Contemporary anti-racism
A review of effective practice

Jehonathan Ben, David Kelly, and Yin Paradies

Introduction

Anti-racism can be minimally defined as ‘forms of thought and/or practice that seek to confront, eradicate and/or ameliorate racism’ (Bonnett, 2000, p. 4) and as ‘ideologies and practices that affirm and seek to enable the equality of races and ethnic groups’ (Bonnett, 2006, p. 1099). Anti-racism practice has expanded remarkably over the past decades (Paradies, 2016; Pettigrew & Tropp, 2006). At the same time, evidence as to what works in confronting, eradicating and ameliorating racism, or, complementarily, how to enable and affirm racial/ethnic equality, remains limited. Paluck (2016, p. 147), for example, asks, ‘What do social scientists know about reducing prejudice in the world?’ before concluding that we know ‘very little’. And indeed, relative to the amount of anti-racism work underway, few evaluations have discerned interventions’ causal effects, limiting our understanding of interventions’ effectiveness. Real-world field experiments with longitudinal bearings (Paluck & Green, 2009; Paluck, Green, & Green, 2018) are especially well placed to answer questions about the extent and manners by which racism may be curbed, but remain particularly uncommon.

In this chapter, we examine research on anti-racism practice, focusing on effective approaches to tackling racism and interrelated phenomena like prejudice and racial/ethnic discrimination. In focusing on effectiveness, we examine the extent to which interventions produce measurable, positive changes. We draw especially on recent meta-analyses, reviews, and experimental (field- and laboratory- based) studies. First, we briefly summarise four central approaches to tackling racism, and synthesise key findings concerning effective anti-racism practices per approach as well as across approaches. We then consider the possibility that hindrance to anti-racism efforts may come from initiatives themselves, resulting in counterproductive ‘backlash effects’, and we discuss how these may be avoided. The chapter concludes with a broader discussion of current knowledge and implications for future research directions.

Approaches to anti-racism

Anti-racism approaches are highly diverse, spanning everything from prejudice reduction to conflict resolution to collective action (Paradies, 2016); and from reducing the incidence of racism to empowering racialised subjects to fostering a radical indifference to race (Hage,
This includes examples such as virtual reality experiments (Banakou, Hanumanthu, & Slater, 2016) and participatory theatre projects (Sonn et al., 2015). Here we focus on some of the most commonly used anti-racism approaches and their effectiveness in addressing racism, namely: (1) intergroup contact; (2) training and education; (3) communications and media campaigns; and (4) organisational development. These approaches sometimes overlap or can be used in combination to reinforce one another; organisational development may, for instance, feature a component of diversity training, while diversity training and media campaigns may involve a degree of intergroup contact.

**Intergroup contact**

Intergroup contact is a broad anti-racism approach that has been extensively implemented and studied, and has arguably become the most important approach for reducing prejudice (Paluck, Green, & Green, 2018). Grounded in Gordon Allport’s (1954) intergroup contact theory, it is predicated on five ‘optimal contact conditions’ for successfully reducing intergroup conflict and increasing harmony: (1) equal status between interacting groups; (2) common goals between groups; (3) intergroup cooperation; (4) support from authorities, law, or custom; and (5) situations that allow for developing personal acquaintance and friendships through meaningful, repeated contact (Al Ramiah & Hewstone, 2013; Pettigrew & Tropp, 2006). Contact may take direct (face-to-face) or indirect forms (i.e. as imagined, extended, vicarious or virtual) that can both be effective in reducing prejudice (Brown & Paterson, 2016). Educational settings like schools and universities are the most popular sites for intergroup contact interventions, followed by workplaces and organisations, and a host of other settings (Kalinoski et al., 2013; Pettigrew & Tropp, 2006).

Intergroup contact reduces prejudice through various mechanisms. Some of the best-studied and most important mechanisms include affective processes that decrease intergroup threat, anxiety and symbolic threat (i.e. anticipating harmful consequences), enhance self-disclosure, increase empathy and perspective taking, and alter group norms and social categorisations (Dovidio et al., 2017). The impact of intergroup contact on attitudes can generalise beyond the individual out-group members encountered and towards their greater group (Lemmer & Wagner, 2015; Pettigrew & Tropp, 2006). The extent to which the individuals encountered (whether directly or indirectly) are perceived as typical members of the out-group makes generalisation more likely (Al Ramiah & Hewstone, 2013; Brown & Paterson, 2016; Dovidio et al., 2017). Meta-analyses have shown that programme effects can persist after the programme has ended (Beelmann & Heinemann, 2014; Lemmer & Wagner, 2015).

Research demonstrates that the quality/favourability of contact has a stronger effect on attitudes than contact quantity, while the duration of contact also matters, with sustained contact becoming more positive over time, up to a point of diminishing returns (Dovidio et al., 2017). A balanced ratio of majority to minority group members in contact situations makes contact more effective in reducing prejudice, as it can maximise opportunities for interaction and reduce perceived intergroup threat (Al Ramiah & Hewstone, 2013). As to participants, college-aged students are more strongly impacted compared to adults (Pettigrew & Tropp, 2006), and participants who are highly prejudiced and/or for whom contact experiences are relatively novel may be more strongly impacted as having ample room for attitudinal change (Al Ramiah & Hewstone, 2013). While meeting Allport’s conditions for optimal contact is associated with greater prejudice reduction, not all conditions may be required to reduce prejudice, and contact may not always lead to positive attitudes, whereas
the attitudes of majority group members toward minority group members are often more strongly affected than vice versa (e.g. Dovidio et al., 2017; Pedersen et al., 2011; Ülger et al., 2017).

**Anti-racism training and education**

Various anti-racism initiatives use forms of education and training to enhance cultural competency or challenge discrimination, often at workplaces and schools. The most common intervention is diversity training, which draws on programmes that aim to increase positive intergroup behaviours and decrease prejudice or discrimination towards (perceived) out-group members (Pendry, Driscoll, & Field, 2007, p. 29). Training can decrease discriminatory attitudes and beliefs among most participants, but disturbingly can also increase them for a small, yet sizeable group (Paradies et al., 2009).

A meta-analysis of curricular and co-curricular diversity-related activities found they had moderate effect in reducing racial bias, where effects were stronger for white students compared with non-white students (Denson, 2009). Another meta-analysis of diversity training programmes, of which over a third focused on race, ethnicity, culture and/or religion, found considerable effects on cognitive-based and skill-based outcomes, and somewhat smaller effects on affective-based outcomes (Kalinoski et al., 2013). Other reviews have been less supportive of diversity training. A recent review suggests that training ‘can lead to both positive and negative social justice outcomes’, including studies that find that training can both reduce and increase discrimination (Alhejji et al., 2016, p. 5). A review of the impact of diversity training on management composition in private organisations found that it was generally ineffective in increasing the share of black American managers (Kalev, Dobbin, & Kelly, 2006).

Discussing effective manners of training and education, several studies stress the importance of explicitly discussing racism, within a safe space for open and frank dialogue (Paradies et al., 2009; Pedersen et al., 2011). Bezrukova, Jehn, and Spell (2012) point to effective aspects of diversity training such as using multiple instruction methods, and integrating training as part of systematic, planned organisational development rather than as standalone components. Focusing on participants from a range of racial/ethnic/cultural/religious backgrounds was deemed more effective than focusing on a specific group (Bezrukova, Jehn, & Spell, 2012; Paradies et al., 2009). In workplaces, engaging managers in promoting diversity, propelling them to increase their contact with members of different out-groups and encouraging their accountability are effective in reducing bias and increasing workplace diversity (Dobbin & Kalev, 2016). Both voluntary and compulsory training may be effective under certain situations. Although both have been subject to critique, for example for ‘preaching to the converted’ when voluntary, or for inducing resistance when forced (Bezrukova, Jehn, & Spell, 2012; Dobbin & Kalev, 2016).

To be effective, training should be neutral and informal, and provide accurate information about the out-group based on multiple disciplines, preferably using multiple instruction methods (e.g. readings, audio-visuals, small group discussion and role plays) (Pedersen et al., 2011; Torino, 2015). Training should be tailored to each organisation, linked to operational goals, and specifically address behaviour, while trainers should engage participants respectfully and interactively, build and invoke social norms, enhance awareness, attitudes and skills, and encourage intergroup contact where appropriate. They preferably should be ‘insiders’, from various racial/ethnic/cultural backgrounds, and with experience and/or qualifications in organisational change (Paradies et al., 2009).
Communications and media campaigns

Communications and media campaigns against racism usually consist of large-scale initiatives drawing on various media forms and platforms, sometimes use social marketing techniques, and are frequently assessed via experiments (naturalistic and lab-based). Media and communications can aggravate stereotyping, prejudice and discriminatory behaviour towards different racial/ethnic groups (Willard, Isaac, & Carney, 2015), but they can also raise awareness of race-based discrimination, impact attitudes and behaviours, and help develop or strengthen positive social norms (Paluck & Green, 2009; Paradies et al., 2009).

Although many ‘real-life’ campaigns exist, their impact has rarely seen rigorous assessment (Donovan & Vlais, 2006). A review of 13 media interventions addressing different forms of prejudice that used field experiments, noted that interventions were mostly conducted in schools and showed ‘suggestive’ results in their impact on empathy, perspective taking and social norms, as well as on using narratives for persuasive purposes (Paluck & Green, 2009). Other reviews have portrayed a similarly mixed picture of the effectiveness of anti-racism campaigns using communications and social marketing approaches to reduce discrimination and support diversity (Aboud et al., 2012; Donovan & Vlais, 2006; Rankine, 2014).

Popular racism-reduction methods that rely on audio-visual media (e.g. television and film) are vicarious and imagined intergroup contact (i.e. observing or imagining other people in intergroup contact situations) (Cadenas et al., 2018; Murrar & Brauer, 2018). Vicarious contact may produce constructive perceptions of the out-group and positive emotional responses towards them. It may capitalise on exposure to typical, favourable and salient counter-stereotypical media exemplars, and on identification with in-group characters that engage in positive contact. Other initiatives provide new information to challenge existing stereotypes and norms, or invoke emotions that are conducive to tackling prejudice (e.g. empathy). Studies assessing the impact of such media forms, often in laboratory contexts demonstrate mixed findings (e.g. Castelli et al., 2012; Igartua & Frutos, 2017; Vittrup & Holden, 2011).

Communications campaigns have stronger effects when they address specific negative beliefs rather than focus on generating positive feelings, and when focusing on various individuals from one ethnic/racial group at the time rather than promote broad purposes like ‘diversity’ and ‘multiculturalism’ (Donovan & Vlais, 2006). To increase their effectiveness, campaigns should identify beliefs that underlie expressions of racism, challenge racism and promote anti-racism as a prescriptive norm, and highlight perceived, appreciated similarities between groups, especially where negative beliefs are based on ignoring such similarities (Donovan & Vlais, 2006). Campaigns should involve the affected group as a visible part of the campaign, engage media personnel to change media representations, use advocacy and activism to generate wider support and impact policy, provide opportunities for discussion and interaction across groups and aim to pre-emptively address counterarguments (Donovan & Vlais, 2006). Finally, lab experiments suggest that nonverbal behaviours (in audio-visual representations) carry strong impact and need to be considered (Castelli et al., 2012), and that messages’ content should be explicated (Vittrup & Holden, 2011).

Organisational development

Organisational development has been the least reviewed of the approaches we discuss here. Its projects typically use development and change processes to assess or ‘audit’ organisational functions in order to address discrimination and endorse diversity (Paradies et al., 2009, p. 52).
Such projects may implement new organisational policies, plans or operational processes, model and enforce non-discriminatory standards, and work to impact social norms and wider societal change. They may use the three aforementioned approaches, as well as develop resources (e.g. teacher professional development, journalist guides), draw on organisational leadership and deploy conflict resolution approaches (Paradies et al., 2009, pp. 52–53). Multiple studies have documented the effects of individual initiatives, suggesting promising results in relation to organisations in areas such as healthcare (Weech-Maldonado et al., 2018), education (Hagopian et al., 2018) and workplaces (Ferdinand, Paradies, & Kelaher, 2017; Treanerry, Franklin, & Paradies, 2012).

Practices considered effective include development of a shared organisational vision, clear goals, measurable outcomes, and organisational accountability, as well as customisation based on local social, political and other contexts (Paradies et al., 2009; Treanerry, Franklin, & Paradies, 2012). Initiatives are more impactful when cultivating transparency, trust and the exchange of information (Ferdinand, Paradies, & Kelaher, 2017; Paradies et al., 2009). Organisational development tends to involve multiple layers and elements and large-scale public institutions. Although such complexity is important, it may also introduce inherent challenges to organisational development that must be engaged with (Ferdinand, Paradies, & Kelaher, 2017; Spaaij et al., 2016). Using a ‘whole of organisation’ approach (Treanerry, Franklin, & Paradies, 2012), and a detailed strategic plan addressing multiple aspects of organisational functioning are also considered effective practices (Treanerry, Franklin, & Paradies, 2012, and see discussion in Abramovitz & Blitz, 2015).

Effective practice across approaches

Various practices have demonstrated effectiveness across two or more of the approaches we discussed. At the outset, interventions should be carefully planned, mapped, and well developed, attending to areas like their objectives and materials, while involving a management group and various stakeholders (Donovan & Vlais, 2006; Paradies et al., 2009). Interventions that are theory-driven or based on solid theoretical foundations have been considered more effective in curbing prejudice (Aboud et al., 2012; Pettigrew and Tropp, 2006). The significant place of evaluation research throughout interventions’ lifespan has been reiterated across approaches, involving the use of formative research, preliminary testing of objectives and methods (Aboud et al., 2012; Donovan and Vlais, 2006), allocation of sufficient resources for planning and implementation (Paradies et al., 2009; Rankine, 2014), and preferably using evaluations that consist of pre- and post-testing, randomisation, and delayed outcome measures (Paradies et al., 2009; Pedersen et al., 2011).

Strong support from organisational leaders and champions has been considered crucial for programme effectiveness (Paradies et al., 2009; Treanerry, Franklin, & Paradies, 2012). There is usually an advantage to longer programmes and to programmes that consist of many sessions (Aboud et al., 2012; Kalinoski et al., 2013; Paradies et al., 2009; Pedersen et al., 2011), and wide support for programmes that emphasise invoking empathy (Dovidio et al., 2017; Mazziotta et al., 2011; Paluck & Green, 2009; Pedersen et al., 2011). Last, some have recommended that these approaches work best when integrated and that initiatives are more impactful when collaborating with other organisations involved in anti-racism work (Bezrukova, Jehn, & Spell, 2012; Paradies et al., 2009). The use of multiple, multi-level (e.g. state authorities, organisations), reinforcing approaches, can render interventions more effective (Paradies et al., 2009; and see; Ferdinand, Paradies, & Kelaher, 2017; Johnson, Antle, & Barbee, 2009; Weech-Maldonado et al., 2018 for effective combinations of organisational development and training).
Within the scope of this chapter, we are unable to discuss in detail real-world examples of practices that embody many of these recommended approaches previously. For further reading, refer to exemplary work on intergroup dialogue (Rodríguez et al., 2018); training (Johnson, Antle, & Barbee, 2009); organisational development (Weech-Maldonado et al., 2018); and media campaigns (Paluck, 2009, 2010).

**Backlash effects**

Anti-racism practice can have unintended consequences that may deter its effectiveness. The possibility of such backlash should be considered and pre-empted. Backlash towards anti-racism programmes, policies or practices, denotes forms of resistance that have the potential to strengthen prejudiced attitudes and negative relations between in-groups and out-groups. It can occur everywhere, from small-scale training to national populations, in relation, for example, to multicultural policies (Hewitt, 2005). Backlash works in ways that range from affective-based measures such as negative emotions, through to cognitive forms such as attitudes and perceptions, and is expressed in behaviours that negatively affect the outcomes of programmes, policies or practices (Kidder et al., 2004, p. 77). Anticipating backlash can itself become a form of backlash, by precluding the initiation or implementation of anti-racism initiatives, which can manifest as a refusal or withdrawal of basic resources and services (Bakan and Kobayashi, 2004, 2007a, 2007b).

Research on perspective taking demonstrates that imagining one’s ‘self’ increases the potential to negatively evaluate oneself which can entrench racial prejudice (Vorauer & Sasaki, 2014). Perceived threat to notions of ‘self’ in interpersonal contexts (e.g. because of high dissimilarity with others) may also provoke backlash effects (Sassenrath, Hodges, & Pfatteicher, 2016). Backlash is always a risk in intergroup encounters that confront negative behaviours (Focella, Bean, & Stone, 2015). Some studies have explored how intergroup encounters, where minority groups seek to reduce harm and prejudice by confronting perpetrators, can produce backlash effects that increase prejudiced attitudes (Vorauer & Sasaki, 2009). Studies on confronting prejudice report various forms of backlash, including dislike for the person and their perceived in-group (Czopp, Monteith, & Mark, 2006), while racial discrimination reported by African American participants in educational settings resulted in the stigmatisation of them as complainers (Kaiser & Miller, 2001). Other studies found that out-group members who confronted discriminatory behaviour were more likely to be negatively assessed (or have negative attitudes of them reinforced) by in-group members who were being discriminatory (Gulker, Mark, & Monteith, 2013; Rasinski & Czopp, 2010). This is especially true when persons who hold strong views of meritocracy are confronted, where evaluations of the ‘confronter’ are particularly negative (Schultz & Maddox, 2013), and likewise for those who adopt a colour-blind perspective (Zou & Dickter, 2013).

Forms of framing in anti-racist interventions play a significant role in manifestations of backlash. Framing diversity as ‘good’ within organisations (and not as ‘fair’) may broaden definitions of diversity to include axes of difference beyond race, which can lead, unintentionally, to deprioritising hiring applicants from racial minority backgrounds (Trawalter, Driskell, & Davidson, 2016). The colour-blind approach may often be seen as remedy to such paradoxical framings, although it can reinforce exclusive institutions that maintain unequal power structures in society (Smith & Mayorga-Gallo, 2017). In institutions where cultural diversity is widespread, multicultural policies are generally likely to reduce stereotyping and prejudice, whereas colour-blind practices and policies (ignoring or avoiding race and racial categories) may enhance stereotyping and prejudice, and may leave discrimination undetected (Plaut, 2014). Also, while a focus
on multiculturalist frameworks produces a higher rate of success in reducing biases than colour-blind frameworks, negative outcomes routinely occur, and multiculturalist frameworks can ironically produce higher instances of racialised essentialism, reproducing beliefs that race is biologically determined and fixed (Wilton, Apfelbaum, & Good, 2019).

When diversity inclusion frameworks are imposed upon – or topped-down from within – organisations, research has also found paradoxical outcomes. Dobbin, Schrage, and Kalev (2015) examined the effects of workplace innovations like training, on managerial diversity in 816 U.S. workplaces over 30 years. Accountability in the implementation of such innovations, such as monitoring the impact of hiring reforms through ‘diversity managers’, improved outcomes and reduced the potential for workplace backlash. However, compulsory accountability frameworks often backfired, suggesting that frameworks should in such contexts be willingly implemented. Using data on approximately 500 high-profile employment discrimination lawsuits resolved in U.S. federal courts between 1996 and 2008, Hirsh and Cha (2017) found that court-mandated policy changes to reduce bias expanded opportunities for white women but not for other demographic groups, while policies to increase awareness of rights were associated with declines in managerial diversity. Verdicts with the most costly monetary payouts did not expand managerial diversity compared to more modest payouts, and can further a lack of diversity.

Rutchick and Eccleston (2010) show that when minority groups invoke shared identity characteristics with the majority group during diversity training, this may lead to negative outcomes. The level of backlash exhibited in these experimental studies is predominantly determined by the strength of relationship and identification that groups have with a (perceived) larger homogenous whole, like the nation. The extent to which groups identify as being emblematic of such an overarching whole determines the degree to which messages that invoke dominant-group diversity will be received as intended (Falomir-Pichastor & Frederic, 2013; Steffens et al., 2017).

Discussion

Assessments of intergroup contact, training and education, communications and media campaigns and organisational development have varied in their approaches and conclusions. Intergroup contact interventions have been frequently evaluated, resulting in a broad evidence base that suggests that contact can often reduce racism, especially prejudice. Training and education initiatives, and particularly cultural diversity/competence programmes, have been widespread, yet not much is known about the extent to which, and circumstances under which, they effectively address racism. Concerns about null and adverse effects have made diversity training a particularly contentious area, as suggested by several study titles, like ‘Why diversity programs fail’ (Dobbin & Kalev, 2016) and ‘Pointless diversity training’ (Noon, 2018). Other areas of anti-racism practice have seen far less evaluation. Communications and media campaigns show promise but also mixed findings, and have been scarcely evaluated outside the lab. Organisational development initiatives have been discussed individually (or as part of education/training initiatives), but to our knowledge have yet to be collectively reviewed or assessed regarding their effects.

Anti-racism’s limited evidence base calls for further comprehensive, fine-grained analysis. Field experiments are clearly a priority in this field because of the dynamic, real-life nature of many anti-racism initiatives and the change they seek to instigate. In addition to using randomisation and control, there remains a strong need for assessments to go beyond pre- and immediate post-test measures. Given that intervention effects may transform post-intervention (for example
diminish, or show delayed improvement) it is crucial that we develop better understanding of what happens long after initiatives end. Systematic reviews and meta-analyses of the relationship between training, communications and organisational development initiatives (and initiative components) and the reduction of racial prejudice, discrimination, and contingent outcomes pertaining to racism are also much-needed.

Recently, several studies have emerged that constitute promising future directions in anti-racism practice and research. Examples of innovative, effective methodologies include reading popular books (e.g. Harry Potter) to reduce prejudice that capitalise on processes like identification and perspective taking (Vezzali et al., 2015), embodiment of another racial/ethnic group’s (e.g. black) virtual body to reduce implicit racial bias (Banakou, Hanumanthu, & Slater, 2016; Peck et al., 2013), and exposure to extreme racist audio-visual content which can, paradoxically, lead individuals to reassess their current (less extreme) racist attitudes and beliefs (Hameiri, Bar-Tal, & Halperin, 2019; Hameiri et al., 2014, 2016).

Based on available research, anti-racism initiatives are particularly effective when they are carefully developed, theory- and evidence-based, longer-term, draw on clear objectives and explicit messages, and include rigorous, ongoing evaluation research (e.g. Johnson, Antle, & Barbee, 2009; Paluck, 2009, 2010; Rodríguez et al., 2018; Weech-Maldonado et al., 2018). Reviews also repeatedly stress, and successful initiatives demonstrate, the significance of collaboration, support from and ongoing engagement with and between stakeholders, from institutional leadership to affected groups. There are indications that the integration of various approaches, initiative components, methods and materials is effective in addressing racism. Emerging scholarship cautions us against ways in which anti-racism initiatives may do more harm than good and have led to efforts to understand how we can best avoid backlash. Key suggestions include avoiding negations, given that injunctions to ‘do not’ create unintended associations between subjects (Gawronski et al., 2008), enhancing participants’ self-affirmation (Stone et al., 2011; Watt et al., 2009), and being wary of framing diversity within an overriding identity category (e.g. ‘the nation’) that the dominant in-group is protective of (Falomir-Pichastor & Frederic, 2013; Steffens et al., 2017).

References


