the Indian sample as substantive findings

---

[21] The estimated correlation between the frequency of positive contact and attitudes was larger than in the remaining samples, while the correlation between negative contact and attitudes was smaller. This might be because negative contact is more normative in that context, and thus might contain comparatively less information that is capable of influencing attitudes. While I would be curious to explore that hypothesis, it is outside the scope of this dissertation.

regarding links between valuing diversity and intergroup contact.

7.4. Discussion

The multilevel meta-analysis presented here indicates that the data contained in this dissertation provides consistent evidence for the main finding reported: the association of both positive and negative contact with outgroup attitudes is partly mediated through changes in the valuing of diversity. There is significant heterogeneity between the studies, which suggests that future research into moderators and boundary conditions would be fruitful. Nevertheless, the extent of heterogeneity, as indicated by the τ^2 estimates, is relatively small for all correlations except for that between the frequency of positive and negative contact, and the link between valuing diversity and outgroup attitudes, even though the effect sizes stem from studies with different age groups (adolescents, university students and adults), in different cultural and varying intergroup contexts. The implication of these findings will be discussed in the next chapter.

CHAPTER 8:

General discussion and conclusions

This chapter will summarise the findings of this dissertation, highlight the resulting contributions to the literature and reflect on some key features of the research design. On that basis, I will then discuss limitations and future directions.

8.1. Brief summary of the empirical chapters

The research started from three main findings in the literature review: (1) identity-conscious approaches to diversity (i.e., those that explicitly value diversity) lead to better intergroup outcomes than identity-blind approaches (Leslie et al., 2020), (2) intergroup contact is an effective path towards positive intergroup outcomes (Lemmer & Wagner, 2015; Pettigrew & Tropp, 2006), and yet (3) their relationship has received little empirical attention and been conceptualised divergently (Harper & Yeung, 2013; Tropp & Bianchi, 2006). Six empirical chapters then addressed this under-explored link.

8.1.1. Chapter 2: Contact shapes the valuing of diversity over time

This chapter contained the results of a longitudinal study of first-year psychology students in two English universities ($N = 211$). The students were surveyed at the start of their first academic year, and then three months later. A cross-lagged panel model revealed that positive and negative contact experiences with Black British people contributed to increases/decreases in the valuing of diversity over time. The reverse paths from valuing diversity to the frequency of positive and negative contact were weaker and not statistically significant. Therefore, it appeared warranted to focus on valuing diversity as a potential outcome of intergroup contact, which shaped the perspective of the subsequent chapters.

This finding confirmed the only previous longitudinal result on the link

between intergroup contact and valuing diversity (Harper & Yeung, 2013), thus showing that their finding was likely substantive rather than an artefact of a key methodological weakness (the absence of the baseline level of valuing diversity from the model). Conversely, it failed to support the suggestion advanced in previous cross-sectional research that valuing diversity predicts intergroup contact (Tropp & Bianchi, 2006). While the results presented here do not rule out that valuing diversity might shape future contact, they strongly suggest that the path from contact to valuing diversity is stronger, at least in the short run.

8.1.2. *Chapter 3: Valuing diversity mediates various effects of intergroup contact*

Having shown that intergroup contact appears to shape the valuing of diversity in Chapter 2, I moved on to consider valuing diversity as a potential mediator of contact effects in Chapter 3. To gain a comprehensive understanding of mediational pathways, I contrasted valuing diversity with the two best-established mediators of the effects of contact on attitudes, intergroup empathy and anxiety (Pettigrew & Tropp, 2008). Given that contact has been shown to have distinct effects on different intergroup outcomes (Tropp & Pettigrew, 2005a), which likely come about through different pathways, I included both cognitive and affective outcomes, both with regard to measures of outgroup attitudes and to behavioural intentions. Additionally, I addressed the relative dearth of research into the mediation of the effects of negative contact (Wallrich et al., 2020b) by considering both positive and negative contact as predictors.

This design resulted in eight mediation models (2 predictors x 4 outcomes), which I tested through a survey of 224 psychology undergraduates in five English universities. As hypothesized, valuing diversity was the strongest mediator in

explaining the links between positive and negative contact and the cognitive

outcomes (cognitive prejudice and support for pro-diversity policies). It also

significantly contributed to explaining the association between contact and affective

prejudice, though with a much smaller effect size; regarding bystander intervention

intentions, valuing diversity did not make a significant contribution. The point

estimates for the direct and indirect effects of negative contact were consistently

(though not always significantly) smaller for negative than for positive contact, but

the pattern of results did not change depending on the valence of contact under

consideration.

Overall, these findings suggested that valuing diversity can make a critical

contribution to explaining the links between contact and cognitive outcomes, which

tend to be weaker than the links to affective outcomes (Tropp & Pettigrew, 2005a),

but which are critical to move towards more just and equitable intergroup

arrangements (Dixon et al., 2010).

8.1.3. *Chapter 4: Replication of the mediation in a German probability sample*

The results so far are drawn from samples of psychology undergraduates in

the UK, which limits their generalisability. Therefore, I used data from the German

General Social Survey (ALLBUS) to replicate the key finding of Chapter 3.

ALLBUS uses a cluster-randomised sampling approach ($N = 2,618$) that enables one,

when using the appropriate survey weights, to derive results that generalize to the

German adult population. In addition, the analyses to date did not consider whether

valuing diversity reaches beyond *allophilia*, i.e., a general positive disposition

towards outgroups (Pittinsky et al., 2011). Therefore, I extended the analysis by

including generalised attitudes towards foreigners as a parallel mediator to valuing

diversity in explaining the links between positive and negative contact and approach intentions.

The results replicate the findings of Chapter 3 in another context with yet another outcome measure. When choosing among potential neighbourhoods to live in, positive and negative contact predicted higher/lower intentions to approach the more diverse settings. This effect was mostly mediated through generalised attitudes towards foreigners (i.e., the implied potential neighbours), but valuing diversity made a significant contribution on top of that. The correlation between these two mediators was substantial, $r = .53$, $p < .001$, yet far too small to suggest that the constructs are redundant. Concerning the specific outcome measure considered, it is encouraging that intergroup contact appears to reduce the desire to perpetuate social segregation, and that changes in valuing diversity might contribute to that. This supports the contention that valuing diversity deserves particular attention when it comes to understanding the pathways from contact to positive intergroup relations in a variety of contexts.

8.1.4. *Chapter 5: Impact of a contact intervention on valuing diversity*

Chapter 2 used a longitudinal design to show that intergroup contact experiences predicted changes in the valuing of diversity over time. However, these experiences were highly varied and not deliberately designed, so that they are of limited use when it comes to the application of contact theory to real-world interventions. Therefore, I shifted my focus in Chapter 5 to the evaluation of a residential contact intervention to see whether it affected the valuing of diversity and what moderators predicted differential effects on this outcome. Based on related previous research, I tested the relevance of conversations with other participations about differences (Vezzali et al., 2017) and that of self-expansion orientation (Dys-

Steenbergen et al., 2016). An understanding of such moderators of intervention effects is critical in enabling practitioners to design interventions in a way that maximises their impact. Additionally, expanding upon previous chapters, I contrasted two facets of valuing diversity that had not been systematically distinguished so far, testing whether there were distinct effects on a personal preference for being in a diverse setting and on a recognition of the instrumental value of diversity in terms of more effective problem-solving.

The research took place in the context of the English National Citizen Service (NCS), a residential summer programme for 16-year olds that brings together diverse groups to focus on positive youth development and social action. Using pre-and post-intervention measures from a 2018 impact survey ($N = 705$), I found that participants' preference for working in diverse teams increased over the course of the contact intervention, while the conviction that diversity in teams has instrumental value was not affected by the intervention overall. Regarding moderators, frequent conversations about differences were associated with greater increases in *both* types of diversity, while self-expansion orientation had a weaker effect that was only significant with regard to the preference for working in diverse teams.

The results of Chapter 5 showed that contact interventions can lead to shifts in the valuing of diversity, though possibly more with regard to personal preferences rather than instrumental convictions. The results of the moderation analyses suggested that encouraging more conversations about differences might enhance the intervention's effectiveness.

8.1.5. Chapter 6: Promoting conversations about differences to enhance intervention effects

Building on the findings of Chapter 5, I designed a randomised controlled

trial to take place in the context of the NCS. Given that (self-reported) conversations about differences were associated with greater increases in valuing diversity, I wanted to test whether scheduling such conversations would increase the effectiveness of the intervention. The trial tested the effect of two discussion conditions against an active control, with one condition focused on a general appreciation of diversity in a multicultural society, and another on the instrumental value of diversity in teams. Regarding outcomes, I considered a preference to work in diverse teams separately from an appreciation of diversity in society. The outcomes were based on follow-up measures collected at the end of the NCS programme, three weeks after the discussions took place.

After a period of data collection that was severely limited by organizational disarray on the part of my research partner, I could analyse results from 653 participants. Both intervention conditions did not affect participants' valuing of diversity in society; they both increased the valuing of diversity in teams only for participants from a minority-ethnic background. White participants were unaffected regarding their valuing of diversity, but there was a trend towards reduced interest in future intergroup contact, particularly among those that had discussed the valuing of diversity in teams. While it is encouraging that a brief discussion intervention led to more positive attitudes among minority-ethnic participants, the null or potentially even negative findings for White participants highlight that the design of brief interventions to increase the valuing of diversity is difficult.

8.1.6. Chapter 7: Internal meta-analysis: aggregate evidence for the mediation

In the final empirical chapter, I returned to the idea that valuing diversity is shaped by contact experience, and that it mediates the effects of positive and

negative contact on outgroup attitudes. I aggregated all datasets used in this dissertation (and an additional unpublished dataset), so that the evidence for the model could be assessed based on six samples with 3,981 participants, which contained a total of 57 effect sizes. Since these effect sizes were obviously not independent, a multi-level meta-analytical model had to be used to account for shared variance between measures of the same construct and within each study. To produce a conservative test of the hypotheses and to be able to generalise beyond the studies conducted here, I used a random-effects model.

The results provided support for the hypothesized model. There was significant but limited heterogeneity between the samples, and the meta-analytic structural equation model contained significant indirect effects from both positive and negative contact on outgroup attitudes via valuing diversity. This continued to be the case when the largest sample (the German ALLBUS data) was downweighed. However, the size of the mediation effect was modest, which will be further discussed below.

8.2. Contributions to the literature

The present work contributes to the literature with regard to understanding the relationship between intergroup contact and valuing diversity, with regard to the role of valuing diversity as a mediator of contact effects, and with regard to the nature of contact interventions that might influence the valuing of diversity. These will be discussed in turn.

8.2.1. *Intergroup contact shapes the valuing of diversity*

In the literature to date, valuing diversity has been variously conceptualised as a precedent to intergroup contact (Tropp & Bianchi, 2006), a predictor of

diversity-seeking within contact situations (Bahns et al., 2015), or an outcome of intergroup contact (Harper & Yeung, 2013). The results presented here mostly support the final conceptualisation, showing valuing diversity as an outcome of contact. In Chapter 2, the longitudinal paths from contact to valuing diversity were stronger than the paths from valuing diversity to future contact, and the latter lacked statistical significance. Likewise, Chapter 5 showed that a broad contact intervention led to an increased valuing of diversity, particularly when it included frequent conversations with outgroup members about differences.

Based on these findings, Chapters 3, 4 and 7 treat valuing diversity as an outcome of contact, that can then serve as a mediator to explain other effects. However, the question of whether valuing diversity is a precedent *or* outcome might be too simplified – as Paolini et al. (2018) argue, many relationships in social psychology in general, and intergroup contact in particular, are likely to take the form of virtuous or vicious cycles, where two variables mutually and continuously affect each other. The results presented here do not provide strong evidence for this, but there are some hints. The path coefficients from diversity to future contact in Chapter 2 pointed in the expected direction, with valuing diversity being (non-significantly) associated with more positive and less negative intergroup contact in the future. Similarly, the data presented in Chapter 4 showed that Germans who value diversity express greater willingness to move into diverse neighbourhoods (and thus engage in intergroup contact). I will return to this point when discussing future directions below.

8.2.2. Mediation

Identifying mediators of intergroup contact has become a major focus of contact research. The seminal meta-analysis of mediation by Pettigrew and Tropp

(2008) only uncovered 54 studies that included one of the three most frequent mediators up to that date (intergroup empathy, anxiety and knowledge), while GPIR alone has published 25 studies on the mediation of intergroup contact since then. However, partly because Pettigrew and Tropp found that contact mostly influences affective rather than cognitive outcomes (Tropp & Pettigrew, 2005a), and that affective rather than cognitive variables explained these effects (Pettigrew & Tropp, 2008), the literature has focused on affective mediators such as empathy, anxiety and threat, with relatively little focus on attitudes and beliefs. However, intergroup harmony cannot be the only goal in sight of stark inequalities and structural discrimination (Dixon et al., 2010). Therefore, it was important to me to explore how intergroup contact can affect cognitive outcomes and how these effects come about. The results presented here consistently show that valuing diversity serves as a mediator of the effect of contact on a wide range of outcomes that include attitudes and behavioural intentions, with the exception of bystander intervention intentions in Chapter 3.

8.2.2.1. Size and heterogeneity of the mediation effect

The overall contribution of valuing diversity to explaining the link between intergroup contact and outgroup attitudes is comparable to that of knowledge identified by Pettigrew and Tropp (2008, Table 3) in their meta-analysis of mediation. They found a standardised indirect effect of contact on attitudes through increased knowledge about the outgroup of .017, which is smaller than the estimate for valuing diversity in the meta-analysis presented here, yet within its confidence interval. The indirect effects through intergroup empathy and anxiety, as estimated by Pettigrew and Tropp, were larger (.112 and .090, respectively) than the effect through valuing diversity established here. Thus, the general importance of valuing diversity as a mediator might appear to be relatively low.

However, as the comparison of the mediation models in Chapter 4 suggests, such a conclusion might be premature. The results there suggest that the relative contribution of the mediators depends on the outcome under consideration. While the effect of contact on outgroup attitudes (i.e., affective prejudice) appeared to be primarily mediated through empathy and anxiety, this pattern reversed with regard to cognitive prejudice and policy support where valuing diversity was the dominant mediator. Unfortunately, Pettigrew and Tropp (2008) did not test for a moderation of the mediation by the type of outcome under consideration; including this into a meta-analysis of a broader range of mediators might be a fruitful direction for future research. Based on the data presented here, there is consistent evidence for a significant contribution of valuing diversity to explaining how the effects of contact come about, yet the relative importance of this mediator depends on the outcome under consideration and needs to be assessed in future research.

8.2.2.2. Negative contact

Given the field's historic focus on researching positive rather than negative contact (Barlow et al., 2012), it is not surprising that more is known about the mediation of positive than that of negative contact (Wallrich et al., under review). The results of the various studies presented here suggest that valuing diversity mediates the effects of both positive and negative contact and that it might be relatively more important for explaining the effects of negative contact. In the meta-analysis (Chapter 7), valuing diversity's explanatory power for negative contact, measured as the share of the indirect effect of the total effect, was nearly twice that of positive contact, with 17.7% versus 9.6%. In the mediation models in Chapter 3, this was the case for affective outcomes, but not for cognitive outcomes. The greater importance of valuing diversity for explaining negative contact effects might be explained in reference to the finding that negative contact derives particular potency

from its ability to make intergroup boundaries salient (Paolini et al., 2010). Since valuing diversity concerns the group rather than the individual level, it is not surprising that it is more relevant when group distinctions are salient.

8.2.3. Contact interventions and valuing diversity

In addition to establishing that contact shapes the valuing of diversity over time, the results in Chapter 2 highlighted that an exposure to opportunities for contact is not the same as actual contact. Even though the entry into university life provided increased contact opportunities for most students, the value they placed on diversity declined during the two time points. Therefore, deliberate interventions are needed to realise the potential of intergroup contact, and research into how such interventions can be designed is crucial.

In this dissertation, I have shown that an intense residential contact intervention in the form of the English National Citizen service leads to an increase in participants' preference for seeking diversity, at least in the context of teamwork (Chapter 5). This was particularly the case for participants who engaged in conversations about differences during the programme, which then provided an inspiration for designing brief discussion-based interventions that were then the focus of a randomised controlled trial (Chapter 6). The results of the trial were not in line with hypotheses, yet tentative conclusions can be drawn: it appears that participants from a minority-status background benefit particularly from an explicit discussion of the value of diversity within an intergroup context, as they reported a greater desire to work in diverse teams following either intervention. Conversely, this outcome appears harder to achieve for White participants, who reported no changes in their valuing of diversity after the intervention and even showed a (marginally significant) tendency towards a reduced interest in future intergroup contact. This

adds further weight to earlier findings that simple celebrations of diversity might have unintended effects on majority-status participants, for instance when multiculturalism primes led to a greater endorsement of Trump (Osborn et al., 2020). This might be because discussions of diversity can be perceived as exclusionary by Whites (Plaut et al., 2011), so that further intervention research is needed.

8.3. Reflection on research design

Some decisions in the focus and design of my research deserve further discussion. This partly concerns a reflection on lesson learned, and partly a justification of atypical choices.

8.3.1. The case for the use of diverse measures

Each study presented in this dissertation operationalised the valuing of diversity in a somewhat different way, while the measurements of contact frequency also varied. This was partly driven by the need to rely on pre-existing data (the German General Social Survey, Chapter 4) and a limited ability to make changes to the questionnaires used to evaluate the NCS experience (Chapters 5 and 6). However, it was also an intentional choice, as I believe that the use of a variety of measures enables the generation of more valid inferences about underlying constructs. The argument for this would be analogous to the argument for stimulus sampling (Wells & Windschitl, 1999). The case for stimulus sampling starts from the observation that experimental studies based on differences in reactions to two stimuli, e.g., a male and a female experimenter, likely suffer from an issue of construct validity. If one wishes to test whether people maintain greater distance from outgroup than ingroup members, for instance, testing whether "John", who happens to be Black, among many other characteristics, is afforded more space than "George" who happens to be White, yields severely limited results with regard to the

question of interest, even if the difference is found in several samples. Likewise, it appears evident to me that a set of studies that all rely on the same measure for a complex construct will accumulate a consistent systematic error. Even survey measures with strong construct validity will almost inevitably tap into more than one latent construct, which will then enter analyses as invisible confounds. A similar set of studies, which each rely on different measures that all aim to measure the same latent construct while overlapping with different separate constructs, will be more likely to yield reliable insights into that construct.

Evidently, this argument mostly holds with regard to significant and consistent findings, which then jointly strengthen results. Where inconsistencies arise, they become more difficult to explain, since differences may always be due to the specific constructs the measures tap into. However, especially in early-stage research, the increase in robustness seems to outweigh the potential decrease in the accuracy of speculations, which, after all, is all one can offer regarding null-findings in the context of frequentist statistics anyway. Nevertheless, a next step for this line of research would be the development and validation of a longer scale to measure valuing diversity, with the attendant opportunities to consider the dimensionality of the construct and its discriminant validity.

8.3.2. Reproducibility and Open Science

Over the course of my research journey and of a teaching appointment for quantitative research methods, I increasingly came to realise the importance of reproducibility of research, and of Open Science standards. Even before worrying about replicability, full transparency into the data, methods and statistical analyses appears essential, as it allows for rigorous quality checks and supports cumulative science (Munafò et al., 2017). Additionally, a commitment to open science

encompasses the use of open research software, open access to research outputs and open licencing of materials (Bezjak et al., 2018). [22]

A commitment to open science principles (and in my case, also to personal sanity) precludes the use of point-and-click interfaces such as the menu-based approach to SPSS for final analyses. While they are helpful for data exploration, they often make it impossible even for the researcher to reproduce results, at least in the absence of very detailed pre-registrations or lab manuals. Code-based approaches, be they SPSS syntax or a full-fledged statistical programming language, offer much greater transparency and reproducibility, so that I decided to transition towards them. Given that R is an open-source software requirement available to anyone, rather than only to those affiliated to institutions that are willing to pay for licences, I decided to learn how to use it, which is enabled by the availability of many high quality open educational materials, including entire textbooks.

In that, I aimed to create tables, graphs and some textual reports directly from the data, in order to avoid misreporting. This is important because research has shown that half of published psychology articles contains p-values that are misreported, i.e. inconsistent with the reported test statistics and degrees of freedom (Nuijten et al., 2016). Given that many other reports cannot so be checked so easily, it is likely that misreported descriptive statistics are even more frequent. As there is no relationship between the frequency of such errors and journal's data sharing

---

[22] An additional component of Open Science that I am passionate about, but that is of limited relevance to this research project, are open educational materials. It appears highly inefficient that thousands of university teachers reinvent the same materials, and unjust that the fruits of this taxpayer-funded work are only available beyond expensive paywalls. While Open Access to journal articles is presently only achievable for individual researchers with substantial resources, making teaching materials available is an individual choice. I have made that choice, for instance, by making most of my lectures available on YouTube and providing the course materials (expect for admin-related matters) on a freely accessible webpage rather than on a proprietary platform like Moodle. Additionally, I licence my materials under a Creative Commons licence whenever possible. For a graduate level Core Quantitative Research Methods module, for instance, the materials can be found here: https://github.com/LukasWallrich/GoldCoreQuants

policies (Nuijten et al., 2017), it appears likely that these are the primarily the results

of honest mistakes rather than motivated 'adjustments.' In order to create

reproducible tables, I relied on the *gt* package (Iannone et al., 2020) and made some

minor contributions to the development of the *modelsummary* package (Arel-

Bundock, 2020). To create moderation and mediation plots, I built on the

DiagrammeR package (Iannone, 2020), while all other plots were created with

ggplot2 (Wickham, 2016). Furthermore, I created R functions to fully automate

recurring tasks such as the creation of correlation and regression tables, mediation

plots, the reporting of simple slopes, or the setup of a new analysis workflow and

made these accessible to other researchers by releasing the *timesaveR* package on

GitHub (Wallrich, 2021).[23]

The data, analysis code and materials for the studies reported in this

dissertation – with the exception of some datasets that I cannot release publicly due

to licencing requirement – are available on GitHub, from where anyone wanting to

reproduce or extend an analysis can directly download the entire repository

(https://github.com/LukasWallrich/PhD_thesis). For enhanced discoverability, this is

also available as a project on the OSF (https://osf.io/f47cb/).

8.3.3. *Robustness of results and the issue of forking paths*

At least since the reproducibility project found that only 25% of studies in

social psychology could be successfully replicated (Open Science Collaboration,

2015), there has been growing concern regarding the robustness of findings. A key

reason that has been identified for the excessive false-positive rate is that data

---

[23] There are various packages that specifically aim to support the reporting of research results in APA style, particularly the *apa* (Gromer, 2020) and *papaja* (Aust & Barth, 2020) packages. While I used them occasionally, they typically did not meet my specific requirements, even though I will follow the future development of *papaja* closely. An exciting newcomer to this field is the *report*-package (Makowski, Dominique et al., 2020) that is part of the broader *easystats* project, which I will attempt to rely on more in future projects.

analysts wander through "a garden of forking paths", making dozens of choices regarding analytical methods that are individually inconsequential yet can collectively torture the data to confess what the analysts want to hear (Gelman & Loken, 2014). To limit this risk, I minimised the number of choices by refraining from exclusions and data transformations wherever possible. In addition, I attempted to identify the major choices that individually had the potential to alter the results and included the alternative approach as robustness checks. While I expect this to help, an additional effective tool are pre-registrations, which – in retrospect – I did not use sufficiently.

Unfortunately, I only came to realise the importance of pre-registrations towards the end of Year 2 of my PhD. I then pre-registered the field experiment with the NCS (Chapter 6) but based the analysis plan on the targeted sample of at least 5,000 participants, rather than the sample of just over 600 participants that I obtained after my partner organisation entered an existential struggle half-way through the period of data collection. While I attempted to follow the pre-registration in as far as possible, the required changes introduced so many decision points that the value of the pre-registration was severely limited.

8.3.4. Student samples and WEIRD participants

Psychological research has been criticized for its excessive focus on student samples and WEIRD participants – those from majority-White, educated, industrialised and rich countries (e.g., Henrich et al., 2010). While university student samples are convenient and can be studied with strong experimental controls, they are clearly not representative of the population at large. Therefore, I started my research with two student samples, but then also considered a representative population sample and two youth samples drawn from a programme that recruits

broadly and is thus less skewed regarding social class and educational attainment. The converging results give me some confidence that my findings can be reasonably – though tentatively – generalised to Europeans at large.

However, I hardly reached beyond WEIRD samples. As briefly discussed in Chapter 7, I attempted a replication of the mediation model in India, which resulted in null-findings. However, I could not obtain a sufficiently large sample to reliably test whether the divergence was significant – given the data as it is, the test was underpowered, but the divergences were almost as far from statistical significance as possible, with $p = .94$. However, the observed pattern still points to the need for future research on the relationship between intergroup contact and diversity beliefs in non-WEIRD samples.

8.3.5. *Difficulties with applied research*

Initially, I wanted to focus the PhD primarily on collaborative research with the NCS, to identify how contact interventions with youth could effectively approach the theme of diversity. However, despite initial enthusiasm and later contractual agreements, the research was first postponed by a year and then only implemented in a severely curtailed manner. The postponement first occurred because the NCS Trust (the government entity that funded The Challenge, my research partner) enforced the prioritisation of another set of field experiments. When the research then took place, only a very limited number of items could be used to evaluate the intervention, because The Challenge was only concerned *whether* the intervention worked, and not willing to place an extra burden on survey respondents to assess *why* it might work. To top things off, they entered a period of organisational disarray in the middle of the data collection period, which ended in bankruptcy, thereby curtailing data collection early.

However, when the randomised controlled trials were implemented the effects of discussion-based interventions on valuing diversity were nuanced and weaker than expected (see Chapter 6).[24] Overall, this experience both highlights the difficulties of applied research and the need to conduct more of it, in order to ensure that the design of contact interventions maximises their impact.

8.4. Limitations and future directions

Before concluding, it is worth reflecting on key limitations of the empirical work presented in this thesis and to consider emerging directions for future research.

8.4.1. The meaning of diversity

Most of the research presented here asked participants about their attitudes towards diversity, or diverse teams, without further specifying the dimensions of diversity under consideration. Previous research sometimes asked about a single dimension such as racial and ethnic diversity (Tropp & Bianchi, 2006), while others used scales that included a wide range of dimensions yet still turned out to be reliable indicators of an underlying construct (e.g., Bahns et al, 2015, asked about race, ethnicity, sexual orientation, nationality, age and physical ability, which yielded a consistent scale). Asking about diversity without providing such a detailed framing will likely lead participant towards considering the dimensions they find most salient, which will yield responses that refer to diverse attitude objects yet tap into a more subjectively meaningful construct.[25]

---

[24] The collaboration also required an alignment of research priorities, which led to the inclusion of an additional randomised controlled trial regarding sessions to promote bystander intervention. The results are reported in Wallrich, Palmer et al. (2021, under review), and again revealed asymmetries between majority- and minority-ethnic status participants, this time in the opposite direction. An activity that promoted self-disclosure around experiences of discrimination mobilised majority-status participants to take action, while it left minority-status participants unaffected.

[25] A similar decision needs to be made frequently in contact research when it comes to measuring positive and negative contact. Many studies, including those in Chapters 4, 5 and 6, ask

The results presented in this thesis show that valuing diversity, as perceived by participants, is typically shaped by intergroup contact experiences, and typically predicts the outcomes desired for positive contact. However, especially given the ambiguous findings for White participants in the randomised controlled trial of discussion-based interventions, an important future direction is to explore what the subjective meanings of 'diversity' in general and 'valuing diversity' in particular are and how they differ between majority-status and minority-status participants. This would best be approached through qualitative interviews or focus groups. Furthermore, future research that distinguishes between the value placed on different dimensions of diversity would be able to test to what extent valuing diversity might explain the secondary transfer effect (Pettigrew, 2009), i.e. the generalisation of the effects of intergroup contact beyond the specific groups that are encountered. As conceptualised here, a generic increase in the value of diversity should be able to explain a link between contact with one outgroup and attitudes towards another, yet this needs to be tested.

8.4.2. Dynamic relationships

As discussed above, the findings presented here suggest that valuing diversity is an outcome of intergroup contact that can help explain how other outcomes come about. While the longitudinal evidence in Chapter 2 suggests that this direction is dominant, there are (non-significant) indications that valuing diversity might influence future contact, and Chapter 4 shows that valuing diversity predicts

---

participants to simply report the frequency of 'positive' and 'negative' 'contact', leaving it up to them to fill the three terms with meaning, while others ask for more specific reports of experiences, such as the frequency of being welcomed versus excluded (e.g., Reimer et al., 2017, whose measure was used in Chapters 2 and 3 here). Here, the latter approach is also likely to result in data that more closely describes reality and evenly weighs types of positive and negative contact, while the former taps more strongly into participants' subjective experiences. To date, I am not aware of a study that systematically compared these different measures in terms of their relationship and respective explanatory power, so that this would also appear to be a fruitful direction for future research.

intergroup approach intentions. Therefore, there is the potential for a more complex and dynamic relationship between intergroup contact and valuing diversity, in which they reinforce each other, giving rise to a virtuous cycle between positive contact and valuing diversity. While such relationships are often plausible in the field of intergroup contact, they are rarely tested. While a recent review issued a call to address this omission (Paolini et al., 2018), effective tests of models that involve reciprocal relationships require large longitudinal datasets that span at least three timepoints. This was beyond the scope of what I could collect within the confines of this PhD, yet it provides a potentially fruitful direction for future research.

8.4.3. *Evidence for causality in mediation models*

A key contribution of this dissertation was to establish that valuing diversity can be considered as a mediator of contact effects on a variety of intergroup outcomes. However, none of the mediation models were based on experimental data, so that causal inferences are limited. To assess whether the suggested causal relationships are plausible, one needs to consider the order of the variables and the likelihood that omitted variables explain the observed associations. Regarding the ordering, the longitudinal data presented in Chapter 2 suggests that the link from contact to valuing diversity is stronger than the potential link from valuing diversity to contact, while the assessment of the overall effect of the NCS (Chapter 5) also shows that there is an effect of contact (the independent variable) on valuing diversity. Similarly, it appears clear from previous research that contact predicts prejudice, at least as strongly as the reverse (Binder et al., 2009; Levin et al., 2003; Pettigrew, 1997), and it is plausible to assume that this extends to behavioural intentions such as policy support and bystander intervention intentions, even though I am not aware of research on that. The relationship in the mediation models put forward that deserves the greatest scrutiny is that between valuing diversity and other

outcomes of intergroup contact. Here the proposed placement of valuing diversity as a mediator appears plausible, yet further longitudinal or experimental research is needed to test whether the (primary) path leads from valuing diversity to intergroup attitudes and behavioural intentions rather than the reverse.

8.4.4. Boundary conditions

Except for the exploratory study in India, all studies took place in comparatively positive intergroup contexts, where the valuing of diversity was present as a social norm. It appears possible that the relationship between contact and valuing diversity might differ in the absence of such a norm, so that future research should consider more negative intergroup contexts to see whether this moderates effects. Additionally, the research presented here focused on contact between ethnic groups or with foreigners. While these represent some of the most urgent divides, and those most prominently considered in intergroup contact research, other outgroups should be considered in future groups, particularly to test whether the visibility of group boundaries makes a difference. For instance, it might be possible that contacts with people who are visibly different might more naturally lead towards a recognition and appreciation of diversity than contact with those attitudinally different would do (e.g., opposing partisan). This specific point was hinted at by Bahns (2017) who found that in naturally occurring dyads, valuing diversity was associated with people associating with those of different religions, ethnicities and sexual orientation, but not nationalities or attitudes. However, further research is needed.

8.4.5. Development of interventions

From its very start, research on intergroup contact had an applied aim – that of improving intergroup relations. Therefore, research on how its findings can be

applied in interventions should be a part of any research agenda. Here, Chapter 5 tested whether an existing contact intervention increased the valuing of diversity and explored how this effect might be strengthened, while Chapter 6 tested specific brief interventions. The results of Chapter 5 indicate that conversations about differences are likely to help in bringing about more valuing diversity, yet Chapter 6 provides evidence that this might only be the case for ethnic-minority participants, while confirming earlier findings that multiculturalism primes can backfire (Osborn et al., 2020). Therefore, further intervention research is needed, ideally informed by a better understanding of the subjective meaning White participants attached to the concepts of diversity and valuing diversity.

8.4.6. Minority perspectives

Finally, this dissertation largely focused on the perspective of majority-status participants. While their intergroup attitudes are arguably the greater obstacle to positive intergroup outcomes, minority-status groups' perspectives should also be considered. In this dissertation, I could only do so in the context of the National Citizen's Service (Chapters 5 and 6). In Chapter 5, there were no differences between minority- and majority-status participants in either their baselines or the intervention effects. In Chapter 6, however, ethnic-minority participants responded more favourably to the interventions that aimed to celebrate diversity. This in in line with prior findings that minority-status groups are particularly favourable towards identity-conscious approaches (Levin et al., 2003). However, research into the meaning minority-status participants attach to the concept of valuing diversity could help to understand the different effects better and to ensure that adaptations of interventions that address the concerns of majority-status participants do not undermine their effectiveness among minority-status participants.

8.5. Conclusions

This thesis highlights a novel pathway from intergroup contact to prejudice reduction and (intended) behaviour change. Positive intergroup contact increases the valuing of diversity, which in turn is associated with improved outgroup attitudes and intentions to promote inclusive behaviours and policies. Valuing diversity also increases the openness to future contact, which suggests that conditions for contact might be created that give rise to a virtuous cycle of self-reinforcing increases in contact and valuing diversity.

References

Abbott, N., & Cameron, L. (2014). What Makes a Young Assertive Bystander? The Effect of Intergroup Contact, Empathy, Cultural Openness, and In-Group Bias on Assertive Bystander Intervention Intentions: Predictors of Bystander Intervention Intentions. *Journal of Social Issues*, *70*(1), 167–182. https://doi.org/10.1111/josi.12053

Abdeleli, A. (2019). *Switzerland's controversial minaret ban, ten years on*. SWI Swissinfo.Ch. https://www.swissinfo.ch/eng/political-history_the-controversial-minaret-ban--ten-years-on/45399822

Aberson, C. L., & Haag, S. C. (2007). Contact, Perspective Taking, and Anxiety as Predictors of Stereotype Endorsement, Explicit Attitudes, and Implicit Attitudes. *Group Processes & Intergroup Relations*, *10*(2), 179–201. https://doi.org/10.1177/1368430207074726

Abrams, D., Rutland, A., Cameron, L., & Ferrell, J. (2007). Older but wilier: In-group accountability and the development of subjective group dynamics. *Developmental Psychology*, *43*(1), 134–148. https://doi.org/10.1037/0012-1649.43.1.134

Abramyan, H., & Alexander, G. (2020). A distinction with a difference: The influence of social concerns and racial prejudice on support for amnesty. *The Social Science Journal*, 1–13. https://doi.org/10.1016/j.soscij.2019.05.002

Adesokan, A. A., Ullrich, J., van Dick, R., & Tropp, L. R. (2011). Diversity Beliefs as Moderator of the Contact–Prejudice Relationship. *Social Psychology*, *42*(4), 271–278. https://doi.org/10.1027/1864-9335/a000058

Ahn, S., Ames, A. J., & Myers, N. D. (2012). A Review of Meta-Analyses in Education: Methodological Strengths and Weaknesses. *Review of Educational Research*, *82*(4), 436–476.

https://doi.org/10.3102/0034654312458162

Alba, R., & Duyvendak, J. W. (2019). What about the mainstream? Assimilation in super-diverse times. *Ethnic and Racial Studies*, *42*(1), 105–124. https://doi.org/10.1080/01419870.2017.1406127

Alesina, A., & Zhuravskaya, E. (2011). Segregation and the Quality of Government in a Cross Section of Countries. *American Economic Review*, *101*(5), 1872–1911. https://doi.org/10.1257/aer.101.5.1872

Allport, G. W. (1954). *The Nature of Prejudice*. Addison.

Amichai-Hamburger, Y., & McKenna, K. Y. A. (2006). The Contact Hypothesis Reconsidered: Interacting via the Internet. *Journal of Computer-Mediated Communication*, *11*(3), 825–843. https://doi.org/10.1111/j.1083-6101.2006.00037.x

Amin, A. (2016). Ethnicity and the Multicultural City: Living with Diversity: *Environment and Planning A*. https://doi.org/10.1068/a3537

Amir, Y. (1969). Contact hypothesis in ethnic relations. *Psychological Bulletin*, *71*(5), 319.

Andersen, H. S. (2017). Selective moving behaviour in ethnic neighbourhoods: White flight, white avoidance, ethnic attraction or ethnic retention? *Housing Studies*, *32*(3), 296–318. https://doi.org/10.1080/02673037.2016.1208161

Andersson, E. K., Malmberg, B., Costa, R., Sleutjes, B., Stonawski, M. J., & de Valk, H. A. G. (2018). A Comparative Study of Segregation Patterns in Belgium, Denmark, the Netherlands and Sweden: Neighbourhood Concentration and Representation of Non-European Migrants. *European Journal of Population*, *34*(2), 251–275. https://doi.org/10.1007/s10680-018-9481-5

Andersson, R. (2013). Reproducing and reshaping ethnic residential segregation in

Stockholm: The role of selective migration moves. *Geografiska Annaler: Series B, Human Geography*, *95*(2), 163–187.

Apfelbaum, E. P., Norton, M. I., & Sommers, S. R. (2012). Racial Color Blindness: Emergence, Practice, and Implications. *Current Directions in Psychological Science*, *21*(3), 205–209. https://doi.org/10.1177/0963721411434980

Apfelbaum, E. P., Sommers, S. R., & Norton, M. I. (2008). Seeing race and seeming racist? Evaluating strategic colorblindness in social interaction. *Journal of Personality and Social Psychology*, *95*(4), 918–932. https://doi.org/10.1037/a0011990

Arel-Bundock, V. (2020). *Modelsummary: Summary Tables and Plots for Statistical Models and Data: Beautiful, Customizable, and Publication-Ready*. https://vincentarelbundock.github.io/modelsummary/

Árnadóttir, K., Lolliot, S., Brown, R., & Hewstone, M. (2018). Positive and negative intergroup contact: Interaction not asymmetry: Positive and negative contact. *European Journal of Social Psychology*, *48*(6), 784–800. https://doi.org/10.1002/ejsp.2365

Asbrock, F., Kauff, M., Issmer, C., Christ, O., Pettigrew, T. F., & Wagner, U. (2012). Kontakt hilft–auch wenn die Politik es nicht immer leicht macht. *Deutsche Zustände. Folge*, *10*, 199–219.

Aust, F., & Barth, M. (2020). *papaja: Create APA manuscripts with R Markdown*. https://github.com/crsh/papaja

Awad, G. H., Cokley, K., & Ravitch, J. (2005). Attitudes Toward Affirmative Action: A Comparison of Color-Blind Versus Modern Racist Attitudes. *Journal of Applied Social Psychology*, *35*(7), 1384–1399. https://doi.org/10.1111/j.1559-1816.2005.tb02175.x

Ayers, W. (1995). Social imagination: A conversation with Maxine Greene.

International Journal of Qualitative Studies in Education, *8*(4), 319–328. https://doi.org/10.1080/0951839950080401

Bagci, S. C., Kumashiro, M., Rutland, A., Smith, P. K., & Blumberg, H. (2017). Cross-ethnic friendships, psychological well-being, and academic outcomes: Study of South Asian and White children in the UK. *European Journal of Developmental Psychology*, *14*(2), 190–205. https://doi.org/10.1080/17405629.2016.1185008

Bagci, S. C., Rutland, A., Kumashiro, M., Smith, P. K., & Blumberg, H. (2014). Are minority status children's cross-ethnic friendships beneficial in a multiethnic context? *British Journal of Developmental Psychology*, *32*(1), 107–115. https://doi.org/10.1111/bjdp.12028

Bagci, S. C., & Turnuklu, A. (2018). Intended, Unintended, and Unknown Consequences of Contact. *Social Psychology*, *50*(1), 7–23. https://doi.org/10.1027/1864-9335/a000355

Bahns, A. J. (2017). Preference, opportunity, and choice: A multilevel analysis of diverse friendship formation. *Group Processes & Intergroup Relations*, *22*(2), 233–252. https://doi.org/10.1177/1368430217725390

Bahns, A. J., Springer, L. S., & The, C. (2015). Fostering diverse friendships: The role of beliefs about the value of diversity. *Group Processes & Intergroup Relations*, *18*(4), 475–488. https://doi.org/10.1177/1368430214566893

Banting, K., & Kymlicka, W. (2006). *Multiculturalism and the Welfare State: Recognition and Redistribution in Contemporary Democracies*. OUP Oxford.

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and Nurture in Own-Race Face Processing. *Psychological Science*, *17*(2), 159–163. https://doi.org/10.1111/j.1467-9280.2006.01679.x

Barlow, F. K., Paolini, S., Pedersen, A., Hornsey, M. J., Radke, H. R. M., Harwood,

J., Rubin, M., & Sibley, C. G. (2012). The Contact Caveat: Negative Contact Predicts Increased Prejudice More Than Positive Contact Predicts Reduced Prejudice. *Personality and Social Psychology Bulletin*, *38*(12), 1629–1643. https://doi.org/10.1177/0146167212457953

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, *34*(1), 215–220. https://doi.org/10.1093/ije/dyh299

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Batson, C. D., & Shaw, L. L. (1991). Evidence for Altruism: Toward a Pluralism of Prosocial Motives. *Psychological Inquiry*, *2*(2), 107.

Bauer, D. J., & Sterba, S. K. (2011). Fitting multilevel models with ordinal outcomes: Performance of alternative specifications and methods of estimation. *Psychological Methods*, *16*(4), 373.

BBC News. (2010, October 17). Merkel: "Multiculture has failed." *BBC News*. https://www.bbc.com/news/world-europe-11559451

BBC News. (2011, February 5). Multiculturalism has failed—PM. *BBC News*. https://www.bbc.com/news/uk-politics-12371994

BBC News. (2021, March 7). Switzerland referendum: Voters support ban on face coverings in public. *BBC News*. https://www.bbc.com/news/world-europe-56314173

Bell, J. M., & Hartmann, D. (2007). Diversity in Everyday Discourse: The Cultural Ambiguities and Consequences of "Happy Talk." *American Sociological Review*, *72*(6), 895–914. https://doi.org/10.1177/000312240707200603

Bentolila, S., Michelacci, C., & Suarez, J. (2010). Social Contacts and Occupational

Choice. *Economica*, *77*(305), 20–45. https://doi.org/10.1111/j.1468-0335.2008.00717.x

Bezjak, S., Clyburne-Sherin, A., Conzett, P., Fernandes, P. L., Görögh, E., Helbig, K., Kramer, B., Labastida, I., Niemeyer, K., Psomopoulos, F., & others. (2018). *The open science training handbook*. [sn]. https://book.fosteropenscience.eu/

Binder, J., Zagefka, H., Brown, R., Funke, F., Kessler, T., Mummendey, A., Maquil, A., Demoulin, S., & Leyens, J.-P. (2009). Does contact reduce prejudice or does prejudice reduce contact? A longitudinal test of the contact hypothesis among majority and minority groups in three european countries. *Journal of Personality and Social Psychology*, *96*(4), 843–856. https://doi.org/10.1037/a0013470

Birtel, M. D., Reimer, N. K., Wölfer, R., & Hewstone, M. (2020). Change in school ethnic diversity and intergroup relations: The transition from segregated elementary to mixed secondary school for majority and minority students. *European Journal of Social Psychology*, *50*(1), 160–176. https://doi.org/10.1002/ejsp.2609

Blackaby, D. H., Leslie, D. G., Murphy, P. D., & O'Leary, N. C. (2005). Born in Britain: How are native ethnic minorities faring in the British labour market? *Economics Letters*, *88*(3), 370–375. https://doi.org/10.1016/j.econlet.2005.03.008

Bolt, G., van Kempen, R., & van Ham, M. (2008). Minority Ethnic Groups in the Dutch Housing Market: Spatial Segregation, Relocation Dynamics and Housing Policy. *Urban Studies*, *45*(7), 1359–1384. https://doi.org/10.1177/0042098008090678

Boustan, L. P. (2010). Was Postwar Suburbanization "White Flight"? Evidence from

the Black Migration. *The Quarterly Journal of Economics*, *125*(1), 417–443.

https://doi.org/10.1162/qjec.2010.125.1.417

Braddock, J. H., & McPartland, J. M. (1989). Social-Psychological Processes That

Perpetuate Racial Segregation: The Relationship Between School and

Employment Desegregation. *Journal of Black Studies*, *19*(3), 267–289.

https://doi.org/10.1177/002193478901900301

Bråmå, Å. (2006). "White Flight"? The Production and Reproduction of Immigrant

Concentration Areas in Swedish Cities, 1990-2000. *Urban Studies*, *43*(7),

1127–1146. https://doi.org/10.1080/00420980500406736

Brannon, R., Cyphers, G., Hesse, S., Hesselbart, S., Keane, R., Schuman, H.,

Viccaro, T., & Wright, D. (1973). Attitude and Action: A Field Experiment

Joined to a General Population Survey. *American Sociological Review*, *38*(5),

625–636. https://doi.org/10.2307/2094412

Brief, A. P., Dietz, J., Cohen, R. R., Pugh, S. D., & Vaslow, J. B. (2000). Just doing

business: Modern racism and obedience to authority as explanations for

employment discrimination. *Organizational Behavior & Human Decision

Processes*.

Brown, R., González, R., Zagefka, H., Manzi, J., & Čehajić, S. (2008). Nuestra

culpa: Collective guilt and shame as predictors of reparation for historical

wrongdoing. *Journal of Personality and Social Psychology*, *94*(1), 75–90.

https://doi.org/10.1037/0022-3514.94.1.75

Brown, R., & Hewstone, M. (2005). An integrative theory of intergroup contact.

Advances in Experimental Social Psychology, *37*, 255–343.

https://doi.org/10.1016/s0065-2601(05)37005-5

Brown, R., Vivian, J., & Hewstone, M. (1999). Changing attitudes through

intergroup contact: The effects of group membership salience. *European*

Journal of Social Psychology, *29*(5-6), 741–764.

https://doi.org/10.1002/(SICI)1099-0992(199908/09)29:5/6<741::AID-

EJSP972>3.0.CO;2-8

Burgess, S., Wilson, D., & Lupton, R. (2005). Parallel Lives? Ethnic Segregation in

Schools and Neighbourhoods. *Urban Studies*, *42*(7), 1027–1056.

https://doi.org/10.1080/00420980500120741

Buuren, S. van, & Groothuis-Oudshoorn, K. (2010). mice: Multivariate imputation

by chained equations in R. *Journal of Statistical Software*, *45*(3), 1–68.

https://doi.org/10.18637/jss.v045.i03

Cappellari, L., & Tatsiramos, K. (2011). *Friends' networks and job finding rates*

[Working Paper No. 2011-21, Institute of Social and Economic Research

(ISER)]. https://www.econstor.eu/bitstream/10419/65908/1/667475869.pdf

Chen, X., Leung, A. K. -y., Yang, D. Y.-J., Chiu, C., Li, Z., & Cheng, S. Y. Y.

(2016). Cultural Threats in Culturally Mixed Encounters Hamper Creative

Performance for Individuals With Lower Openness to Experience. *Journal of

Cross-Cultural Psychology*, *47*(10), 1321–1334.

https://doi.org/10.1177/0022022116641513

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for

Testing Measurement Invariance. *Structural Equation Modeling: A

Multidisciplinary Journal*, *9*(2), 233–255.

https://doi.org/10.1207/S15328007SEM0902_5

Cheung, M. W.-L. (2015a). MetaSEM: An R Package for Meta-Analysis using

Structural Equation Modeling. *Frontiers in Psychology*, *5*(1521).

https://doi.org/10.3389/fpsyg.2014.01521

Cheung, M. W.-L. (2015b). *Meta-Analysis: A Structural Equation Modeling

Approach*. John Wiley & Sons.

Cheung, M. W.-L. (2020). *Synthesizing Indirect Effects in Mediation Models with Meta-Analytic Methods* [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/df6jp

Christensen, R. H. B. (2019). *ordinal—Regression Models for Ordinal Data*.

Church, C. H. (2008). The Swiss Elections of 21 October 2007: Consensus Fights Back. *West European Politics*, *31*(3), 608–623. https://doi.org/10.1080/01402380801939875

Cialdini, R. B. (2007). Descriptive Social Norms as Underappreciated Sources of Social Control. *Psychometrika*, *72*(2), 263–268. https://doi.org/10.1007/s11336-006-1560-6

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Routledge.

Corrigan, P. W., Rowan, D., Green, A., Lundin, R., River, P., Uphoff-Wasowski, K., White, K., & Kubiak, M. A. (2002). Challenging Two Mental Illness Stigmas: Personal Responsibility and Dangerousness. *Schizophrenia Bulletin*, *28*(2), 293–309. https://doi.org/10.1093/oxfordjournals.schbul.a006939

Crisp, R. J., Stone, C. H., & Hall, N. R. (2006). Recategorization and Subgroup Identification: Predicting and Preventing Threats From Common Ingroups. *Personality and Social Psychology Bulletin*, *32*(2), 230–243. https://doi.org/10.1177/0146167205280908

Crowder, K., & South, S. J. (2008). Spatial Dynamics of White Flight: The Effects of Local and Extralocal Racial Conditions on Neighborhood Out-Migration. *American Sociological Review*, *73*(5), 792–812. https://doi.org/10.1177/000312240807300505

Crystal, D. S., Killen, M., & Ruck, M. (2008). It is who you know that counts: Intergroup contact and judgments about race-based exclusion. *British Journal of Developmental Psychology*, *26*(1), 51–70.

https://doi.org/10.1348/026151007X198910

Cutler, D. M., Glaeser, E. L., & Vigdor, J. L. (1999). The Rise and Decline of the

American Ghetto. *Journal of Political Economy*, *107*(3), 455–506.

https://doi.org/10.1086/250069

Davies, K., Tropp, L. R., Aron, A., Pettigrew, T. F., & Wright, S. C. (2011). Cross-

Group Friendships and Intergroup Attitudes: A Meta-Analytic Review.

Personality and Social Psychology Review, *15*(4), 332–351.

https://doi.org/10.1177/1088868311411103

Dessel, A. B., Goodman, K. D., & Woodford, M. R. (2017). LGBT discrimination on

campus and heterosexual bystanders: Understanding intentions to intervene.

Journal of Diversity in Higher Education, *10*(2), 101–116.

https://doi.org/10.1037/dhe0000015

Devine, P. G., Plant, E. A., Amodio, D. M., Harmon-Jones, E., & Vance, S. L.

(2002). The regulation of explicit and implicit race bias: The role of

motivations to respond without prejudice. *Journal of Personality and Social

Psychology*, *82*(5), 835–848. https://doi.org/10.1037//0022-3514.82.5.835

Dickter, C. L., Kittel, J. A., & Gyurovski, I. I. (2012). Perceptions of non-target

confronters in response to racist and heterosexist remarks. *European Journal

of Social Psychology*, *42*(1), 112–119. https://doi.org/10.1002/ejsp.855

Dixon, J., Durrheim, K., & Tredoux, C. (2005). Beyond the Optimal Contact

Strategy: A Reality Check for the Contact Hypothesis. *American

Psychologist*, *60*(7), 697–711. https://doi.org/10.1037/0003-066X.60.7.697

Dixon, J., Durrheim, K., & Tredoux, C. (2007). Intergroup contact and attitudes

toward the principle and practice of racial equality. *Psychological Science*,

18(10), 867–872.

Dixon, J., Tredoux, C., Davies, G., Huck, J., Hocking, B., Sturgeon, B., Whyatt, D.,

Jarman, N., & Bryan, D. (2020). Parallel lives: Intergroup contact, threat, and the segregation of everyday activity spaces. *Journal of Personality and Social Psychology*, *118*(3), 457–480. https://doi.org/10.1037/pspi0000191

Dixon, J., Tropp, L. R., Durrheim, K., & Tredoux, C. (2010). "Let Them Eat Harmony": Prejudice-Reduction Strategies and Attitudes of Historically Disadvantaged Groups. *Current Directions in Psychological Science*, *19*(2), 76–80. https://doi.org/10.1177/0963721410363366

Dovidio, J. F., & Gaertner, S. L. (2000). Aversive Racism and Selection Decisions: 1989 and 1999. *Psychological Science*, *11*(4), 315–319. https://doi.org/10.1111/1467-9280.00262

Dovidio, J. F., Gaertner, S. L., Ufkes, E. G., Saguy, T., & Pearson, A. R. (2016). Included but Invisible? Subtle Bias, Common Identity, and the Darker Side of "We." *Social Issues and Policy Review*, *10*(1), 6–46. https://doi.org/10.1111/sipr.12017

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the Nature of Prejudice: Automatic and Controlled Processes. *Journal of Experimental Social Psychology*, *33*(5), 510–540. https://doi.org/10.1006/jesp.1997.1331

Dovidio, J. F., Love, A., Schellhaas, F. M. H., & Hewstone, M. (2017). Reducing intergroup bias through intergroup contact: Twenty years of progress and future directions. *Group Processes & Intergroup Relations*, *20*(5), 606–620. https://doi.org/10.1177/1368430217712052

Durkheim, E. (1951). Suicide: A study in sociology (JA Spaulding & G. Simpson, trans.). *Glencoe, IL: Free Press.(Original Work Published 1897)*.

Durrheim, K. (2005). Socio-Spatial Practice and Racial Representations in a Changing South Africa. *South African Journal of Psychology*, *35*(3), 444–

459. https://doi.org/10.1177/008124630503500304

Dys-Steenbergen, O., Wright, S. C., & Aron, A. (2016). Self-expansion motivation improves cross-group interactions and enhances self-growth. *Group Processes & Intergroup Relations*, *19*(1), 60–71. https://doi.org/10.1177/1368430215583517

Eisenberg, N., Sallquist, J., French, D. C., Purwono, U., Suryanti, T. A., & Pidada, S. (2009). The relations of majority-minority group status and having an other-religion friend to Indonesian youths' socioemotional functioning. *Developmental Psychology*, *45*(1), 248–259. https://doi.org/10.1037/a0014028

Eller, A., & Abrams, D. (2004). Come together: Longitudinal comparisons of Pettigrew's reformulated intergroup contact model and the common ingroup identity model in Anglo-French and Mexican-American contexts. *European Journal of Social Psychology*, *34*(3), 229–256. https://doi.org/10.1002/ejsp.194

Ely, R. J., & Thomas, D. A. (2001). Cultural Diversity at Work: The Effects of Diversity Perspectives on Work Group Processes and Outcomes. *Administrative Science Quarterly*, *46*(2), 229–273. https://doi.org/10.2307/2667087

elysee.fr. (2020). *Protéger les libertés en luttant contre le séparatisme islamiste: Conférence de presse du Président Emmanuel Macron à Mulhouse.* elysee.fr. https://www.elysee.fr/emmanuel-macron/2020/02/18/proteger-les-libertes-en-luttant-contre-le-separatisme-islamiste-conference-de-presse-du-president-emmanuel-macron-a-mulhouse

Emerson, M. O., Kimbro, R. T., & Yancey, G. (2002). Contact Theory Extended: The Effects of Prior Racial Contact on Current Social Ties. *Social Science*

Quarterly, *83*(3), 745–761. https://doi.org/10.1111/1540-6237.00112

Enders, C. K., & Bandalos, D. L. (2001). The Relative Performance of Full

Information Maximum Likelihood Estimation for Missing Data in Structural

Equation Models. *Structural Equation Modeling: A Multidisciplinary*

Journal, *8*(3), 430–457. https://doi.org/10.1207/S15328007SEM0803_5

Ethnos. (2006). *The Decline of Britishness: A Research Study*. ETHNOS Research

and Consultancy London.

European Commission. (2007). *Press release: Three in four EU citizens welcome*

dialogue with other cultures [Text].

https://ec.europa.eu/commission/presscorner/detail/en/IP_07_1846

Federal Statistical Office. (2018). *Attitudes towards foreigners*.

https://www.bfs.admin.ch/bfs/en/home/statistiken/bevoelkerung/migration-

integration/zusammenleben-schweiz/einstellungen-auslaender.html

Federal Statistical Office. (2020, August 27). *Permanent and non-permanent*

resident population by canton, residence permit, sex, age class and

citizenship—2010-2019 | Table. Federal Statistical Office.

/content/bfs/en/home/statistiken/bevoelkerung/migration-

integration/auslaendische-bevoelkerung.assetdetail.14087560.html

Finchilescu, G. (2010). Intergroup Anxiety in Interracial Interaction: The Role of

Prejudice and Metastereotypes. *Journal of Social Issues*, *66*(2), 334–351.

https://doi.org/10.1111/j.1540-4560.2010.01648.x

Flores, A. R. (2015). Attitudes toward transgender rights: Perceived knowledge and

secondary interpersonal contact. *Politics, Groups, and Identities*, *3*(3), 398–

416. https://doi.org/10.1080/21565503.2015.1050414

Furrow, R. E. (2019). Regression to the Mean in Pre–Post Testing: Using

Simulations and Permutations to Develop Null Expectations. *CBE—Life*

Sciences Education, *18*(2), le2. https://doi.org/10.1187/cbe.19-02-0034

Gaertner, S. L., Dovidio, J. F., Anastasio, P. A., Bachman, B. A., & Rust, M. C.

(1993). The Common Ingroup Identity Model: Recategorization and the

Reduction of Intergroup Bias. *European Review of Social Psychology*, *4*(1),

1–26. https://doi.org/10.1080/14792779343000004

Gelman, A., & Loken, E. (2014). The statistical crisis in science: Data-dependent

analysis--a "garden of forking paths"--explains why many

statistically significant comparisons don't hold up. *American Scientist*,

102(6), 460–466.

GESIS, L.-I. F. zialwissenschaften. (2017). *Allgemeine Bevölkerungsumfrage der*

Sozialwissenschaften ALLBUS 2016 [Data set]. GESIS Data Archive.

https://doi.org/10.4232/1.12796

Gillborn, D., Demack, S., Rollock, N., & Warmington, P. (2017). Moving the

goalposts: Education policy and 25 years of the Black/White achievement

gap. *British Educational Research Journal*, *43*(5), 848–874.

https://doi.org/10.1002/berj.3297

Glazer, N. (1998). *We Are All Multiculturalists Now*. Harvard University Press.

Gonzalez-Mulé, E., & Aguinis, H. (2018). Advancing Theory by Assessing

Boundary Conditions With Metaregression: A Critical Review and Best-

Practice Recommendations. *Journal of Management*, *44*(6), 2246–2273.

https://doi.org/10.1177/0149206317710723

Goto, S. G., & Chan, D. K.-S. (2005). Becoming friends or remaining foes: An

empirical test of a causal model of intergroup contact across two cultures.

International Journal of Intercultural Relations, *29*(2), 197–216.

https://doi.org/10.1016/j.ijintrel.2005.05.003

GOV.UK. (2019). *Ethnicity facts and figures: Age groups*. https://www.ethnicity-

facts-figures.service.gov.uk/uk-population-by-ethnicity/demographics/age-groups/latest#data-sources

Graf, S., Paolini, S., & Rubin, M. (2014). Negative intergroup contact is more influential, but positive intergroup contact is more common: Assessing contact prominence and contact prevalence in five Central European countries. *European Journal of Social Psychology*, *44*(6), 536–547. https://doi.org/10.1002/ejsp.2052

Graham, S., Munniksma, A., & Juvonen, J. (2014). Psychosocial benefits of cross-ethnic friendships in urban middle schools. *Child Development*, *85*(2), 469–483. https://doi.org/10.1111/cdev.12159

Granovetter, M. S. (2005). The Impact of Social Structure on Economic Outcomes. *Journal of Economic Perspectives*, *19*(1), 33–50. https://doi.org/10.1257/0895330053147958

Greenberg, M. T., & Harris, A. R. (2012). Nurturing Mindfulness in Children and Youth: Current State of Research. *Child Development Perspectives*, *6*(2), 161–166. https://doi.org/10.1111/j.1750-8606.2011.00215.x

Greenwald, A. G., Banaji, M. R., & Nosek, B. A. (2015). Statistically small effects of the Implicit Association Test can have societally large effects. *Journal of Personality and Social Psychology*, *108*(4), 553–561. https://doi.org/10.1037/pspa0000016

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*(1), 17–41. https://doi.org/10.1037/a0015575

Gromer, D. (2020). *apa: Format Outputs of Statistical Tests According to APA Guidelines*. https://github.com/dgromer/apa

Guerra, R., Bierwiaczonek, K., Ferreira, M., Golec de Zavala, A., Abakoumkin, G., Wildschut, T., & Sedikides, C. (2020). An intergroup approach to collective narcissism: Intergroup threats and hostility in four European Union countries. *Group Processes & Intergroup Relations*, 1368430220972178. https://doi.org/10.1177/1368430220972178

Guimond, S., Sablonnière, R. de la, & Nugier, A. (2014). Living in a multicultural world: Intergroup ideologies and the societal context of intergroup relations. *European Review of Social Psychology*, *25*(1), 142–188. https://doi.org/10.1080/10463283.2014.957578

Hadler, M. (2012). The Influence of World Societal Forces on Social Tolerance. A Time Comparative Study of Prejudices in 32 Countries. *The Sociological Quarterly*, *53*(2), 211–237. https://doi.org/10.1111/j.1533-8525.2012.01232.x

Hall, S. (2001). *The multicultural question*. Pavis Centre for Social and Cultural Research, The Open University.

Hallinan, M. T., & Williams, R. A. (1990). Students' Characteristics and the Peer-Influence Process. *Sociology of Education*, *63*(2), 122–132. https://doi.org/10.2307/2112858

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889

Harper, C. E., & Yeung, F. (2013). Perceptions of Institutional Commitment to Diversity as a Predictor of College Students' Openness to Diverse Perspectives. *The Review of Higher Education*, *37*(1), 25–44. https://doi.org/10.1353/rhe.2013.0065

Harris, A. (2013). *Young People and Everyday Multiculturalism*. Routledge.

Harwood, J., Hewstone, M., Paolini, S., & Voci, A. (2005). Grandparent-Grandchild

Contact and Attitudes Toward Older Adults: Moderator and Mediator Effects. *Personality and Social Psychology Bulletin*, *31*(3), 393–406. https://doi.org/10.1177/0146167204271577

Hayward, L. E., Tropp, L. R., Hornsey, M. J., & Barlow, F. K. (2017). Toward a Comprehensive Understanding of Intergroup Contact: Descriptions and Mediators of Positive and Negative Contact Among Majority and Minority Groups. *Personality and Social Psychology Bulletin*, *43*(3), 347–364. https://doi.org/10.1177/0146167216685291

Hayward, L. E., Tropp, L. R., Hornsey, M. J., & Barlow, F. K. (2018). How negative contact and positive contact with Whites predict collective action among racial and ethnic minorities. *British Journal of Social Psychology*, *57*(1), 1–20. https://doi.org/10.1111/bjso.12220

Hedeker, D. (2015). Methods for Multilevel Ordinal Data in Prevention Research. *Prevention Science : The Official Journal of the Society for Prevention Research*, *16*(7), 997–1006. https://doi.org/10.1007/s11121-014-0495-x

Heitmeyer, W. (2004). *Deutsche Zustände. Folge 3* (3rd ed.). Suhrkamp Verlag.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X

Herek, G. M., & Capitanio, J. P. (1996). "Some of My Best Friends" Intergroup Contact, Concealable Stigma, and Heterosexuals' Attitudes Toward Gay Men and Lesbians. *Personality and Social Psychology Bulletin*, *22*(4), 412–424. https://doi.org/10.1177/0146167296224007

Herring, C. (2009). Does Diversity Pay?: Race, Gender, and the Business Case for Diversity. *American Sociological Review*, *74*(2), 208–224. https://doi.org/10.1177/000312240907400203

Hewstone, M., & Brown, R. (Eds.). (1986). *Contact and conflict in intergroup encounters* (Vol. xiii). Basil Blackwell.

Hodson, G. (2011). Do Ideologically Intolerant People Benefit From Intergroup Contact? *Current Directions in Psychological Science*, *20*(3), 154–159. https://doi.org/10.1177/0963721411409025

Holmes, E. P., Corrigan, P. W., Williams, P., Conor, J., & Kubiak, M. A. (1999). Changing Attitudes About Schizophrenia. *Schizophrenia Bulletin*, *25*(3), 10.

Holoien, D. S., & Shelton, J. N. (2012). You deplete me: The cognitive costs of colorblindness on ethnic minorities. *Journal of Experimental Social Psychology*, *48*(2), 562–565. https://doi.org/10.1016/j.jesp.2011.09.010

Hong, Y., Chao, M. M., & No, S. (2009). Dynamic Interracial/Intercultural Processes: The Role of Lay Theories of Race. *Journal of Personality*, *77*(5), 1283–1310. https://doi.org/10.1111/j.1467-6494.2009.00582.x

Howell, J., & Emerson, M. O. (2018). Preserving racial hierarchy amidst changing racial demographics: How neighbourhood racial preferences are changing while maintaining segregation. *Ethnic and Racial Studies*, *41*(15), 2770–2789. https://doi.org/10.1080/01419870.2017.1398833

Hutchison, P., & Rosenthal, H. E. S. (2011). Prejudice against Muslims: Anxiety as a mediator between intergroup contact and attitudes, perceived group variability and behavioural intentions. *Ethnic and Racial Studies*, *34*(1), 40–61. https://doi.org/10.1080/01419871003763312

Iannone, R. (2020). *DiagrammeR: Graph/Network Visualization*. https://CRAN.R-project.org/package=DiagrammeR

Iannone, R., Cheng, J., & Schloerke, B. (2020). *gt: Easily Create Presentation-Ready Display Tables*. https://CRAN.R-project.org/package=gt

Ihlanfeldt, K. R., & Scafidi, B. (2004). Whites' neighbourhood racial preferences and

neighbourhood racial composition in the United States: Evidence from the

multi-city study of urban inequality. *Housing Studies*, *19*(3), 325–359.

https://doi.org/10.1080/0267303042000204278

Islam, M. R., & Hewstone, M. (1993). Dimensions of Contact as Predictors of

Intergroup Anxiety, Perceived Out-Group Variability, and Out-Group

Attitude: An Integrative Model. *Personality and Social Psychology Bulletin*,

19(6), 700–710. https://doi.org/10.1177/0146167293196005

Islam, N. (2018, June 27). *An anti-Muslim narrative has shaped policy for decades.

The travel ban will make it worse.* Vox. https://www.vox.com/first-

person/2018/6/27/17510560/travel-ban-muslim-trump-islamophobia

Ito, T. A., & Urland, G. R. (2003). Race and gender on the brain: Electrocortical

measures of attention to the race and gender of multiply categorizable

individuals. *Journal of Personality and Social Psychology*, *85*(4), 616–626.

https://doi.org/10.1037/0022-3514.85.4.616

Jackman, S. (2020). *pscl: Classes and Methods for R Developed in the Political

Science Computational Laboratory*. United States Studies Centre, University

of Sydney. https://github.com/atahk/pscl/

Jackson, D. L. (2003). Revisiting Sample Size and Number of Parameter Estimates:

Some Support for the N:q Hypothesis. *Structural Equation Modeling: A

Multidisciplinary Journal*, *10*(1), 128–141.

https://doi.org/10.1207/S15328007SEM1001_6

Jamshidian, M., & Jalal, S. (2010). Tests of Homoscedasticity, Normality, and

Missing Completely at Random for Incomplete Multivariate Data.

Psychometrika, *75*(4), 649–674. https://doi.org/10.1007/s11336-010-9175-3

Jamshidian, M., Jalal, S., & Jansen, C. (2014). MissMech: An R Package for Testing

Homoscedasticity, Multivariate Normality, and Missing Completely at

Random (MCAR). *Journal of Statistical Software*, *56*(6), 1–31.

Jansen, W. S., Vos, M. W., Otten, S., Podsiadlowski, A., & Zee, K. I. van der. (2016). Colorblind or colorful? How diversity approaches affect cultural majority and minority employees. *Journal of Applied Social Psychology*, *46*(2), 81–93. https://doi.org/10.1111/jasp.12332

Johnston, B. M., & Glasford, D. E. (2018). Intergroup contact and helping: How quality contact and empathy shape outgroup helping. *Group Processes & Intergroup Relations*, *21*(8), 1185–1201.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2019). *semTools: Useful tools for structural equation modeling*. https://CRAN.R-project.org/package=semTools

Joseph, S., Weatherall, K., & Stringer, M. (1997). Attributions for unemployment in Northern Ireland: Does it make a difference what your name is? *The Irish Journal of Psychology*, *18*, 341–348. https://doi.org/10.1080/03033910.1997.10558151

Jugert, P., Noack, P., & Rutland, A. (2012). Children's cross-ethnic friendships: Why are they less stable than same-ethnic friendships? *European Journal of Developmental Psychology*, *10*(6), 649–662. https://doi.org/10.1080/17405629.2012.734136

Kanas, A., Scheepers, P., & Sterkens, C. (2017). Positive and negative contact and attitudes towards the religious out-group: Testing the contact hypothesis in conflict and non-conflict regions of Indonesia and the Philippines. *Social Science Research*, *63*, 95–110. https://doi.org/10.1016/j.ssresearch.2016.09.019

Karlsson, P., & Bergmark, A. (2015). Compared with what? An analysis of control-group types in Cochrane and Campbell reviews of psychosocial treatment

efficacy with substance use disorders. *Addiction*, *110*(3), 420–428.

https://doi.org/10.1111/add.12799

Karson, K. (2020, June 19). *Divisions emerge on renaming military bases, reparations amid unrest: POLL*. ABC News.

https://abcnews.go.com/Politics/divisions-emerge-renaming-military-bases-reparations-amid-unrest/story?id=71326039

Kauff, M., Asbrock, F., & Schmid, K. (2020). Pro-diversity beliefs and intergroup relations. *European Review of Social Psychology*, *0*(0), 1–36.

https://doi.org/10.1080/10463283.2020.1853377

Kauff, M., & Wagner, U. (2012). Valuable Therefore Not Threatening: The Influence of Diversity Beliefs on Discrimination Against Immigrants. *Social Psychological and Personality Science*, *3*(6), 714–721.

https://doi.org/10.1177/1948550611435942

Kaufmann, E., & Harris, G. (2015). "White Flight" or Positive Contact? Local Diversity and Attitudes to Immigration in Britain. *Comparative Political Studies*, *48*(12), 1563–1590. https://doi.org/10.1177/0010414015581684

Kawabata, Y., & Crick, N. R. (2011). The significance of cross-racial/ethnic friendships: Associations with peer victimization, peer support, sociometric status, and classroom diversity. *Developmental Psychology*, *47*(6), 1763–1775. https://doi.org/10.1037/a0025399

Khan, T. (2020). The Citizenship Amendment Act, 2019: A Religion Based Pathway to Indian Citizenship. *SSRN Electronic Journal*.

https://doi.org/10.2139/ssrn.3665743

Kline, R. B. (2015). *Principles and Practice of Structural Equation Modeling, Fourth Edition*. Guilford Publications.

Knowles, E. D., Lowery, B. S., Hogan, C. M., & Chow, R. M. (2009). On the

malleability of ideology: Motivated construals of color blindness. *Journal of Personality and Social Psychology*, *96*(4), 857–869. https://doi.org/10.1037/a0013595

Kraus, S. J. (1995). Attitudes and the Prediction of Behavior: A Meta-Analysis of the Empirical Literature. *Personality and Social Psychology Bulletin*, *21*(1), 58–75. https://doi.org/10.1177/0146167295211007

Krysan, M. (2002). Whites who say they'd flee: Who are they, and why would they leave? *Demography*, *39*(4), 675–696. https://doi.org/10.1353/dem.2002.0037

Laar, C. V., Levin, S., Sinclair, S., & Sidanius, J. (2005). The effect of university roommate contact on ethnic attitudes and behavior. *Journal of Experimental Social Psychology*, *41*(4), 329–345. https://doi.org/10.1016/j.jesp.2004.08.002

Lakens, D., Hilgard, J., & Staaks, J. (2016). On the reproducibility of meta-analyses: Six practical recommendations. *BMC Psychology*, *4*(1), 24. https://doi.org/10.1186/s40359-016-0126-3

Landmann, H., Gaschler, R., & Rohmann, A. (2019). What is threatening about refugees? Identifying different types of threat and their association with emotional responses and attitudes towards refugee migration. *European Journal of Social Psychology*, ejsp.2593. https://doi.org/10.1002/ejsp.2593

Langer, G. (2020, July 21). *63% support Black Lives Matter as recognition of discrimination jumps: POLL*. ABC News. https://abcnews.go.com/Politics/63-support-black-lives-matter-recognition-discrimination-jumps/story?id=71779435

LaPiere, R. T. (1934). Attitudes vs. Actions. *Social Forces*, *13*(2), 230–237.

Large, P., & Ghosh, K. (2006). Estimates of the population by ethnic group for areas within England. *Population Trends*, *124*, 8–17.

Lemmer, G., & Wagner, U. (2015). Can we really reduce ethnic prejudice outside the

lab? A meta-analysis of direct and indirect contact interventions. *European*

Journal of Social Psychology, *45*(2), 152–168.

https://doi.org/10.1002/ejsp.2079

Leslie, L. M., Bono, J. E., Kim, Y. (Sophia), & Beaver, G. R. (2020). On melting

pots and salad bowls: A meta-analysis of the effects of identity-blind and

identity-conscious diversity ideologies. *Journal of Applied Psychology*,

105(5), 453–471. https://doi.org/10.1037/apl0000446

Leslie, L. M., King, E. B., Bradley, J. C., & Hebl, M. R. (2008). Triangulation

Across Methodologies: All Signs Point to Persistent Stereotyping and

Discrimination in Organizations. *Industrial and Organizational Psychology*,

1(4), 399–404. https://doi.org/10.1111/j.1754-9434.2008.00073.x

Letki, N. (2008). Does Diversity Erode Social Cohesion? Social Capital and Race in

British Neighbourhoods. *Political Studies*, *56*(1), 99–126.

https://doi.org/10.1111/j.1467-9248.2007.00692.x

Levin, S., Matthews, M., Guimond, S., Sidanius, J., Pratto, F., Kteily, N., Pitpitan, E.

V., & Dover, T. (2012). Assimilation, multiculturalism, and colorblindness:

Mediated and moderated relationships between social dominance orientation

and prejudice. *Journal of Experimental Social Psychology*, *48*(1), 207–212.

https://doi.org/10.1016/j.jesp.2011.06.019

Levin, S., van Laar, C., & Sidanius, J. (2003). The Effects of Ingroup and Outgroup

Friendships on Ethnic Attitudes in College: A Longitudinal Study. *Group*

Processes & Intergroup Relations, *6*(1), 76–92.

https://doi.org/10.1177/1368430203006001013

Li, Y., Pickles, A., & Savage, M. (2005). Social Capital and Social Trust in Britain.

European Sociological Review, *21*(2), 109–123. JSTOR.

Lin, L.-C., Huang, P.-H., & Weng, L.-J. (2017). Selecting Path Models in SEM: A

    Comparison of Model Selection Criteria. *Structural Equation Modeling: A*

    Multidisciplinary Journal, *24*(6), 855–869.

    https://doi.org/10.1080/10705511.2017.1363652

Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in

    latent variable panel analyses of longitudinal data. *International Journal of*

    Behavioral Development, *31*(4), 357–365.

    https://doi.org/10.1177/0165025407077757

Loewenthal, K. M. (2001). *An introduction to psychological tests and scales*.

    Psychology Press.

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R.

    Behavior Research Methods, *49*(4), 1494–1502.

    https://doi.org/10.3758/s13428-016-0809-y

MacInnis, C. C., & Page-Gould, E. (2015). How Can Intergroup Interaction Be Bad

    If Intergroup Contact Is Good? Exploring and Reconciling an Apparent

    Paradox in the Science of Intergroup Relations. *Perspectives on*

    Psychological Science, *10*(3), 307–327.

    https://doi.org/10.1177/1745691614568482

MacInnis, C., & Hodson, G. (2012). `Where the rubber hits the road' en route to

    inter-group harmony: Examining contact intentions and contact behaviour

    under meta-stereotype threat. *British Journal of Social Psychology*, *51*(2),

    363–373. https://doi.org/10.1111/j.2044-8309.2010.02014.x

MacKinnon, D. (n.d.). *An Answer to Julie's Question*. Retrieved December 6, 2020,

    from http://www.public.asu.edu/~davidpm/ripl/maloy.htm

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence Limits for

    the Indirect Effect: Distribution of the Product and Resampling Methods.

Multivariate Behavioral Research, *39*(1), 99–128.

https://doi.org/10.1207/s15327906mbr3901_4

Mackinnon, S., Curtis, R., & O'Connor, R. (2020). *A Tutorial in Longitudinal*

Measurement Invariance and Cross-lagged Panel Models Using Lavaan.

PsyArXiv. https://doi.org/10.31234/osf.io/tkzrb

Makowski, Dominique, Lüdecke, Daniel, & Ben-Shachar, Mattan S. (2020). The

report package for R: Ensuring the use of best practices for results reporting.

CRAN. https://github.com/easystats/report

Mallett, R. K., Huntsinger, J. R., Sinclair, S., & Swim, J. K. (2008). Seeing Through

Their Eyes: When Majority Group Members Take Collective Action on

Behalf of an Outgroup. *Group Processes & Intergroup Relations*, *11*(4), 451–

470. https://doi.org/10.1177/1368430208095400

Mangiafico, S. S. (2016). *Summary and analysis of extension program evaluation in*

R, version 1.18. 0. Rutgers Cooperative Extension (https://rcompanion.

org/handbook/).

Maras, P., & Brown, R. (2010). Effects of different forms of school contact on

children's attitudes toward disabled and non-disabled peers. *British Journal*

of Educational Psychology, *70*(3), 337–351.

https://doi.org/10.1348/000709900158164

Massey, D. S., & Denton, N. A. (1993). *American apartheid: Segregation and the*

making of the underclass. Harvard University Press.

McConahay, J. B. (1978). The Effects of School Desegregation upon Students'

Racial Attitudes and Behavior: A Critical Review of the Literature and a

Prolegomenon to Future Research. *Law and Contemporary Problems*, *42*(3),

77. https://doi.org/10.2307/1191344

McKeown, S., & Taylor, L. K. (2017). Intergroup Contact and Peacebuilding:

Promoting Youth Civic Engagement in Northern Ireland. *Journal of Social and Political Psychology*, *5*(2), 415–434.

McKeown, S., Williams, A., & Pauker, K. (2017). Stories that move them: Changing children's behaviour toward diverse peers. *Journal of Community & Applied Social Psychology*, *27*(5), 381–387. https://doi.org/10.1002/casp.2316

McLeod, P. L., Lobel, S. A., & Cox, T. H. (1996). Ethnic Diversity and Creativity in Small Groups. *Small Group Research*, *27*(2), 248–264. https://doi.org/10.1177/1046496496272003

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, *27*(1), 415–444. https://doi.org/10.1146/annurev.soc.27.1.415

McShane, B. B., & Böckenholt, U. (2017). Single-Paper Meta-Analysis: Benefits for Study Summary, Theory Testing, and Replicability. *Journal of Consumer Research*, *43*(6), 1048–1063. https://doi.org/10.1093/jcr/ucw085

Meleady, R., & Forder, L. (2019). When contact goes wrong: Negative intergroup contact promotes generalized outgroup avoidance. *Group Processes & Intergroup Relations*, *22*(5), 688–707. https://doi.org/10.1177/1368430218761568

Meleady, R., Seger, C., & Vermue, M. (2020). Evidence of a dynamic association between intergroup contact and intercultural competence. *Group Processes & Intergroup Relations*. https://doi.org/10.1177/1368430220940400

Mereish, E., & Poteat, V. P. (2015). Effects of heterosexuals' direct and extended friendships with sexual minorities on their attitudes and behaviors: Intergroup anxiety and attitude strength as mediators and moderators. *Journal of Applied Social Psychology*, *45*(3), 147–157. https://doi.org/10.1111/jasp.12284

Meshel, D. S. (1997). The contact hypothesis and the effects of intergenerational

contact on adolescents' attitudes and stereotypes toward older people. *Texas Tech University*, 184.

Miville, M. L., Gelso, C. J., Pannu, R., Liu, W., Touradji, P., Holloway, P., & Fuertes, J. (1999). Appreciating similarities and valuing differences: The Miville-Guzman Universality-Diversity Scale. *Journal of Counseling Psychology*, *46*(3), 291–307. https://doi.org/10.1037/0022-0167.46.3.291

Morrison, K. R., Plaut, V. C., & Ybarra, O. (2010). Predicting Whether Multiculturalism Positively or Negatively Influences White Americans' Intergroup Attitudes: The Role of Ethnic Identification. *Personality and Social Psychology Bulletin*, *36*(12), 1648–1661. https://doi.org/10.1177/0146167210386118

Mousa, S. (2018). Overcoming the Trust Deficit: Intergroup Contact and Associational Life in post-ISIS Iraq. *Working Paper, Accessed from Http://Aalims.Org/Uploads/Mousa_Overcoming-the-Trust-Deficit_Aalims2018.Pdf*, 48.

Mulvey, K. L., Palmer, S. B., & Abrams, D. (2016). Race-Based Humor and Peer Group Dynamics in Adolescence: Bystander Intervention and Social Exclusion. *Child Development*, *87*(5), 1379–1391. https://doi.org/10.1111/cdev.12600

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 1–9. https://doi.org/10.1038/s41562-016-0021

Murrar, S., Campbell, M. R., & Brauer, M. (2020). Exposure to peers' pro-diversity attitudes increases inclusion and reduces the achievement gap. *Nature Human Behaviour*, *4*(9), 889–897. https://doi.org/10.1038/s41562-020-0899-5

Nagda, B. (Ratnesh) A. (2006). Breaking Barriers, Crossing Borders, Building Bridges: Communication Processes in Intergroup Dialogues. *Journal of Social Issues*, *62*(3), 553–576. https://doi.org/10.1111/j.1540-4560.2006.00473.x

NCS Trust. (2020). *National Citizen Service Annual Report and Accounts 2018/2019*. GOV.UK. https://www.gov.uk/government/publications/national-citizen-service-annual-report-and-accounts-20182019

Neville, H. A., Poteat, V. P., Lewis, J. A., & Spanierman, L. B. (2014). Changes in White college students' color-blind racial ideology over 4 years: Do diversity experiences make a difference? *Journal of Counseling Psychology*, *61*(2), 179–190. https://doi.org/10.1037/a0035168

Norton, M. I., & Sommers, S. R. (2011). Whites See Racism as a Zero-Sum Game That They Are Now Losing. *Perspectives on Psychological Science*, *6*(3), 215–218. https://doi.org/10.1177/1745691611406922

Norton, M. I., Sommers, S. R., Apfelbaum, E. P., Pura, N., & Ariely, D. (2006). Color Blindness and Interracial Interaction: Playing the Political Correctness Game. *Psychological Science*, *17*(11), 949–953. https://doi.org/10.1111/j.1467-9280.2006.01810.x

Nuijten, M. B., Borghuis, J., Veldkamp, C. L. S., Dominguez-Alvarez, L., van Assen, M. A. L. M., & Wicherts, J. M. (2017). Journal Data Sharing Policies and Statistical Reporting Inconsistencies in Psychology. *Collabra: Psychology*, *3*(31). https://doi.org/10.1525/collabra.102

Nuijten, M. B., Hartgerink, C. H. J., van Assen, M. A. L. M., Epskamp, S., & Wicherts, J. M. (2016). The prevalence of statistical reporting errors in psychology (1985–2013). *Behavior Research Methods*, *48*(4), 1205–1226. https://doi.org/10.3758/s13428-015-0664-2

Oberski, D. (2014). lavaan.survey: An R Package for Complex Survey Analysis of Structural Equation Models. *Journal of Statistical Software*, *57*(1), 1–27.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). https://doi.org/10.1126/science.aac4716

Osborn, H. J., Sosa, N., & Rios, K. (2020). Perceiving demographic diversity as a threat: Divergent effects of multiculturalism and polyculturalism. *Group Processes & Intergroup Relations*, *23*(7), 1014–1031. https://doi.org/10.1177/1368430219880606

Oudenhoven, J. P. V., Groenewoud, J. T., & Hewstone, M. (1996). Cooperation, ethnic salience and generalization of interethnic attitudes. *European Journal of Social Psychology*, *26*(4), 649–661. https://doi.org/10.1002/(SICI)1099-0992(199607)26:4<649::AID-EJSP780>3.0.CO;2-T

Page-Gould, E., Mendoza-Denton, R., & Tropp, L. R. (2008). With a little help from my cross-group friend: Reducing anxiety in intergroup contexts through cross-group friendship. *Journal of Personality and Social Psychology*, *95*(5), 1080–1094. https://doi.org/10.1037/0022-3514.95.5.1080

Pais, J. F., South, S. J., & Crowder, K. (2009). White Flight Revisited: A Multiethnic Perspective on Neighborhood Out-Migration. *Population Research and Policy Review*, *28*(3), 321–346. https://doi.org/10.1007/s11113-008-9101-x

Palmer, S. B., & Cameron, L. (2010). Bystander intervention in subtle and explicit racist incidents. *Paper Presented at EASP Small Group Meeting on 'Developmental Perspectives on Intergroup Prejudice: Advances in Theory, Measurement, and Intervention', Lisbon, Portugal.*

Paluck, E. L., Green, S. A., & Green, D. P. (2019). The contact hypothesis re-evaluated. *Behavioural Public Policy*, *3*(02), 129–158. https://doi.org/10.1017/bpp.2018.25

Paolini, S., Harwood, J., Hewstone, M., & Neumann, D. L. (2018). Seeking and
avoiding intergroup contact: Future frontiers of research on building social
integration. *Social and Personality Psychology Compass*, *12*(12), e12422.
https://doi.org/10.1111/spc3.12422

Paolini, S., Harwood, J., & Rubin, M. (2010). Negative Intergroup Contact Makes
Group Memberships Salient: Explaining Why Intergroup Conflict Endures.
Personality and Social Psychology Bulletin, *36*(12), 1723–1738.
https://doi.org/10.1177/0146167210388667

Paolini, S., Harwood, J., Rubin, M., Husnu, S., Joyce, N., & Hewstone, M. (2014).
Positive and extensive intergroup contact in the past buffers against the
disproportionate impact of negative contact in the present. *European Journal
of Social Psychology*, *44*(6), 548–562. https://doi.org/10.1002/ejsp.2029

Paolini, S., & McIntyre, K. (2019). Bad Is Stronger Than Good for Stigmatized, but
Not Admired Outgroups: Meta-Analytical Tests of Intergroup Valence
Asymmetry in Individual-to-Group Generalization Experiments. *Personality
and Social Psychology Review*, *23*(1), 3–47.
https://doi.org/10.1177/1088868317753504

Parents Involved in Community Schools v. Seattle School Dist. No. 1, 551 US ___
(US Supreme Court 2007).

Park, B., & Judd, C. M. (2005). Rethinking the Link Between Categorization and
Prejudice Within the Social Cognition Perspective. *Personality and Social
Psychology Review*, *9*(2), 108–130.
https://doi.org/10.1207/s15327957pspr0902_2

Pettigrew, T. F. (1979). The Ultimate Attribution Error: Extending Allport's
Cognitive Analysis of Prejudice. *Personality and Social Psychology Bulletin*,
5(4), 461–476. https://doi.org/10.1177/014616727900500407

Pettigrew, T. F. (1997). Generalized Intergroup Contact Effects on Prejudice. *Personality and Social Psychology Bulletin*, *23*, 173–185. https://doi.org/10.1177/0146167297232006

Pettigrew, T. F. (2009). Secondary transfer effect of contact: Do intergroup contact effects spread to noncontacted outgroups? *Social Psychology*, *40*(2), 55–65. https://doi.org/10.1027/1864-9335.40.2.55

Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, *90*(5), 751–783. https://doi.org/10.1037/0022-3514.90.5.751

Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, *38*(6), 922–934. https://doi.org/10.1002/ejsp.504

Pettigrew, T. F., & Tropp, L. R. (2011). *When Groups Meet: The Dynamics of Intergroup Contact*. Psychology Press.

Pettigrew, T. F., Wagner, U., & Christ, O. (2010). Population Ratios and Prejudice: Modelling Both Contact and Threat Effects. *Journal of Ethnic and Migration Studies*, *36*(4), 635–650. https://doi.org/10.1080/13691830903516034

Pica-Smith, C. (2009). Children Speak about Interethnic and Interracial Friendships in the Classroom: Lessons for Teachers. *Multicultural Education*, *17*(1), 38–47.

Pittinsky, T. L., Rosenthal, S. A., & Montoya, R. M. (2011). Measuring positive attitudes toward outgroups: Development and validation of the Allophilia Scale. In *Moving beyond prejudice reduction: Pathways to positive intergroup relations* (pp. 41–60). American Psychological Association. https://doi.org/10.1037/12319-002

Plaut, V. C., Garnett, F. G., Buffardi, L. E., & Sanchez-Burks, J. (2011). "What

about me?" Perceptions of exclusion and Whites' reactions to multiculturalism. *Journal of Personality and Social Psychology*, *101*(2), 337–353. https://doi.org/10.1037/a0022832

Plaut, V. C., Thomas, K. M., & Goren, M. J. (2009). Is Multiculturalism or Color Blindness Better for Minorities? *Psychological Science*, *20*(4), 444–446. https://doi.org/10.1111/j.1467-9280.2009.02318.x

Preacher, K. J., & Selig, J. P. (2012). Advantages of Monte Carlo Confidence Intervals for Indirect Effects. *Communication Methods and Measures*, *6*(2), 77–98. https://doi.org/10.1080/19312458.2012.679848

Purdie-Vaughns, V., Steele, C. M., Davies, P. G., Ditlmann, R., & Crosby, J. R. (2008). Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology*, *94*(4), 615–630. https://doi.org/10.1037/0022-3514.94.4.615

Putnam, R. D. (1995). Bowling alone: America's declining social capital. *Journal of Democracy*, *6*(1), 65–78.

Putnam, R. D. (2007). E Pluribus Unum: Diversity and Community in the Twenty-first Century The 2006 Johan Skytte Prize Lecture. *Scandinavian Political Studies*, *30*(2), 137–174. https://doi.org/10.1111/j.1467-9477.2007.00176.x

Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences*, *114*(41), 10870–10875. https://doi.org/10.1073/pnas.1706255114

Quito, A. (2016). *Switzerland's largest political party insists on depicting foreigners as black sheep*. Quartz. https://qz.com/617050/switzerlands-largest-political-party-insists-on-depicting-foreigners-as-black-sheep/

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Raftery, A. E. (1995). Bayesian Model Selection in Social Research. *Sociological Methodology*, *25*, 111–163. https://doi.org/10.2307/271063

Ramiah, A. A., Hewstone, M., Voci, A., Cairns, E., & Hughes, J. (2013). It's never too late for 'us' to meet 'them': Prior intergroup friendships moderate the impact of later intergroup friendships in educational settings. *British Journal of Educational Psychology*, *83*(1), 57–75. https://doi.org/10.1111/j.2044-8279.2011.02054.x

Rattan, A., & Ambady, N. (2013). Diversity ideologies and intergroup relations: An examination of colorblindness and multiculturalism. *European Journal of Social Psychology*, *43*(1), 12–21. https://doi.org/10.1002/ejsp.1892

Redick, T. S., Shipstead, Z., Wiemers, E. A., Melby-Lervåg, M., & Hulme, C. (2015). What's Working in Working Memory Training? An Educational Perspective. *Educational Psychology Review*, *27*(4), 617–633. https://doi.org/10.1007/s10648-015-9314-6

Reibel, M., & Regelson, M. (2011). Neighborhood Racial and Ethnic Change: The Time Dimension in Segregation. *Urban Geography*, *32*(3), 360–382. https://doi.org/10.2747/0272-3638.32.3.360

Reimer, N. K., Becker, J. C., Benz, A., Christ, O., Dhont, K., Klocke, U., Neji, S., Rychlowska, M., Schmid, K., & Hewstone, M. (2017). Intergroup contact and social change: Implications of negative and positive contact for collective action in advantaged and disadvantaged groups. *Personality and Social Psychology Bulletin*, *43*(1), 121–136. https://doi.org/10.1177/0146167216676478

Ridout, M. S. (1991). Testing for random dropouts in repeated measurement data.

Biometrics, *47*(4), 1617–1619. https://doi.org/10.2307/2532413

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons.

Rubio-Aparicio, M., López-López, J. A., Viechtbauer, W., Marín-Martínez, F., Botella, J., & Sánchez-Meca, J. (2020). Testing Categorical Moderators in Mixed-Effects Meta-analysis in the Presence of Heteroscedasticity. *The Journal of Experimental Education*, *88*(2), 288–310. https://doi.org/10.1080/00220973.2018.1561404

Ryan, C. S., Hunt, J. S., Weible, J. A., Peterson, C. R., & Casas, J. F. (2007). Multicultural and Colorblind Ideology, Stereotypes, and Ethnocentrism among Black and White Americans. *Group Processes & Intergroup Relations*, *10*(4), 617–637. https://doi.org/10.1177/1368430207084105

Sager, L. (2012). Residential Segregation and Socioeconomic Neighbourhood Sorting: Evidence at the Micro-neighbourhood Level for Migrant Groups in Germany. *Urban Studies*, *49*(12), 2617–2632. https://doi.org/10.1177/0042098011429487

Sanders, M., Chonaire, A. N., Carr, D., Heal, J., & Anik, L. (2017). Increasing social trust with an ice-breaking exercise–an RCT carried out with NCS participants. *Behaviour Insights Team*. http://www.ncsyes.co.uk/sites/default/files/Increasing%20social%20trust%20 with%20an%20ice-breaking%20exercise%20- %20BIT%20March%202017.pdf

Saucier, D. A., & Miller, C. T. (2003). The Persuasiveness of Racial Arguments as a

Subtle Measure of Racism. *Personality and Social Psychology Bulletin*, *29*(10), 1303–1315. https://doi.org/10.1177/0146167203254612

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, *1*(2), 143–186. https://doi.org/10.1080/0022250X.1971.9989794

Schmidt, P., & Weick, S. (2017). Kontakte und die Wahrnehmung von Bedrohungen besonders wichtig für die Einschätzung von Migranten: Einstellungen der deutschen Bevölkerung zu Zuwanderern von 1980 bis 2016. *Informationsdienst Soziale Indikatoren*, *57*, 1–7.

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining Power and Sample Size for Simple and Complex Mediation Models. *Social Psychological and Personality Science*, *8*(4), 379–386. https://doi.org/10.1177/1948550617715068

Schönwälder, K., & Söhn, J. (2009). Immigrant Settlement Structures in Germany: General Patterns and Urban Levels of Concentration of Major Groups. *Urban Studies*, *46*(7), 1439–1460. https://doi.org/10.1177/0042098009104575

Schütz, H., & Six, B. (1996). How strong is the relationship between prejudice and discrimination? A meta-analytic answer. *International Journal of Intercultural Relations*, *20*(3), 441–462. https://doi.org/10.1016/0147-1767(96)00028-4

Seeman, T. E. (1996). Social ties and health: The benefits of social integration. *Annals of Epidemiology*, *6*(5), 442–451.

Seider, S. (2009a). Overwhelmed and immobilized: Raising the consciousness of privileged young adults about world hunger and poverty. *International Studies Perspectives*, *10*(1), 60–76. https://doi.org/10.1111/j.1528-3585.2008.00358.x

Seider, S. (2009b). The trouble with teaching ethics on trolley cars and train tracks.

Journal of Moral Education, *38*(October 2013), 219–236.

https://doi.org/10.1080/03057240902792744

Selvanathan, H. P., Techakesari, P., Tropp, L. R., & Barlow, F. K. (2017). Whites for

racial justice: How contact with Black Americans predicts support for

collective action among White Americans. *Group Processes & Intergroup

Relations*, 1368430217690908. https://doi.org/10.1177/1368430217690908

Sengupta, N. K., & Sibley, C. G. (2013). Perpetuating One's Own Disadvantage:

Intergroup Contact Enables the Ideological Legitimation of Inequality.

Personality and Social Psychology Bulletin, *39*(11), 1391–1403.

https://doi.org/10.1177/0146167213497593

Sherif, M., & Sherif, C. W. (1953). *Groups in harmony and tension; an integration

of studies on intergroup relations*. Harper.

Shin, T., Davison, M. L., & Long, J. D. (2017). Maximum likelihood versus multiple

imputation for missing data in small longitudinal samples with nonnormality.

Psychological Methods, *22*(3), 426–449. https://doi.org/10.1037/met0000094

Social Integration Commission. (2014a). *Report 1: How integrated is modern

Britain?* http://socialintegrationcommission.org.uk/SIC_Report_WEB.pdf

Social Integration Commission. (2014b). *Report 2: Social integration: A wake-up

call*. http://socialintegrationcommission.org.uk/a-wake-up-call-social-

integration-commission.pdf

Stahl, G. K., Maznevski, M. L., Voigt, A., & Jonsen, K. (2010). Unraveling the

effects of cultural diversity in teams: A meta-analysis of research on

multicultural work groups. *Journal of International Business Studies*, *41*(4),

690–709. https://doi.org/10.1057/jibs.2009.85

Stapleton, L. M. (2008). Variance Estimation Using Replication Methods in

Structural Equation Modeling With Complex Sample Data. *Structural

Equation Modeling: A Multidisciplinary Journal, *15*(2), 183–210. https://doi.org/10.1080/10705510801922316

Stathi, S., Crisp, R. J., & Hogg, M. A. (2011). Imagining intergroup contact enables member-to-group generalization. *Group Dynamics: Theory, Research, and Practice*, *15*(3), 275–284. https://doi.org/10.1037/a0023752

Stearns, E., Buchmann, C., & Bonneau, K. (2009). Interracial friendships in the transition to college: Do birds of a feather flock together once they leave the nest. *Sociology of Education*, 173–195.

Stephan, W. G., Ybarra, O., Martnez, C. M., Schwarzwald, J., & Tur-Kaspa, M. (1998). Prejudice toward Immigrants to Spain and Israel: An Integrated Threat Theory Analysis. *Journal of Cross-Cultural Psychology*, *29*(4), 559–576. https://doi.org/10.1177/0022022198294004

Stevens, F. G., Plaut, V. C., & Sanchez-Burks, J. (2008). Unlocking the Benefits of Diversity: All-Inclusive Multiculturalism and Positive Organizational Change. *The Journal of Applied Behavioral Science*, *44*(1), 116–133. https://doi.org/10.1177/0021886308314460

Stolle, D., Petermann, S., Schmid, K., Schönwälder, K., Hewstone, M., Vertovec, S., Schmitt, T., & Heywood, J. (2013). Immigration-Related Diversity and Trust in German Cities: The Role of Intergroup Contact. *Journal of Elections, Public Opinion & Parties*, *23*(3), 279–298. https://doi.org/10.1080/17457289.2013.809350

Stürmer, S., Benbow, A. E. F., Siem, B., Barth, M., Bodansky, A. N., & Lotz-Schmitt, K. (2013). Psychological foundations of xenophilia: The role of major personality traits in predicting favorable attitudes toward cross-cultural contact and exploration. *Journal of Personality and Social Psychology*, *105*(5), 832–851. https://doi.org/10.1037/a0033488

Swart, H., Hewstone, M., Christ, O., & Voci, A. (2011). Affective mediators of

    intergroup contact: A three-wave longitudinal study in South Africa. *Journal*

    of Personality and Social Psychology, *101*(6), 1221–1238.

    https://doi.org/10.1037/a0024450

Swim, J. K., & Miller, D. L. (1999). White Guilt: Its Antecedents and Consequences

    for Attitudes Toward Affirmative Action. *Personality and Social Psychology*

    Bulletin, *25*(4), 500–514. https://doi.org/10.1177/0146167299025004008

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Biosocial Science*,

    1(S1), 173–191.

Talaska, C. A., Fiske, S. T., & Chaiken, S. (2008). Legitimating Racial

    Discrimination: Emotions, Not Beliefs, Best Predict Discrimination in a

    Meta-Analysis. *Social Justice Research*, *21*(3), 263–396.

    https://doi.org/10.1007/s11211-008-0071-2

Tam, T., Hewstone, M., Harwood, J., Voci, A., & Kenworthy, J. (2006). Intergroup

    Contact and Grandparent–Grandchild Communication: The   Effects

    of Self-Disclosure on Implicit and Explicit Biases Against Older People.

    Group Processes & Intergroup Relations, *9*(3), 413–429.

    https://doi.org/10.1177/1368430206064642

Tausch, N., Hewstone, M., Kenworthy, J. B., Psaltis, C., Schmid, K., Popan, J. R.,

    Cairns, E., & Hughes, J. (2010). Secondary transfer effects of intergroup

    contact: Alternative accounts and underlying processes. *Journal of*

    Personality and Social Psychology, *99*(2), 282–302.

    https://doi.org/10.1037/a0018553

Taylor, D. M. (2014). The social psychology of racial and cultural diversity: Issues

    of assimilation and multiculturalism. In *Bilingualism, multiculturalism, and*

    second language learning (pp. 21–40). Psychology Press.

Teney, C. (2011). Endorsement of Assimilationism among Ethnic Minority and

    Majority Youth in a Multination–Multiethnic Context: The Case of Brussels.

    European Sociological Review, *27*(2), 212–229.

    https://doi.org/10.1093/esr/jcq003

Tharoor, I. (2020). Analysis | Macron's liberal double standards. *Washington Post*.

    https://www.washingtonpost.com/world/2020/12/02/macron-france-race-

    press-freedom/

The Challenge. (2019). *The Challenge Impact Report 2017-18* (murr; p. 28).

    Archived at

    https://www.fortismere.haringey.sch.uk/_site/data/files/sixth%20form/careers

    %20info/F685642451F419766DE7EEBB0C0D9BCD.pdf

Thomsen, J. P. F., & Rafiqi, A. (2017). The Contact-Prejudice Relationship Among

    Ethnic Minorities: Examining the Facilitative Influence of Religiosity. *Social*

    Science Quarterly, *98*(5), 1571–1586. https://doi.org/10.1111/ssqu.12368

Todd, A. R., & Galinsky, A. D. (2012). The reciprocal link between multiculturalism

    and perspective-taking: How ideological and self-regulatory approaches to

    managing diversity reinforce each other. *Journal of Experimental Social*

    Psychology, *48*(6), 1394–1398. https://doi.org/10.1016/j.jesp.2012.07.007

Trawalter, S., Richeson, J. A., & Shelton, J. N. (2009). Predicting Behavior During

    Interracial Interactions: A Stress and Coping Approach: *Personality and*

    Social Psychology Review. https://doi.org/10.1177/1088868309345850

Tropp, L. R., & Bianchi, R. A. (2006). Valuing diversity and interest in intergroup

    contact. *Journal of Social Issues*, *62*(3), 533–551.

    https://doi.org/10.1111/j.1540-4560.2006.00472.x

Tropp, L. R., Hawi, D. R., Laar, C. V., & Levin, S. (2012). Cross-ethnic friendships,

    perceived discrimination, and their effects on ethnic activism over time: A

longitudinal investigation of three ethnic minority groups. *British Journal of Social Psychology*, *51*(2), 257–272. https://doi.org/10.1111/j.2044-8309.2011.02050.x

Tropp, L. R., O'Brien, T. C., & Migacheva, K. (2014). How Peer Norms of Inclusion and Exclusion Predict Children's Interest in Cross-Ethnic Friendships: Peer Norms and Cross-Ethnic Friendships. *Journal of Social Issues*, *70*(1), 151–166. https://doi.org/10.1111/josi.12052

Tropp, L. R., & Pettigrew, T. F. (2005a). Differential Relationships Between Intergroup Contact and Affective and Cognitive Dimensions of Prejudice. *Personality and Social Psychology Bulletin*, *31*(8), 1145–1158. https://doi.org/10.1177/0146167205274854

Tropp, L. R., & Pettigrew, T. F. (2005b). Relationships Between Intergroup Contact and Prejudice Among Minority and Majority Status Groups. *Psychological Science*, *16*(12), 951–957. https://doi.org/10.1111/j.1467-9280.2005.01643.x

Tunger-Zanetti, A. (2021). *Verhüllung: Die Burka-Debatte in der Schweiz* (1. Edition). Hier und Jetzt.

Turner, R. N., & Cameron, L. (2016). Confidence in Contact: A New Perspective on Promoting Cross-Group Friendship Among Children and Adolescents. *Social Issues and Policy Review*, *10*(1), 212–246. https://doi.org/10.1111/sipr.12023

Turner, R. N., Hewstone, M., & Voci, A. (2007). Reducing explicit and implicit outgroup prejudice via direct and extended contact: The mediating role of self-disclosure and intergroup anxiety. *Journal of Personality and Social Psychology*, *93*(3), 369–388. https://doi.org/10.1037/0022-3514.93.3.369

Turner, R. N., West, K., & Christie, Z. (2013). Out-group trust, intergroup anxiety, and out-group attitude as mediators of the effect of imagined intergroup contact on intergroup behavioral tendencies. *Journal of Applied Social*

Psychology, *43*(S2), E196–E205. https://doi.org/10.1111/jasp.12019

Uslaner, E. M. (2010). Segregation, mistrust and minorities. *Ethnicities*, *10*(4), 415–434. https://doi.org/10.1177/1468796810378325

van Dyk, A. C. (1990). Voorspellers van etniese houdings in 'n noue kontaksituasie [Determinants of ethnic attitudes in a close contact situation]. *South African Journal of Psychology*, *20*(3), 206–214. https://doi.org/10.1177/008124639002000311

Varshney, A. (2001). Ethnic Conflict and Civil Society: India and Beyond. *World Politics*, *53*(03), 362–398. https://doi.org/10.1353/wp.2001.0012

Vasta, E. (2007). From ethnic minorities to ethnic majority policy: Multiculturalism and the shift to assimilationism in the Netherlands. *Ethnic and Racial Studies*, *30*(5), 713–740. https://doi.org/10.1080/01419870701491770

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S* (Fourth). Springer. http://www.stats.ox.ac.uk/pub/MASS4

Verkuyten, M. (2005). Ethnic Group Identification and Group Evaluation Among Minority and Majority Groups: Testing the Multiculturalism Hypothesis. *Journal of Personality and Social Psychology*, *88*(1), 121–138. https://doi.org/10.1037/0022-3514.88.1.121

Verkuyten, M., & Martinovic, B. (2006). Understanding multicultural attitudes: The role of group status, identification, friendships, and justifying ideologies. *International Journal of Intercultural Relations*, *30*(1), 1–18. https://doi.org/10.1016/j.ijintrel.2005.05.015

Vertovec, S. (2007). Super-diversity and its implications. *Ethnic and Racial Studies*, *30*(6), 1024–1054. https://doi.org/10.1080/01419870701599465

Vertovec, S. (2010). Towards post-multiculturalism? Changing communities, conditions and contexts of diversity. *International Social Science Journal*,

61(199), 83–95. https://doi.org/10.1111/j.1468-2451.2010.01749.x

Vertovec, S., & Wessendorf, S. (2010). *The multiculturalism backlash: European discourses, policies, and practices*. Routledge.

Vezzali, L., Andrighetto, L., Capozza, D., Di Bernardo, G. A., & Saguy, T. (2017). Discussing differences between groups: The content of intergroup encounters and motivation for social change among members of advantaged groups. *Journal of Theoretical Social Psychology*, *1*(2), 52–59. https://doi.org/10.1002/jts5.12

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Visintin, E. P., Voci, A., Pagotto, L., & Hewstone, M. (2017). Direct, extended, and mass-mediated contact with immigrants in Italy: Their associations with emotions, prejudice, and humanity perceptions. *Journal of Applied Social Psychology*, *47*(4), 175–194. https://doi.org/10.1111/jasp.12423

Voci, A., & Hewstone, M. (2003). Intergroup contact and prejudice toward immigrants in Italy: The mediational role of anxiety and the moderational role of group salience. *Group Processes & Intergroup Relations*, *6*(1), 37–54. https://doi.org/10.1177/1368430203006001011

Vollhardt, J. R. (2010). Enhanced external and culturally sensitive attributions after extended intercultural contact. *British Journal of Social Psychology*, *49*(2), 363–383. https://doi.org/10.1348/014466609X459899

Vorauer, J. (2008). Unprejudiced and self-focused: When intergroup contact is experienced as being about the ingroup rather than the outgroup. *Journal of Experimental Social Psychology*, *44*(3), 912–919. https://doi.org/10.1016/j.jesp.2007.10.002

Vorauer, J., Gagnon, A., & Sasaki, S. J. (2009). Salient Intergroup Ideology and

Intergroup Interaction. *Psychological Science*, *20*(7), 838–845. https://doi.org/10.1111/j.1467-9280.2009.02369.x

Vosgerau, J., Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2019). 99% impossible: A valid, or falsifiable, internal meta-analysis. *Journal of Experimental Psychology: General*, *148*(9), 1628–1639. https://doi.org/10.1037/xge0000663

Walker, I., & Smith, H. J. (2002). *Relative Deprivation: Specification, Development, and Integration*. Cambridge University Press.

Wallrich, L. (2021). *TimesaveR* [R]. https://github.com/LukasWallrich/timesaveR

Wallrich, L., Birtel, M. D., Abbott, N., West, K., Drury, L., Meleady, R., & Gkinopoulos, T. (under review). *The Interaction between Positive and Negative Contact: The Mediating Roles of Intergroup Anxiety, Empathy and Threat*. https://osf.io/jpqfr/

Wallrich, L., Palmer, S. B., & Rutland, A. (2021). *Adolescents challenging discrimination: The benefits of a perspective-taking and action-planning intervention on self-efficacy*. OSF Preprints. https://doi.org/10.31219/osf.io/rzvud

Wallrich, L., West, K., & Rutland, A. (2020a). Painting All Foreigners With One Brush? How the Salience of Muslims and Refugees Shapes Judgements. *Journal of Social and Political Psychology*, *8*(1), 246-265–265. https://doi.org/10.5964/jspp.v8i1.1283

Wallrich, L., West, K., & Rutland, A. (2020b). *The dynamic relationship between valuing diversity and intergroup contact: Is valuing diversity a precursor to contact and mediator of diverse contact effects?* OSF Preprints. https://doi.org/10.31219/osf.io/tynju

Wallrich, L., West, K., & Rutland, A. (2021). Civic understanding mediates of the

effect of educational tracks on voting intentions in German schools. *Citizenship, Social and Economics Education*, 20471734211016500. https://doi.org/10.1177/20471734211016501

Wang, C. S., Kenneth, T., Ku, G., & Galinsky, A. D. (2014). Perspective-Taking Increases Willingness to Engage in Intergroup Contact. *PLOS ONE*, *9*(1), e85681. https://doi.org/10.1371/journal.pone.0085681

Wells, G. L., & Windschitl, P. D. (1999). Stimulus Sampling and Social Psychological Experimentation. *Personality and Social Psychology Bulletin*, *25*(9), 1115–1125. https://doi.org/10.1177/01461672992512005

West, K., Greenland, K., & Laar, C. van. (2021). Implicit racism, colour blindness, and narrow definitions of discrimination: Why some White people prefer 'All Lives Matter' to 'Black Lives Matter.' *British Journal of Social Psychology*. https://doi.org/10.1111/bjso.12458

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399. https://doi.org/10.1002/sim.4067

Whitley, B. E. J., & Webster, G. D. (2018). The Relationships of Intergroup Ideologies to Ethnic Prejudice: A Meta-Analysis: *Personality and Social Psychology Review*. https://doi.org/10.1177/1088868318761423

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wilder, D. A. (1984). Intergroup contact: The typical member and the exception to the rule. *Journal of Experimental Social Psychology*, *20*(2), 177–194. https://doi.org/10.1016/0022-1031(84)90019-2

Williams, R. M. J. (1947). The reduction of intergroup tensions: A survey of research on problems of ethnic, racial, and religious group relations. *Social Science*

Research Council Bulletin.

Wilson, S. J., Polanin, J. R., & Lipsey, M. W. (2016). Fitting meta-analytic structural equation models with complex datasets. *Research Synthesis Methods*, *7*(2), 121–139. https://doi.org/10.1002/jrsm.1199

Wölfer, R., Jaspers, E., Blaylock, D., Wigoder, C., Hughes, J., & Hewstone, M. (2017). Studying Positive and Negative Direct and Extended Contact: Complementing Self-Reports With Social Network Analysis. *Personality and Social Psychology Bulletin*, *43*(11), 1566–1581. https://doi.org/10.1177/0146167217719732

Wurdock, C. J. (1981). Neighborhood Racial Transition: A Study of the Role of White Flight. *Urban Affairs Quarterly*, *17*(1), 75–89. https://doi.org/10.1177/004208168101700105

Yogeeswaran, K., Verkuyten, M., & Ealam, B. (2020). A way forward? The impact of interculturalism on intergroup relations in culturally diverse nations. *Group Processes & Intergroup Relations*, 1368430220918651. https://doi.org/10.1177/1368430220918651

Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2, Pt.2), 1–27. https://doi.org/10.1037/h0025848

Zebrowitz, L. A., White, B., & Wieneke, K. (2008). Mere Exposure and Racial Prejudice: Exposure to Other-Race Faces Increases Liking for Strangers of That Race. *Social Cognition*, *26*(3), 259–275. https://doi.org/10.1521/soco.2008.26.3.259

Zhang, J. (2004). A Dynamic Model of Residential Segregation. *The Journal of Mathematical Sociology*, *28*(3), 147–170. https://doi.org/10.1080/00222500490480202

Zhang, Z. (2014). Monte Carlo based statistical power analysis for mediation

models: Methods and software. *Behavior Research Methods*, *46*(4), 1184–

1198. https://doi.org/10.3758/s13428-013-0424-0

Appendices

Appendix 2.1 – Cross-lagged panel models with observed variables

Instead of the models with latent variables presented in the main text, the models below are based on observed variables, namely the mean response per scale.

Figure A2.1. *Cross-lagged panel model connecting contact and valuing diversity*
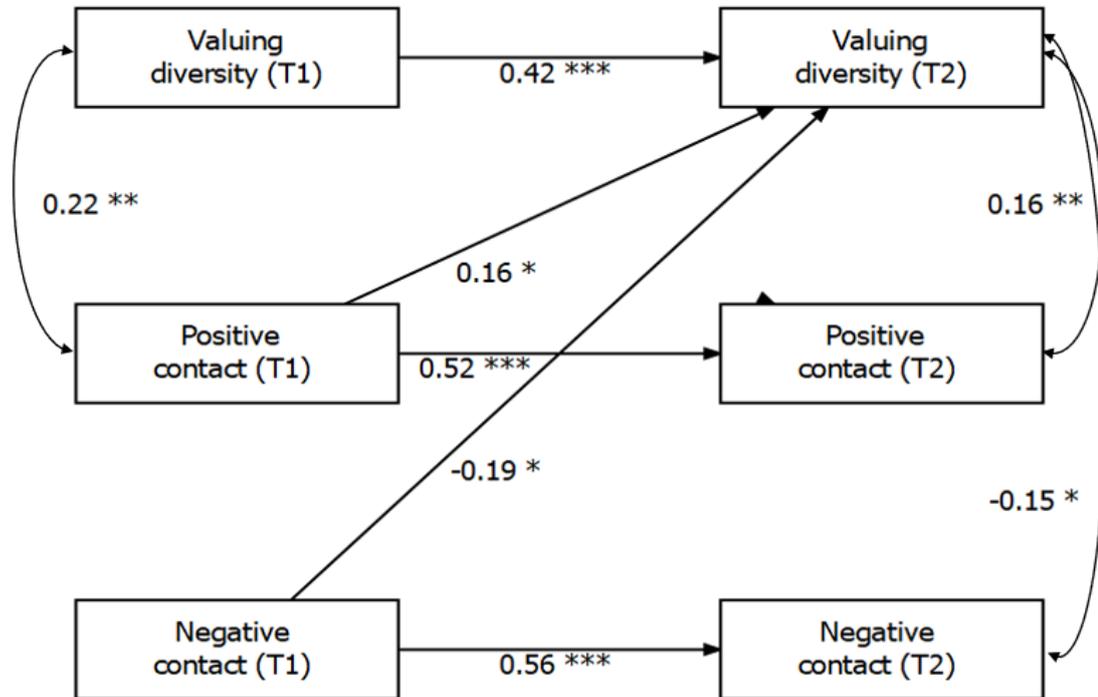
Notes: Standardized coefficients; paths with *p*-values above .1 are not shown for simplicity. * $p < .05$, ** $p < .01$, *** $p < .001$

Table A2.1. *Results of the Cross-Lagged Panel Model Connecting Contact and Valuing Diversity*

Predictor	Valuing diversity (T2)		Positive contact (T2)		Negative contact (T2)	
	β [95% CI]	p	β [95% CI]	p	β [95% CI]	p
Valuing diversity (T1)	0.42 [0.28, 0.56]	< .001	0.10 [-0.03, 0.24]	.133	-0.07 [-0.20, 0.07]	.353
Positive contact (T1)	0.16 [0.02, 0.30]	.024	0.52 [0.38, 0.66]	< .001	-0.03 [-0.17, 0.11]	.666
Negative contact (T1)	-0.19 [-0.33, -0.04]	.014	-0.00 [-0.15, 0.14]	.959	0.56 [0.42, 0.71]	< .001

Notes:

Standardised coefficients estimated with full-information maximum-likelihood.

T1: Measured at timepoint 1   T2: Measured at timepoint 2

Appendix 4.1 – Logistic regression as robustness check

Table A4.1. *Proportional odds logistic regression models predicting neighbourhood choice*

	OR [95% CI]	*Stand. OR [95% CI]*
Valuing diversity	1.66 *** [1.47, 1.87]	1.44 *** [1.32, 1.57]
Positive attitude towards foreigners	2.07 *** [1.92, 2.23]	2.42 *** [2.20, 2.65]
Political orientation *(right-wing)*	1.15 *** [1.10, 1.21]	1.27 *** [1.18, 1.38]
Age	0.96 *** [0.96, 0.97]	0.51 *** [0.47, 0.55]
Gender *(female)*	1.11 [0.97, 1.28]	1.11 [0.97, 1.28]
Region *(East)*	0.48 *** [0.40, 0.58]	0.48 *** [0.40, 0.58]
Education	1.25 *** [1.17, 1.33]	1.27 *** [1.18, 1.36]
N	2618	
R^2	.43	

Note: Given that dummy variables lose their interpretability when standardised, OR are only standardised for continuous predictors. † $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Appendix 5.1: Permutation tests of programme effect

Regression to the mean implies that participants who give low responses at the start are more likely to give higher responses later, while the reverse is true for those who give high responses. When scales are bounded, as they were in this case, the concern becomes even more acute since a sizeable share of participants can only move in one direction. Therefore, it can be instructive to estimate a null distribution that takes the observed regression to the mean into account.

Given that regression to the mean refers to a statistical relationship, it does not depend on the timing of measures. Rather, it is as likely that 'outlier' responses at the start revert to the mean as it is that mean responses at the start are followed by extreme responses. Therefore, permutation tests that respect the pairing of responses but randomise their order can result in a null-distribution that takes into account the observed regression to the mean (Furrow, 2019). They cannot (easily) account for the nesting of participants in cohorts, so that they are only suitable as a supplementary analysis here. I run such permutation tests with 5,000 permutations; the resulting null-distribution for pre-post changes depending on the initial response is shown in the violin plots in Figures S5.1 and S5.2. In each of the Figures, the orange/light line indicates the change score for each initial response below which 95% of the permutation results fall, which serves as a threshold for statistical significance. The green/darker line shows the observed change for participants who gave that initial response. Figures A5.1 shows that with regard to participants preference for working in diverse teams, the observed changes exceeded those that could reasonably be expected under the null-hypothesis for participants who reported an initial preference for working in diverse teams of at least 3. This was the case for 88% of participants, which confirms the finding that the intervention had a significant impact on that outcome. For the smaller group of participants with lower initial response, the

220

estimated programme effect was still larger than the mean effect under the null

hypothesis, but it did not reach statistical significance.

Figure A5.1.

*Results of permutation tests of impact of intervention on preference for diversity*
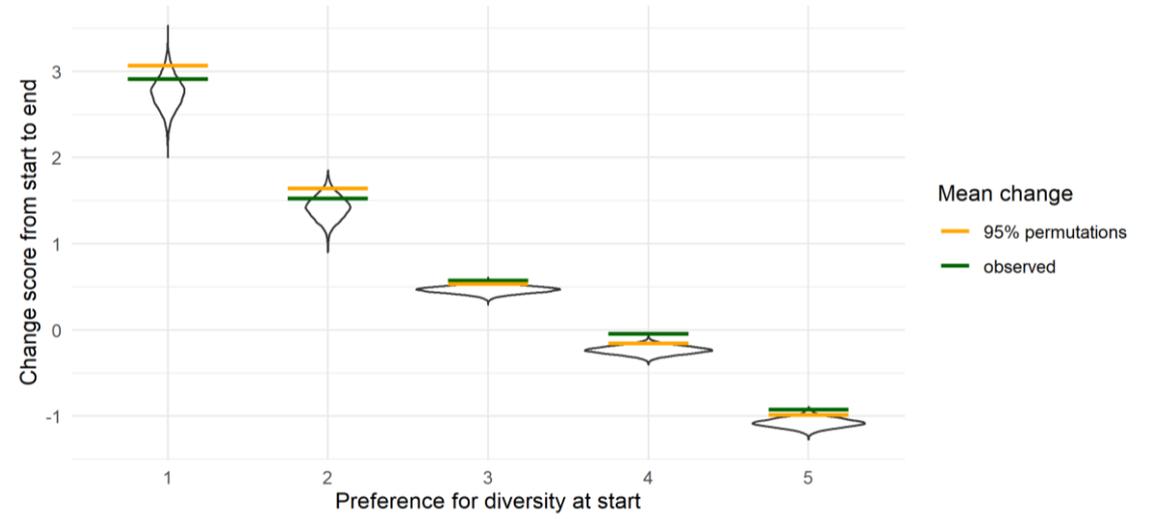
Figure A5.2, conversely, confirms that the programme had no significant

effect on participants' beliefs regarding the instrumental value of diversity. In fact,

for most baseline responses, the observed changes were below those to be expected

due to regression to the mean.

Figure A5.2. *Results of permutation tests of impact of intervention on instrumental valuing of diversity*
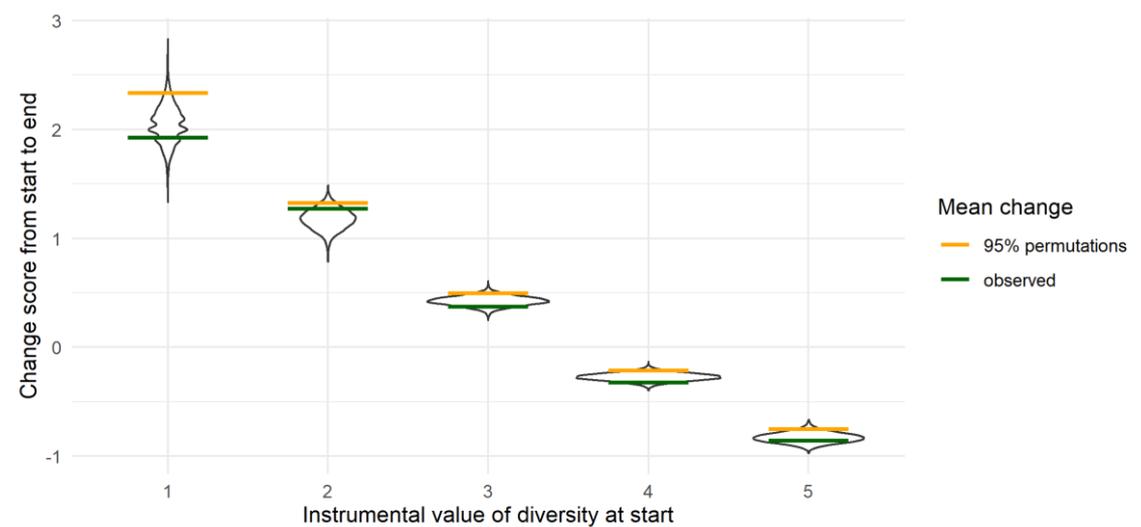
Appendix 6.1: Descriptive statistics and correlations across conditions

Table A6.1. *Descriptive statistics and correlations for full sample*

Variable	M (SD)	N	1	2	3	4	5
1. Valuing diverse teams	2.85 (1.02)	538					
2. Valuing diverse societies	3.00 (1.14)	280	.47 *** [.35, .58]				
3. Contact valence	3.86 (0.78)	994	.10 * [.01, .18]	.16 ** [.04, .28]			
4. Frequency of discussing differences	3.37 (0.91)	997	.08 † [-.01, .16]	.20 ** [.08, .31]	.32 *** [.27, .38]		
5. Interest in contact	3.39 (0.85)	384	.08 [-.06, .21]	.15 * [.03, .26]	.33 *** [.24, .42]	.25 *** [.15, .34]	
6. Commitment to address diversity	3.73 (0.87)	979	.15 *** [.06, .23]	.13 * [.01, .25]	.33 *** [.27, .39]	.32 *** [.26, .37]	.26 *** [.17, .35]

Note. M and *SD* are used to represent mean and standard deviation, respectively. *N* denotes the number of responses to that variable. Values in square brackets indicate the 95% confidence interval for each correlation (calculated with *pairwise* deletion of cases with missing data).

† $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$