

# Cohesion through participation? Youth engagement, interethnic attitudes, and pathways of positive and negative intergroup contact among adolescents: a quasi-experimental field study

James Laurence<sup>a,b</sup>

<sup>a</sup>Cathie Marsh Institute for Social Research, University of Manchester, Manchester, UK; <sup>b</sup>Wissenschaftszentrum Berlin für Sozialforschung, WZB Berlin Social Science Center, Berlin, Germany

## ABSTRACT

The paper explores how youth engagement (i.e. organised social participation in a group, club, or activity) can impact young people's interethnic attitudes, via pathways of positive and negative interethnic contact. To do so, it examines processes of interethnic cohesion occurring on a large-scale, nationally-implemented UK youth engagement scheme. Employing a quasi-experimental approach, using pre-test/post-test data on a sample of participants and a (propensity-score matched) control-group, analyses demonstrate that participation leads to positive changes in young people's interethnic attitudes, evident at least 4–6 months after participation ended. This improvement in attitudes is driven primarily by increases in young people's positive interethnic contact, while participation has no impact on young people's levels of negative interethnic contact. However, the impact of participation on interethnic attitudes depends on how much positive contact young people had *prior* to taking part: young people who joined the scheme with less frequent positive contact see substantially larger improvements in their levels of positive contact which, in turn, drives even greater improvements in their interethnic attitudes. These findings provide encouraging evidence that sites of youth engagement, especially national engagement schemes, can foster intergroup cohesion among adolescents; especially among those with less frequent positive contact in their daily lives.

## ARTICLE HISTORY

Received 19 March 2019  
Accepted 30 November 2019

## KEYWORDS

Youth engagement and social participation; positive and negative intergroup contact; interethnic attitudes and prejudice; youth and adolescence; quasi-experimental

## Introduction

As societies become increasingly ethnically diverse, the question of how socially cohesive we are, and how we build interethnic cohesion where it is needed most, is of growing interest (e.g. Europe Commission 2016; MHCLG 2018). In many ways, interethnic cohesion has come a long way. In the UK, for example, segregation for all ethnic groups is declining, leading to more mixing (Simpson 2012); on several indicators both majority and minority

**CONTACT** James Laurence  [james.laurence@manchester.ac.uk](mailto:james.laurence@manchester.ac.uk)

 Supplemental data for this article can be accessed <https://doi.org/10.1080/1369183X.2019.1700787>.

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group  
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ethnic groups continue to report increasingly positive attitudes towards one another (Storm, Sobolewska, and Ford 2017); while the latest census shows the mixed-ethnicity group is one of the fastest growing in the UK. However, despite these positive signals, signs of potential fractures remain: segregation is persistently high among certain ethnic groups (Catney 2013); research suggests ethnically diverse communities tend to be less cohesive places (van der Meer and Tolsma 2014), and, when segregated, may generate greater perceived-threat and interethnic tensions (Laurence et al. 2018); while events like Brexit highlight persistent anxiety towards immigration among sections of society (Ford and Goodwin 2017). Understanding how to preserve the gains made – and further augment cohesion – in diverse societies is thus of high importance.

This question is particularly important for young people. In 2011 in England, around 20% of the country described themselves as non-White British. However, in 2017, around 40% of births were recorded by their mothers as non-White British (Haines 2017). As our world is to our parent's generation, the world that children are growing up into will be increasingly different from that of today. Cultivating intergroup cohesion among young people is therefore crucial to equip them for a future where diversity is increasingly the norm. The question is, what means are available to help build positive intergroup relations among young people.

One approach may be through encouraging greater youth engagement; that is, organised, formal social participation in groups, clubs, or activities. Among adults, formal social participation has long been posited as a key driver of cohesion, especially tolerance and trust, through opportunities it provides to form ties with those different from ourselves (Blau 1977; Putnam 2000). In many ways, youth engagement activities form the youth-equivalent of adult associational involvement, providing 'young people the opportunity to interact and connect with a group of individuals in the pursuit of common goals' (Smith 1999, 555). Accordingly, youth engagement could help foster more positive relations among young people, particularly through increasing opportunities for more optimal, positive intergroup mixing (Watkins, Larson, and Sullivan 2007; Conner and Erickson 2017; Knifsend and Juvonen 2017; Laurence 2019b). However, there are also reasons to be cautious. Factors may inhibit or even adversely affect the ability of youth engagement to produce positive relations. For example, processes of homophily may inhibit positive contact, even under optimal conditions (Maoz 2002). Meanwhile, although mixing between different groups may bring opportunities for positive contact, it may also increase opportunities for negative contact, which can actually harm intergroup relations (Guffler and Wagner 2017; Laurence, Schmid, and Hewstone 2017).

The primary aim of this study therefore is to explore what effect youth engagement has on young people's interethnic attitudes, and the role that pathways of positive *and* negative interethnic contact play in this relationship. However, the effectiveness of engagement for improving participants' intergroup relations could depend on their experiences of contact with other ethnic groups *prior* to participating; in particular, how much *pre*-participation positive or negative contact they had. On one hand, youth engagement could be particularly beneficial for those joining with less frequent positive contact, or more frequent negative contact. By providing greater opportunities for positive contact among those who normally have few, or providing more positive interactions among those for whom negative experiences are the norm, youth engagement could lead to larger improvements in their outgroup attitudes. Yet, studies also show that individuals with fewer experiences

of positive, or more experiences of negative, contact can feel more anxious, or more threatened, by ethnic outgroups, which can, in turn, affect how they respond to contact situations (Amir 1969; Stephan and Stephan 1985; Hewstone and Brown 1986; Plant and Devine 2003). Potentially, those who, prior to participating, had fewer positive/more negative experiences of contact may actively try to minimise their intergroup contact during participation, try to keep it more superficial (and thus less optimal), and experience any contact they do have less positively/more negatively (Stephan and Stephan 1985; Van Zomeren, Fischer, and Spears 2007; Binder et al. 2009). As a result, participation could actually have a weaker positive-, or even a negative-, effect on their intergroup attitudes. The secondary aim of this study therefore is to test whether the impact of engagement on youth interethnic attitudes is conditional on their prior outgroup contact experiences.

To pursue these aims this study takes advantage of a unique data opportunity to explore processes of intergroup cohesion occurring on a large-scale, nationally-implemented youth engagement scheme: the UK National Citizen Service (NCS). NCS brings together young people to engage in a 3–4 week programme of civic and social activities (National Audit Office 2017). The scheme is open to all youth aged 16–17 (and some aged 15<sup>1</sup>), and involves high levels of uptake (in 2018 one in six of all eligible young people that year took part) (National Citizen Service Trust 2018). With pre-test/post-test data on a sample of participants and (propensity-score matched) controls, this study applies a quasi-experimental, intervention-design to explore how discrete periods of youth engagement can impact dimensions of intergroup cohesion. In doing so, it aims to significantly advance the nascent work in this area, including: applying robust, quasi-experimental approaches to address endogeneity in prior studies; studying a large-scale, nationally-implemented engagement scheme to strengthen the generalisability of findings; exploring the role of both positive and negative contact pathways for understanding how participation impacts intergroup attitudes; and examining which factors can affect *when* participation may have stronger/weaker impacts on young people's attitudes towards other ethnic groups.

## Theoretical framework

### *Youth engagement, (optimal) contact and interethnic cohesion*

Fostering mixing between ethnic groups has long been held as a key means of strengthening interethnic relations in society, as formulated under the contact hypothesis (Allport 1954; Pettigrew and Tropp 2006). Briefly, the concept follows that positive interethnic contact can help improve intergroup attitudes by reducing intergroup stereotyping and anxiety, and increasing empathy towards outgroups. However, the conditions under which contact occurs are believed to be critical to its effectiveness: 'optimal contact' for improving intergroup attitudes should be equal status, involving co-operative, common goal-orientated interactions, where positive cross-group mixing is sanctioned by authority (for example, through positive norms) (Allport 1954).

Extensive evidence shows that when some or all of these conditions are met intergroup contact can be especially effective at improving intergroup attitudes (although even positive contact in the absence of these conditions can be beneficial) (Hewstone and Brown 1986; Pettigrew and Tropp 2006). The difficulty is that much of the contact in society

may not be of the optimal kind for significant prejudice-reduction (Dixon, Durrheim, and Tredoux 2005); in fact, at times it may be more sub-optimal, risking the development of negative contact (Barlow et al. 2012). For example, interethnic contact in neighbourhoods can be intermittent or superficial (Dixon, Durrheim, and Tredoux 2005; Pettigrew and Tropp 2006). In schools, environments tend to be individualistic-, not co-operative-, orientated spaces, which may impede optimal contact (Watkins, Larson, and Sullivan 2007; Knifsend and Juvonen 2017). School status hierarchies, reinforced by peer-group dynamics, may also inhibit equal-status contact (Hamm, Bradford Brown, and Heck 2005; Bekhuis, Ruiters, and Coenders 2013). In light of this, one site that may be more conducive to optimal contact is spaces of youth engagement.

Youth engagement involves formal social participation through an organised club, group, or activity, undertaken with peers out of the home (Dworkin, Larson, and Hansen 2003; McGee et al. 2006). A key feature of youth engagement is that it involves discretionary, 'structured' or 'instrumental' (that is task-/skill-related) co-operative-orientated activities, within the structural-parameters of a defined organisation/group; this distinguishes it from more informal social participation, such as hanging out, or watching TV with friends (Larson and Verma 1999; Gilman 2001; Park 2004; Bundick 2011). Such youth engagement may occur in schools, for example through extra-curricular activities (Knifsend and Juvonen 2017) and service-learning programmes (Melchior et al. 1999), or outside of school environments, such as uniformed-groups or community-projects (Larson and Verma 1999; Pancer 2015). These sites of youth engagement may be effective at generating conditions more conducive to optimal interethnic contact (Watkins, Larson, and Sullivan 2007; Conner and Erickson 2017; Knifsend and Juvonen 2017; Laurence 2019b). They tend to involve common-goal orientated activities, requiring sustained co-operation between participants (Dworkin, Larson, and Hansen 2003; Watkins, Larson, and Sullivan 2007). Operating within the structural-parameters of a defined-group may cut across ethnic-group membership, broadening conceptions of the in-group via emerging superordinate identities (Gaertner et al. 2008; Levy and Hughes 2009). The contact is also more likely to occur on an equal-basis, especially when occurring outside of school environments, where peer-group pressures, homophily, and status-hierarchies can inhibit positive intergroup mixing (Watkins, Larson, and Sullivan 2007; Knifsend and Juvonen 2017). In addition, engagement activities often create positive social norms to promote positive mixing, sanctioned by group-leaders (Watkins, Larson, and Sullivan 2007; Pancer 2015; Conner and Erickson 2017).

If youth engagement can foster more positive interethnic contact then it could, in turn, build more positive interethnic attitudes. However, there are reasons why the impact of youth engagement may be limited. To be sure, even when there are opportunities to mix, intergroup contact may not occur (Maoz 2002). Contact occurring through youth engagement may also remain suboptimal (casual, short-term, superficial), limiting its effectiveness for prejudice-reduction (Al Ramiah and Hewstone 2012). Of particular concern is that, when sub-optimal contact conditions are met, or when an encounter involves hostility or harm, contact can actually be experienced negatively, resulting in increased prejudice (Amir 1969; Lolliot et al. 2015), and harming well-being more generally (Laurence 2019a). Studies of both neighbourhoods (Koopmans and Veit 2014; Laurence, Schmid, and Hewstone 2017) and schools (ten Berge, Lancee, and Jaspers 2017; Bayram Özdemir et al. 2018) show that as opportunities for positive-mixing increase so

too do opportunities for negative mixing (Pettigrew 2008). Accordingly, youth engagement could also generate more negative contact experiences, which not only impede any positive impact of participation on youth attitudes but potentially generate worse outcomes than if young people had not participated at all (Conner and Erickson 2017; Guffler and Wagner 2017).

A growing body of work demonstrates that, on the whole, youth engagement appears to improve young people's intergroup relations; particularly through fostering positive intergroup contact. Involvement in extra-curricular activities (such as hobby clubs, after-school groups, sports teams) is found to be positively related to cross-group ties (Moody 2001; Brown et al. 2003), and past engagement can predict future positive intergroup relations *via* more frequent positive contact (Knifsend and Juvonen 2017; Laurence 2019b). Reviews of US service-learning programmes, which combine in-school lessons on civic/social issues with community-service, also show participants tend to report improved intergroup attitudes and positive outgroup contact (Melchior et al. 1999; Morgan and Streb 2001; Holsapple 2012). Qualitative accounts of young people's engagement (e.g. in mentoring-, sports-, community-groups) detail how prejudice-reduction appears to emerge from more optimal-contact: co-operative, common-goal interactions, sanctioned by program-organisers, occurring on a more equal basis, outside of school peer-alliances (Reynolds 2007; Watkins, Larson, and Sullivan 2007; Lee et al. 2010).

Other studies, however, produce somewhat more mixed findings. Maoz (2002) performed a one-year study of 46 separate Arab/Jewish civic/social interventions in Israel, explicitly designed to foster positive contact. While most programmes led to some mixing, in 15% of cases contact was 'very low or entirely absent'. Therefore, even when opportunities for contact are maximised, contact does not necessarily materialise. In a quasi-experimental study of a 3-month Malaysian youth engagement programme (involving civic-skill training, identity building, and community-service), Al Ramiah and Hewstone (2012) found little evidence that participants experienced improvements in contact-quality or interethnic attitudes. Similarly, a quasi-experimental test of a large-scale, 10-month, youth engagement scheme in the US (AmeriCorps), found endorsements of intergroup contact were lower post-participation (Frumkin et al. 2009).

The type of contact participants experience appears crucial for understanding such differences in how participation affects intergroup attitudes (Erickson and O'Connor 2000; Conner and Erickson 2017; Guffler and Wagner 2017). In a study of Arab/Jewish youth in Israel, Guffler and Wagner (2017) found that involvement in an engagement programme had a short-term positive-effect on Arab youth. However, Jewish youth reported worsening attitudes post-participation, driven by reports of negative contact during engagement. Similar accounts have been observed in quasi-experimental tests of some service-learning studies. When schemes achieved more optimal-contact intergroup attitudes improved; yet, some engagement worsened intergroup attitudes through increasing sub-optimal and negative contact experiences (Erickson and O'Connor 2000; Conner and Erickson 2017).

Taken together, participation appears able to improve young people's intergroup attitudes *via* increasing interethnic contact. However, this depends on if, and what type of, contact occurs. When this contact is sub-optimal, or even negative, it may not only

suppress potential positive effects of engagement but lead attitudes to become worse than if one had not participated at all.

### *Heterogeneity in the impact of youth engagement*

While current studies are instructive in beginning to understand if, and why, youth engagement affects intergroup attitudes, research largely focuses on the impact of engagement across *all* participants; that is, the average impact of engagement on participants' attitudes. This approach, however, may mask important heterogeneity in how participation affects different groups of young people. One key driver of heterogeneity may be how much positive or negative intergroup contact a participant had *prior* to engaging.

On one hand, we might expect that engagement could be particularly effective at improving the outgroup attitudes of young people who, before engaging, had fewer experiences of positive contact, or more experiences of negative contact. Among those joining with fewer positive contact experiences, participation may provide greater opportunities to experience such contact. This could result in bigger increases in their levels of positive contact from participation, in turn, driving a more positive impact on their intergroup attitudes (relative to those who joined already reporting frequent positive contact). Similarly, youth who, prior to engaging, have more frequent negative contact in their daily lives may gain more opportunities to experience positive exchanges with outgroups through participation. This may lead to a larger decline in their negative contact, in turn, driving a more positive impact of participation on their intergroup attitudes.

On the other hand, we might also expect that having less positive contact or more negative contact prior to engagement could, instead of leading to more positive outcomes, actually limit participation's effectiveness. As discussed, factors such as the situational-context of contact (e.g. is it voluntary/involuntary, intimate/superficial), or its nature (being helped or harmed), can drive the valence of contact experiences (how positive/negative it is) (Amir 1969; Pettigrew 2008; Lolliot et al. 2015). However, alongside this, one's prior feelings towards outgroups can also affect how one responds to contact situations (Stephan and Stephan 1985; Plant and Devine 2003; Van Zomeren, Fischer, and Spears 2007; Binder et al. 2009). For example, feeling greater anxiety towards outgroups, or viewing outgroups as more threatening, can lead individuals to actively avoid outgroup contact where possible, and when not possible, try to keep interactions at a more superficial level, impeding optimal contact (Binder et al. 2009). These feelings can also affect the valence of an individual's contact when it does occur: individuals with higher intergroup anxiety/threat are found to experience any contact they do have less positively, and are likely to experience it more negatively (Stephan and Stephan 1985; Van Zomeren, Fischer, and Spears 2007).

One key driver of intergroup anxiety is having had less positive contact with outgroups in the past; which, for example, can predispose individuals to expect outgroup encounters to be negative (Stephan and Stephan 1985; Plant and Devine 2003). Anxiety and/or threat are also higher among individuals who have had more experiences of negative contact in the past (Amir 1969; Hewstone and Brown 1986). Taken together, if participants who had less positive contact or more negative contact prior to engaging also have greater outgroup anxiety and/or threat, then such participants may actively try to minimise their contact

during participation, try to keep it more superficial (and thus less optimal), and experience any contact they do have less positively/more negatively. This might result in a weaker positive-, or even a negative-, effect of participation on their intergroup attitudes. To the best of our knowledge, little research has explored this.

### **Present study and aims**

This study aims to explore: how involvement in youth engagement activities affects young people's interethnic attitudes; how far any impacts are driven by changing patterns of positive and negative intergroup contact; and what factors may drive heterogeneity in the impact of youth engagement (particularly how much positive/negative contact participants had *prior* to engaging). To pursue these aims, it will explore processes of intergroup cohesion occurring on a large-scale, nationally implemented youth engagement scheme – the UK National Citizen Service (NCS). NCS brings together young people aged 15–17, to undertake a period of engagement in social and civic activities in small teams of around 12–15 people, over a period of 3–4 weeks (National Audit Office 2017). The majority of young people go through the scheme during the summer (June through August), during which they undertake three phases of activities.<sup>2</sup> The first phase entails a residential-period away from home (usually at an outdoors activity centre) where participants undertake various 'outdoor pursuits' (e.g. raft building, climbing), helping develop team-buildings skills. During the second phase participants experience a second-period of living away from home, usually within university accommodation, during which time they are involved in various courses and activities aimed at developing life-skills, such as communication-skills or social action awareness. At the same time, living together with other team members away from home aims to foster experiences of independent living. In the final phase, participants return to their communities to design and implement a social action project, over a period of 60-hours; for example, building a communal garden in their local area (National Audit Office 2017). Further information on the programme-design can be found in the NCS audit-report (National Audit Office 2017).

The scheme itself is open to all young people in the country, aged 16–17, and some aged 15.<sup>3</sup> Young people who participate are also, at least on key socio-demographics, broadly representative of the UK youth population, with some over-representation among females (59% on the scheme versus 49% in society), ethnic minorities (32% versus 20%), young people eligible for free school meals (17% versus 8%), and those from the top quantile of community deprivation (13% versus 11%) (National Audit Office 2017). However, the scheme is also designed so that young people participate with peers drawn from the wider geographical area (UK government Local Authority districts) in which they live (National Audit Office 2017). Insofar as different Local Authorities have different levels of ethnic diversity, the opportunities young people (especially White British participants) have for mixing with ethnic outgroups in their teams may differ across the country. In more diverse Local Authorities, participants may have greater opportunities for mixing than their peers undertaking the scheme in more homogeneous Local Authorities (*see* 'Discussion' section).

This scheme offers a key test-bed to explore the impact of youth engagement on intergroup cohesion. NCS aims to promote a more cohesive society by providing opportunities for young people to mix with others from different backgrounds. The design of the scheme

also fulfils many of the criteria for why youth engagement should be effective for fostering optimal intergroup mixing: the activities are common-goal orientated, necessitating sustained, co-operative interactions; the scheme aims to foster positive norms of intergroup mixing, sanctioned by authority and promoted via team leaders; participating in small-teams of 12–15 people provides the structural-parameters of a defined-group; while group size likely limits opportunities for homophilic tendencies. Furthermore, as participation occurs outside peer-group pressures and status-hierarchies of schools, mixing is more likely to be on an equal-basis.

## Data and methods

### *Study design and sample*

Given NCS is a nationally-implemented programme, available to all age-eligible young people, random assignment of individuals into participation was not possible. Pre-participation and post-participation data are therefore collected on a sample of participants and a control-sample of young people who did not engage, constituting a non-equivalent control-group design. Data collection was commissioned by the Department for Digital, Culture, Media and Sport (DCMS). Questionnaires were administered to all NCS participants who took part over a four-week evaluation window in 2015. The evaluation-period was selected to generate a representative sample of participants. Participants were surveyed prior to beginning phase one, and then followed-up 4–6 months after participation had ended. The control-group is composed of young people who ‘expressed an interest’ in participating but did not engage during the summer 2015. These young people had attended a recruitment event for NCS and provided their contact details for further information but did not go on to participate that summer.<sup>4</sup> A random sample of these young people was selected to form the control-group, surveyed over the same pre-/post-participation period as the participants.

A mixed-mode approach was taken to the surveys, including postal-surveys at baseline and postal-/online-surveys at follow-up. At baseline, the response rates among the participant and control-groups were 85% and 47% respectively. The higher response-rate evident among participants is likely driven by questionnaires being part of broader information gathered about participants prior to beginning the programme. The lower response-rate among controls is in line with other (government) administered surveys which aim to reach young people outside of schools (a necessity, given not all of the target population were in school) (ARK 2009; Jessiman and Drever 2010; Gireesh, Das, and Viner 2018). Among those who responded to the initial questionnaire, a random-sample of  $n = 3985$  participants and  $n = 3985$  controls was selected to be re-contacted (the choice to re-contact only a sub-set of participants/controls was undertaken due to financial-restrictions on evaluation costs). Post-participation follow-up response rates for participants and controls were 48% and 51%, which is broadly in line with similar two-wave youth studies that are not part of parental panel surveys (Brown et al. 2008; Keating et al. 2010).

Missing data (from non-response and within-case missingness) can bias our ability to generalise our findings to the scheme as a whole. Several baseline factors among participants and controls are associated with post-test survey non-response, including being male, reporting lower generalised trust and exhibiting less frequent civic engagement.

Reassuringly, these factors predicting non-response were consistent among both participants and the control-group, strengthening confidence in the similarity of the groups (an exception being that only control-group members, not participants, with lower social-confidence were somewhat less likely to respond to the post-test survey). Missing within-case data is also below 2% for most variables (and no significant differences were observed between participants/controls). However, we undertake extensive testing using weighting and multiple-imputation to address potential sample-bias (see below).

### *Analytic approach*

Given the study's non-equivalent control-group design, we apply a difference-in-difference (DiD) approach, which compares pre-test/post-test changes in outcomes between the participant and control-group; the difference in these changes represents the impact of participation. Unbiased causal-estimates rely on the validity of the parallel-trends assumption: that participants/controls would have behaved in the same way over the test-period had participation not occurred. This assumption is strengthened when the similarity of participant- and control-groups is maximised. We take two steps to achieve this. Firstly, as outlined, control-group members are sampled from an 'expression of interest' pool of those who had proactively engaged with the NCS recruitment process. In theory, these individuals should be more similar to our sample of participants than young people in general (e.g. McAdam 1986). Secondly, we apply propensity-score matching (PSM) to further enhance similarity of participants/controls across a range of pre-participation characteristics.

Together this constitutes a propensity score matched difference-in-difference approach (PSM-DiD). This approach has several advantages. It accounts for time-invariant unobserved heterogeneity among individuals (e.g. personality-traits); knowing the exact timing of participation relative to pre-test/post-test measures helps address reverse causality and time-variant unobserved heterogeneity; while it also accounts for secular trends in our outcomes which might be occurring among all young people during this period (e.g. maturation-effects). In addition, a control-group helps address issues of regression to the mean; especially when testing for differences in effects based on pre-participation levels of contact (although see Daw and Hatfield 2018). At the same time, however, given selection into the participant-/control-groups was non-random, differences may remain between participants and the control-group introducing bias into the analysis if those who selected on to the scheme are systematically different in some way from those who did not. For example, such differences could lead to different secular-trends in the key outcomes between groups, potentially accounting for any apparent effects of participation (see 'Discussion' section for further details).

### *Measures*

#### *Inter-ethnic attitudes*

Ethnic outgroup attitudes are measured using a feeling thermometer question:

Everybody has different views about different groups of people. Imagine a thermometer that runs from zero to a hundred degrees, where 0–49 means you feel colder (less favourable); 51–100 degrees means you feel warmer (more favourable); and 50 means you don't feel

particularly warm or cold. Using this thermometer please write in how you feel about people from a *different* race or ethnicity to you.

Feeling thermometer measures display good test-retest reliability, correlate well with specific views of outgroups, and demonstrate strong discriminant validity with implicit-bias tests (Blair et al. 2010; Lolliot et al. 2015).

### *Positive and negative intergroup contact*

To measure patterns of positive and negative intergroup contact respondents were asked two questions:

People report having positive and negative social contact with others from all kinds of backgrounds. Thinking of your own experiences with people from a *different* race or ethnicity to you, how often, if at all, would you say have had ...

‘(a) positive or good experiences. For example someone being friendly to you, or making you feel welcome?’; and (b) ‘negative or bad experiences. For example someone being mean to you, or making you feel unwelcome?’ A 5-option likert scale of responses was used, including: ‘Never’ to ‘Very Often’. Single-item questions can provide effective global indicators of how much positive and negative contact a respondent experiences (Lolliot et al. 2015).

### *Matching and Modelling*

The first step of the PSM-DiD approach is to generate a matched sample of participants/controls using propensity-score matching. We match on an extensive range of pre-participation measures (see [Appendix 1](#) for full descriptives), including: multiple socio-demographic indicators; an index of social confidence;<sup>5</sup> an index of network-diversity e.g. support ties across different schools, races, religions, sexualities, incomes;<sup>6</sup> prior engagement behaviours; and region of residence. We also match on pre-participation positive/negative contact, and feeling thermometer scores, to strengthen claims of strong-ignorability. An iterative process was undertaken to achieve best balance across groups, using multiple matching methods, including restricting to regions of common support, trimming (5% level), and caliper/bandwidth specifications, with tests of matching quality employed to examine balance across bias and variance.<sup>7</sup> Based on the diagnostics, we take an Epanechnikov kernel-density matching approach with a bandwidth of 0.06 (see Supplementary-Appendix A.1 for a comparison across strategies, A.2 for pre-/post-matched sample characteristics and A.3-A.4 for post-matching diagnostics). Kernel-density matching compares each participant with all available control observations, weighting observations according to their distance (propensity score) from participant cases.

The second step of the PSM-DiD approach employs regression modelling to estimate the DiD-scores, incorporating the PSM kernel-density weights into the models. Regression modelling allows us to account for the nested nature of our data (observations in individuals), using multi-level mixed-effects linear regression models. Bootstrapped standard errors are used (1000 reps). The DiD-estimator is specified using an interaction term between the pre-test/post-test identifier and the participant/control-group identifier. To explore whether participation exerts different effects conditional on individuals’ pre-

participation levels of positive/negative contact we employ a Difference-in-Difference-in-Difference (DiDiD) approach. This is specified with a three-way interaction term between the pre-test/post-test identifier, the participant/control-group identifier and pre-participation level of either positive or negative contact.<sup>8</sup> Analyses presented here are conducted on the unweighted sample, with listwise deletion of missing cases (analytic sample of  $n = 1379$  participants/ $n = 1910$  controls). To explore the role of bias from non-response and within-case missingness we employ a range of weighting and multiple-imputation approaches. These demonstrate highly consistent findings with those reported here (see Supplementary-Appendix B).

## Results

### Overall impact of participation

The first stage is to test the overall impact of participation on young people's interethnic attitudes. Model 1 (Table 1) demonstrates that the DiD-score (the impact of participation on intergroup attitudes) is significant and positive: DiD: 2.58 [CI: 0.69, 4.46]. We therefore observe that participants report an increase in warmth towards outgroups of 2.58 points, evident at least 4–6 months after participation has ended. We next examine how far this impact is driven by any changes in positive or negative outgroup mixing. Models 2 and 3 test the overall impact of participation on young people's frequency of positive contact and negative contact. We observe that participants experience a significance increase in positive contact after participation (DiD: 0.13 [CI: 0.06, 0.21]) (Model 2), while participation has a small negative but non-significant impact on negative contact (DiD:  $-0.03$  [CI:  $-0.11, 0.05$ ]) (Model 3). Model 4 then tests whether the observed increases in positive contact can account for the positive impact of participation on outgroup warmth, re-running the overall impact of participation on interethnic attitudes (in Model 1) but now including positive and negative contact. Firstly, we observe that positive contact has a strong positive impact on intergroup-attitudes while negative contact has a negative

**Table 1.** Overall impact of participation on outgroup warmth and positive/negative contact.

Dependent variable	Model 1 Feeling thermometer	Model 2 Positive contact	Model 3 Negative contact	Model 4 Feeling thermometer
<i>Baseline: Pre-test period</i>				
Post-test period	2.746*** (0.671)	-0.047 (0.027)	0.007 (0.030)	3.068*** (0.669)
<i>Baseline: Control-group</i>				
Participants	0.797 (0.827)	0.004 (0.030)	0.033 (0.031)	0.762 (0.803)
Post-test * Participants (DiD)	2.576** (0.962)	0.133*** (0.037)	-0.026 (0.041)	1.732 (0.962)
Positive contact				6.107*** (0.392)
Negative contact				-2.439*** (0.379)
Constant	73.204*** (0.576)	4.151*** (0.021)	2.240*** (0.022)	53.402*** (2.107)
N (individuals)	3289	3289	3289	3289

Notes: kernel-density (Epanechnikov) propensity-score weighted; bootstrapped standard errors in parentheses (1000 reps); \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$  (two-tailed tests).

impact on intergroup attitudes. Secondly, on accounting for changes in contact, the positive impact of participation (DiD-score) is reduced by 34% and is no longer significant.

Engagement therefore leads to warmer attitudes towards ethnic outgroups. Over a third of this positive impact appears to come from increases in positive contact, and, after adjusting for this, participation no longer has a significant relationship with outgroup warmth. Interestingly, despite likely having more opportunities for negative contact during participation we observe no change in participants' levels of negative contact. While these results are promising, the actual size of the impact of youth engagement on intergroup attitudes is relatively small. However, simply looking at the average impact of participation (across all participants) may mask important heterogeneity in effects between different individuals.

### ***Heterogeneous impacts of participation by pre-participation levels of positive and negative contact***

As discussed, the impact of participation on interethnic attitudes could depend on how much positive/negative contact young people had *prior* to participating. To explore this, the first step is to test whether the amount of pre-participation positive or negative contact an individual had conditions (by which we mean, moderates) how participation impacts their outgroup warmth. We employ a DiDiD-approach, modelled using an interaction term between the participant/control-group identifier, pre-test/post-test identifier, and an individual's level of pre-participation positive contact or negative contact. Looking first at the pre-participation positive contact coefficient (Model 1, Table 2), we first observe that its coefficient is strong and positive i.e. young people joining the scheme with more frequent positive mixing report higher outgroup warmth. Secondly, we see that the DiDiD-term for pre-participation positive contact is negative and significant:  $-2.93$  [CI:  $-5.49, -0.37$ ]. Therefore, the *less* positive contact a participant had prior to engaging the *stronger* the positive impact of participation on their interethnic attitudes.<sup>9</sup> To explore this finding in more detail Figure 1 shows the impact of participation (the DiD-score) on outgroup warmth among all participants, and then among participants who, prior to participation, had positive contact 'very often', 'sometimes', and 'never' (derived from Model 1, Table 2). Participants who join with high-levels of positive contact ('very often') experience no impact of engaging on their interethnic attitudes. However, participants who joined reporting they 'sometimes', or especially 'never', have positive contact with outgroups, see a substantially stronger positive impact.

Model 2 tests whether participation's impact on outgroup warmth depends on how much pre-participation negative contact an individual had prior to joining. The coefficient for pre-participation negative contact is strong and negative i.e. young people who, pre-participation, had more frequent negative contact in their daily lives report colder outgroup attitudes. However, while the DiDiD-term for pre-participation negative contact is positive it is also non-significant:  $1.36$  [CI:  $-1.06, 3.78$ ]. The impact of participation on young people's outgroup warmth does not appear to depend on how much negative contact they brought with them to the scheme.

These findings demonstrate that behind the overall impact of participation on outgroup warmth is important heterogeneity based on how much pre-participation positive (but not negative) contact a young person had prior to engaging.<sup>10</sup> We next want to explore what

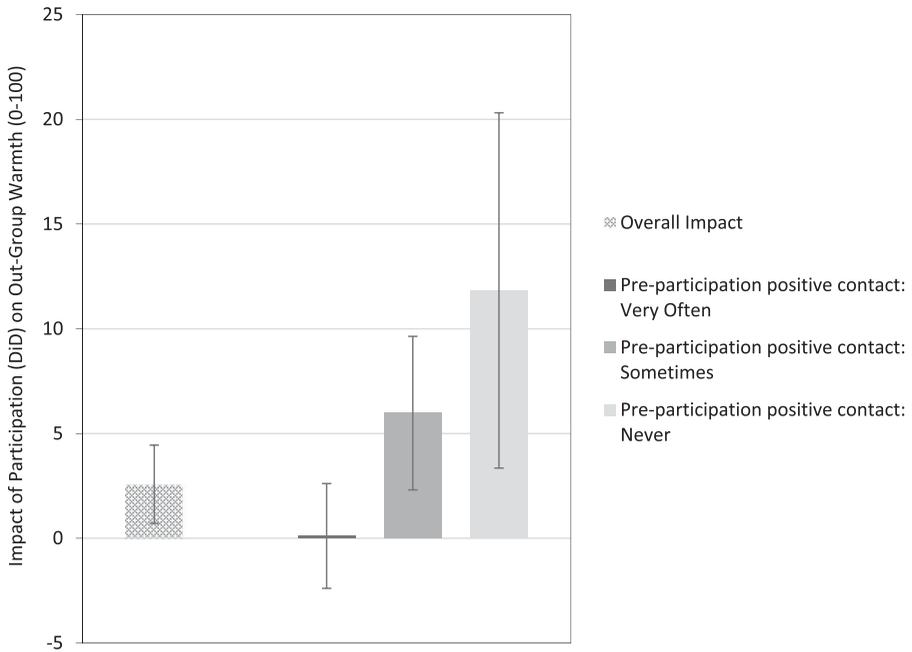
**Table 2.** Heterogeneous impacts of participation on outgroup warmth and positive/negative contact by level of pre-participation positive/negative contact.

Dependent variable	Model 1 Feeling thermometer	Model 2 Feeling thermometer	Model 3 Positive contact	Model 4 Negative contact	Model 5 Feeling thermometer	Model 6 Feeling thermometer
<i>Baseline: Pre-test period</i>						
Post-test period	2.395 (3.738)	2.704 (1.991)	2.506*** (0.195)	1.681*** (0.171)	-8.523* (3.810)	-6.123 (5.033)
<i>Baseline: Control-group</i>						
Baseline: Pre-test period	-0.392 (4.508)	-0.555 (2.453)	-0.002 (0.012)	0.006 (0.011)	-0.348 (4.501)	-2.763 (5.859)
Post-test * Participants (DiD)	14.765** (5.609)	-0.523 (2.837)	0.685* (0.277)	0.092 (0.253)	10.266 (5.548)	9.575 (7.078)
Positive contact					5.519*** (0.563)	4.954*** (0.591)
Negative contact						-2.329*** (0.564)
Pre-participation positive contact	7.267*** (0.732)		0.999*** (0.002)	-0.001 (0.002)	1.764 (0.930)	1.469 (0.973)
Pre-participation positive contact * Participants	0.066 (0.875)		-0.533*** (0.035)	-0.079* (0.033)	2.749** (0.889)	2.526** (0.922)
Post-test Pre-participation positive contact * Participants	0.254 (1.061)		0.001 (0.002)	0.001 (0.002)	0.238 (1.059)	0.481 (1.107)
Pre-participation positive contact * Participants (DiDiD)	-2.931* (1.307)		-0.154** (0.051)	-0.009 (0.048)	-2.015 (1.289)	-2.001 (1.340)
Pre-participation negative contact		-4.641*** (0.766)	0.000 (0.002)	1.002*** (0.002)		-0.209 (1.004)
Pre-participation negative contact * Participants		0.012 (0.894)	-0.145*** (0.035)	-0.590*** (0.035)		-0.649 (1.021)
Post-test Pre-participation negative contact * Participants		0.674 (1.059)	-0.000 (0.002)	-0.002 (0.002)		0.665 (1.087)
Pre-participation negative contact * Participants (DiDiD)		1.357 (1.236)	0.034 (0.047)	-0.036 (0.048)		0.286 (1.296)
Constant	43.137*** (3.106)	83.575*** (1.709)	0.002 (0.012)	-0.006 (0.011)	43.093*** (3.095)	52.332*** (4.040)
N (individuals)	3289	3289	3289	3289	3289	3289

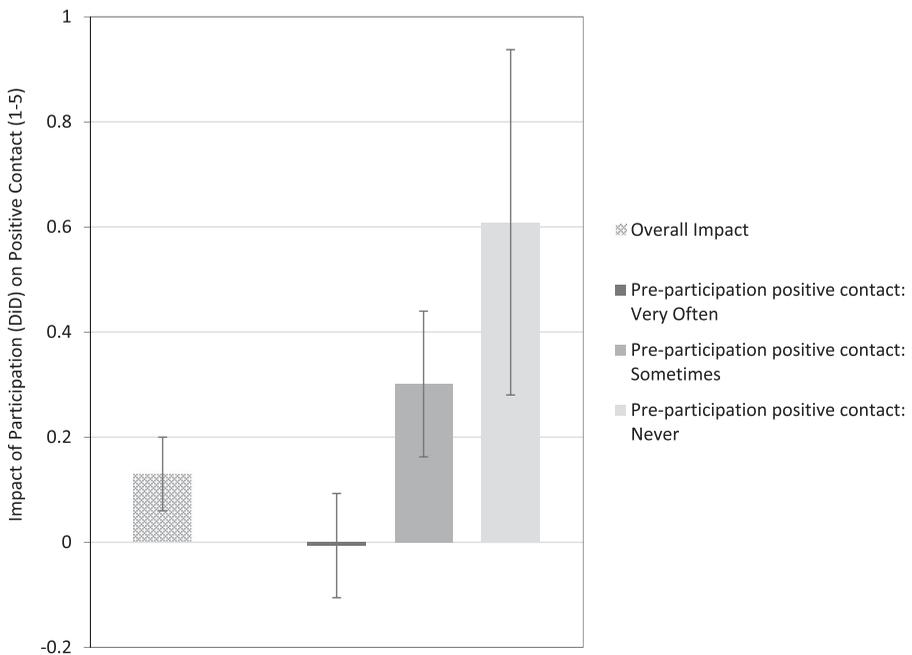
Notes: kernel-density (Epanechnikov) propensity-score weighted; bootstrapped standard errors in parentheses (1000 reps); \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$  (two-tailed tests).

drives this heterogeneity. One possibility is that the impact of participation on an individual's frequency of positive/negative contact may also depend on how much, and what type, of contact they had prior to taking part. Model 3 tests whether the impact of participation on young people's positive outgroup contact is also dependent on how much positive contact they had, or how much negative contact they had, prior to taking part. Again, a DiDiD-approach is applied, using interaction terms between the participant/control identifier, pre-test/post-test identifier, and one's level of pre-participation positive/negative contact.

We observe that the DiDiD-term for pre-participation negative contact is weakly positive and non-significant: DiDiD: 0.034 [CI: -0.05, 0.13]: participation therefore does not



**Figure 1.** Impact of participation on outgroup warmth: Overall and by levels of pre-participation positive contact.



**Figure 2.** Impact of participation on positive contact: Overall and by levels of pre-participation positive contact.

have a different effect on how much positive contact one gains from engagement based on how much negative contact they had prior to joining. However, the DiDiD-term for pre-participation positive contact is significant and negative:  $-0.15$  [CI:  $-0.25, -0.05$ ]. Participating therefore has a *stronger* positive impact on a young person's level of positive contact if they joined with *less* frequent positive contact.<sup>11</sup> Figure 2 shows the size of the impact of participation (DiD-score) on positive contact among: all participants; and then subdivided by whether an individual reported having positive contact 'very often', 'sometimes', and 'never' prior to participation. We observe that behind the overall impact of participation on positive contact lies important heterogeneity: among those who joined already reporting high levels of positive contact, participation leads to no change in their levels of positive contact (as might be expected); however, among individuals who joined reporting they only 'sometimes', and especially 'never', have positive outgroup contact we see a much bigger positive change from engaging.<sup>12</sup>

Model 4 then tests whether the impact of participation on rates of negative outgroup-contact is dependent on pre-participation levels of positive/negative contact. However, both DiDiD-terms are weak and non-significant. The impact of participation on a young person's level of negative contact does not appear to depend on the contact (either positive or negative) they had prior to joining.

Taken together, these findings show that young people who had less frequent positive contact prior to participating experience larger increases in both their outgroup warmth and their level of positive contact. We next explore whether their additional gains in outgroup warmth can be accounted for by their additional gains in positive contact. To test this, Model 5 replicates the analysis of Model 1 (Table 2). However, we now include the pre-test/post-test measure of positive contact. We observe that the DiDiD-term is reduced by 34% and is now no longer significant. In other words, a large part of why participation leads to bigger improvements in warmth among those with lower pre-participation positive contact can be accounted for by their larger gains in positive contact. Model 6 then enters the negative contact relationships into the model; however, the results do not change.

For robustness, we tested whether the observed heterogeneity in participation-effects among those with more/less pre-participation positive contact is driven by heterogeneous effects across other pre-participation characteristics e.g. participation could have a stronger effect on those with lower social efficacy, which is what is driving the observed contact-heterogeneity (see Appendix 1 for full list of pre-participation characteristics tested). However, the substantive findings remain unchanged. Of particular interest are tests for whether participation exhibited different effects for white participants and non-white participants.<sup>13</sup> Although some small differences appear across groups none were significant. Future research with a larger sample of groups is required to test this in more detail.

## Discussion

This study aimed to explore the role youth engagement can play in improving adolescents' intergroup cohesion. We demonstrate that discrete periods of participation can lead to improvements in young people's warmth towards other ethnic groups, can increase their rates of positive inter-ethnic contact, but has no effect on their rates of negative

inter-ethnic contact. Critically, a substantial portion (and the statistical significance) of the positive relationship between participation and outgroup warmth is mediated by the increases in positive intergroup contact. However, the impact of engagement on intergroup cohesion also depends, to a large extent, on how much positive contact participants had *prior* to taking part. Young people who joined the programme with less frequent positive contact saw substantially larger improvements in both their outgroup-warmth and their levels of positive contact; these additional gains in positive contact appear to account for a large portion (and the statistical significance) of why they also experienced larger improvements in outgroup warmth.

Several key insights emerge from these findings. Firstly, these results demonstrate robust evidence that youth engagement can act a key driver of youth intergroup cohesion. Furthermore, the observed positive impacts are evident at least 4–6 months after participation ended, suggesting that young people are able to take the gains made during engagement with them after leaving (at least in the short- to medium-term). The quasi-experimental approach applied increases our confidence these impacts are likely causal.

Secondly, the results suggest a key reason youth engagement appears to work is through the opportunities it provides for positive intergroup contact, and that, despite the likelihood that participating could also increase opportunities for negative contact there is little evidence it does so here. However, positive contact only accounts for around a third of the impact of participation. This may be a consequence of using a single measure of positive contact, and further measures (e.g. number of outgroup friends, more detailed indicators of contact-quality) may account for more. However, participation could also have a positive effect on interethnic attitudes outside of contact pathways. For example, instilling a stronger civic identity may cultivate a more inclusive superordinate identity (Houlette et al. 2004). Participation can increase social efficacy which, in turn, may reduce social anxiety towards intergroup encounters (Plant and Devine 2003; Mellor et al. 2008). Or, by generating a greater sense of empowerment, participation may reduce the kinds of feelings of alienation linked to outgroup hostility (Watkins, Larson, and Sullivan 2007). Therefore, while positive contact constitutes a key channel through which participation leads to improved attitudes, other pathways may be operating alongside this.

Thirdly, these findings also highlight the importance of looking behind overall effects when judging the efficacy of engagement for fostering cohesion. When examining the overall impact of engagement, the scheme appears only marginally effective. However, this is partly a consequence of many participants already evincing high degrees of intergroup cohesion (particularly high levels of positive-contact), who, as we observe, see only small improvements from participating.<sup>14</sup> Among those who join with less positive contact we observe much stronger positive impacts on their warmth towards outgroups. In other words, those who could benefit the most from involvement see the strongest effects.

Notwithstanding the new insights gleaned, this study has limitations. Studying NCS provides a key insight into how a nationally-implemented scheme, open to all age-eligible young people, can affect intergroup cohesion across the country. However, differences in the impact of engagement may exist based on how ethnically diverse one's co-participants are during involvement in the scheme. As mentioned, participants are likely to have unequal opportunities to mix with outgroups during engagement depending on where they participated in the country; for example, in areas where ethnic diversity is low

opportunities for mixing are likely lower (National Audit Office 2017). Accordingly, participation could exhibit weaker/stronger effects on outgroup warmth among those who had fewer/greater opportunities for positive contact during engagement. In this way, our findings provide an overall test of how participation affects intergroup cohesion across all teams in the country. Future research which is able to link in team ethnic diversity data will help explore how far the diversity of one's co-participants may lead to better/worse outcomes across different teams.

A second limitation is potential threats to the internal-validity of the findings, given random-assignment into participation was not possible. As outlined, a key assumption for valid inference using a non-equivalent control-group is that, absent of participation, both participants and controls would have exhibited similar pre-test/post-test trends in the outcomes. However, group-differences between participants and controls could mean that participants already had increasing rates of positive contact/outgroup warmth *prior* to engaging, which is driving the apparent impact of engagement. As outlined, drawing the control-group from a pool of young people who had proactively engaged in NCS recruitment, and performing extensive matching, should strengthen the parallel-trends assumption. While we cannot further investigate this here, the heterogeneity in impacts we observe provides some support for a participation-effect: the size of the increase in warmth (12-points) among those participants who joined with low levels of positive contact, occurring over a 4–6-month period, is unlikely to reflect a secular trend for this group.

A third limitation to the study is threats to the external validity of the findings from missingness, and our ability to generalise from the study to the scheme as a whole. At least on key socio-demographics, the analytic sample is highly similar to the composition of the participants who took part during the evaluation window (see Supplementary-appendix B.1). In addition, analyses applying weights and multiple-imputation returned highly consistent findings, although such techniques are only as effective as the covariates available to weight on/impute with. One particular concern is if participants who did not experience a positive impact of participation were also less likely to complete the follow-up survey; for example, if they had a negative experience on the scheme. If this is the case, the positive findings could be upwardly biased. Unfortunately, we cannot unpack such bias from the current analysis; future research which minimises non-response will be crucial to validating the results reported here.

Lastly, a key question which emerges from such a study is whether the results could be generalisable to all young people. If participation was something every young person undertook would we observe the same results. As in any research into the effects of participation some form of selection is nominally at work which may bias our understanding of its effects. For example, more extroverted, prosocial young people may be more likely to select on to the scheme, who are also more interested in meeting different people, predisposing them to react more positively to the experience. Concurrently, individuals who are more averse to meeting new people, or mixing with other groups, for whom participation could have a weaker impact, may simply avoid taking part. Caution is thus required in generalising to all young people.

## Conclusion

This study provides compelling evidence that youth engagement, in particular discrete periods of participation through youth engagement schemes, can build intergroup

cohesion. It also demonstrates that engagement appears particularly effective at boosting interethnic cohesion among young people who normally have less positive contact in their daily lives. Engagement could therefore be an effective tool to help build cohesion among young people; especially where it is weakest. To be sure, how effective such schemes can be for building intergroup cohesion across society as a whole depends on who takes part. However, nationally-implemented engagement schemes like the one studied here appear to be effective at recruiting groups who normally evince weaker patterns of engagement; for example, young people from disadvantaged backgrounds (Gilman 2001; National Audit Office 2017). In this way, through providing space for all those who want to engage, regardless of who they are or where they live, and the provision of subsidised costs for those who normally might struggle to engage, nationally-implemented engagements schemes could be particularly effective at reaching young people who have fewer opportunities to participate, helping further augment the positive impact participation can have on society.

## Notes

1. Only 15-year olds who turn 16 by August 31 of the participation year are eligible.
2. A smaller cohort of participants attend NCS in Autumn and Spring.
3. See endnote 1
4. This may be because they chose never to participate or participated later in year/the following year.
5. Based on a four-item index on how confident respondents feel about different areas of their life (5-point likert scale 'not at all' to 'very confident'). 'How do you feel about the following things, even if you have never done them before, " ... Meeting new people"; " ... Having a go at things that are new to me"; " ... Working with other people in a team"; " ... Being the leader of a team"' (Eigen Factor: 2.16; minimum loading: 0.69; alpha coefficient: 0.83).
6. Questions on bridging ties load on to a single index of tie-diversity (Eigen value: 2.13; minimum loading: 0.47; alpha coefficient: 0.78).
7. All matching processes were undertaken using the `psmatch2` function in Stata (Leuven and Sianesi 2018).
8. We also explored splitting the samples based on pre-participation contact (e.g. low/high pre-participation positive contact) and matching within these. Results were highly similar.
9. Similar substantive findings are returned if we include pre-participation positive contact as dummy variables.
10. Another possibility is that it is not prior levels of positive contact *per se*, but how much out-group contact a young person had in general before participating; potentially, those with higher levels of out-group contact in general may simply have more fixed attitudes towards out-groups (limiting the impact of participation). Creating a measure of total pre-participation out-group contact (frequency of positive plus negative contact) showed that this did not moderate the impact of participation in the same way. In other words, it is specifically prior levels of positive contact which matter for how participation impacts out-group attitudes.
11. Similar substantive findings are returned if we include pre-participation positive contact as dummy variables.
12. We also tested whether the effect of participation on positive/negative depended on participants' pre-participation level of outgroup warmth e.g., were participants who had colder attitudes towards out-groups less likely to experience positive contact? We observe no significant relationships.
13. The data does not allow one to disaggregate white British from the white category.
14. This may also be a consequence of the way our measures are constructed. For example, if an individual joins the scheme already reporting positive contact 'very often' then even if their

positive contact does increase further we cannot capture this as respondents have already selected the maximum amount of contact.

## Acknowledgements

This work received generous support from Nerissa Steel, Charlotte Kume-Holland, and Anna Romaniuk at the UK Cabinet Office; Maria Willoughby at the UK Department for Digital, Culture, Media and Sport; Steven Webster, Paul Adnitt, Will Damazer, and Michael Lynas at the National Citizen Service Trust; and Daniel Cameron and James Stannard at Ipsos MORI.

## Disclosure statement

Access to the data was provided by the National Citizen Service Trust (which manages the youth engagement scheme studied). The author's institution also received a small funding grant from the National Citizen Service Trust to allow the author to write-up a report based on the independent research into the impacts of participating on the youth engagement scheme.

## Funding

This work was supported with funding from the Economic and Social Research Council [grant number ES/S013121/1] and the Alexander von Humboldt Stiftung/Foundation [Ref: GBR 1195617 HFST-E].

## References

- Allport, Gordon W. 1954. *The Nature of Prejudice*. Reading, MA: Addison-Wesley.
- Al Ramiah, Ananthi, and Miles Hewstone. 2012. "Rallying Around the Flag: Can an Intergroup Contact Intervention Promote National Unity?" *British Journal of Social Psychology* 51: 239–256.
- Amir, Yehuda. 1969. "Contact Hypothesis in Ethnic Relations." *Psychological Bulletin* 71: 319–342.
- ARK. 2009. *Young Life and Times Survey 2009: Technical Notes*. Belfast: Department for Social Development in Northern Ireland.
- Barlow, Fiona Kate, Stefania Paolini, Anne Pedersen, Matthew J. Hornsey, Helena R. M. Radke, Jake Harwood, Mark Rubin, and Chris G. Sibley. 2012. "The Contact Caveat: Negative Contact Predicts Increased Prejudice More Than Positive Contact Predicts Reduced Prejudice." *Personality and Social Psychology Bulletin* 38: 1629–1643.
- Bayram Özdemir, Sevgi, Shuyan Sun, Liliia Korol, Metin Özdemir, and Håkan Stattin. 2018. "Adolescents' Engagement in Ethnic Harassment: Prejudiced Beliefs in Social Networks and Classroom Ethnic Diversity." *Journal of Youth and Adolescence* 47: 1151–1163.
- Bekhuis, Hidde, Stijn Ruiter, and Marcel Coenders. 2013. "Xenophobia among Youngsters: The Effect of Inter-Ethnic Contact." *European Sociological Review* 29: 229–242.
- Binder, Jens, Hanna Zagefka, Rupert Brown, Friedrich Funke, Thomas Kessler, Amelie Mummendey, Annemie Maquil, Stephanie Demoulin, and Jacques-Philippe Leyens. 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: 843–856.
- Blair, Irene V., Charles M. Judd, Edward P. Havranek, and John F. Steiner. 2010. "Using Community Data to Test the Discriminant Validity of Ethnic/Racial Group IATs." *Zeitschrift für Psychologie / Journal of Psychology* 218: 36–43.
- Blau, P. M. 1977. *Inequality and Heterogeneity - A Primitive Theory of Social Structure*. New York: Free Press.

- Brown, Kendrick T., Tony N. Brown, James S. Jackson, Robert M. Sellers, and Warde J. Manuel. 2003. "Teammates On and Off the Field? Contact With Black Teammates and the Racial Attitudes of White Student Athletes." *Journal of Applied Social Psychology* 33: 1379–1403.
- Brown, R., T. Kessler, J. Leyens, and A. Mummendey. 2008. *Longitudinal Three-Nation Study on Relations Between Ethnic Minorities and Host Societies among School Students, 2004–2005*. [Data Collection]. UK Data Service. SN: 5676. doi:10.5255/UKDASN- 5676-1.
- Bundick, Matthew J. 2011. "Extracurricular Activities, Positive Youth Development, and the Role of Meaningfulness of Engagement." *The Journal of Positive Psychology* 6: 57–74.
- Catney, Gemma. 2013. *Has Neighbourhood Ethnic Segregation Decreased?* Manchester: Centre on Dynamics of Ethnicity: University of Manchester.
- Conner, Jerusha, and Joseph Erickson. 2017. "When Does Service-Learning Work? Contact Theory and Service-Learning Courses in Higher Education." *Michigan Journal of Community Service Learning* 23: 53–65.
- Daw, Jamie R., and Laura A. Hatfield. 2018. "Matching and Regression to the Mean in Difference-in-Differences Analysis." *Health Services Research* 53: 4138–4156.
- Dixon, John, Kevin Durrheim, and Colin Tredoux. 2005. "Beyond the Optimal Contact Strategy: A Reality Check for the Contact Hypothesis." *American Psychologist* 60: 697–711.
- Dworkin, Jodi B., Reed Larson, and David Hansen. 2003. "Adolescents' Accounts of Growth Experiences in Youth Activities." *Journal of Youth and Adolescence* 32: 17–26.
- Erickson, J., and S. O'Connor. 2000. "Service-learning's Effect on Prejudice: Does it Reduce or Promote it." In *Transforming Education, Transforming the World: The Integration of Service-Learning and Multicultural Education Into Higher Education*, edited by C. O'Grady, 59–70. Erlbaum Mahwah, NJ: Lawrence Erlbaum Associates.
- Europe Commission. 2016. *Action Plan on the Integration of Third Country Nationals*. Brussels: European Commission.
- Ford, Robert, and Matthew Goodwin. 2017. "Britain After Brexit: A Nation Divided." *Journal of Democracy* 28: 17–30.
- Frumkin, Peter, JoAnn Jastrzab, Margaret Vaaler, Adam Greeney, Robert T. Grimm, Kevin Cramer, and Nathan Dietz. 2009. "Inside National Service: AmeriCorps' Impact on Participants." *Journal of Policy Analysis and Management* 28: 394–416.
- Gaertner, Samuel L., John F. Dovidio, Rita Guerra, Margarida Rebelo, Maria Benedicta Monteiro, Blake M. Riek, and Melissa A. Houlette. 2008. "The Common in-Group Identity Model: Applications to Children and Adults." In *Intergroup Attitudes and Relations in Childhood Through Adulthood*, edited by S. R. Levy and M. Killen, 204–219. New York, NY: Oxford University Press.
- Gilman, Rich. 2001. "The Relationship Between Life Satisfaction, Social Interest, and Frequency of Extracurricular Activities among Adolescent Students." *Journal of Youth and Adolescence* 30: 749–767.
- Gireesh, Aswathikutty, Shikta Das, and Russell M. Viner. 2018. "Impact of Health Behaviours and Deprivation on Well-Being in a National Sample of English Young People." *BMJ Paediatrics Open* 2: e000335–e35.
- Guffler, Kerstin, and Ulrich Wagner. 2017. "Backfire of Good Intentions: Unexpected Long-term Contact Intervention Effects in an Intractable Conflict Area." *Peace and Conflict: Journal of Peace Psychology* 23: 383–391.
- Haines, Nicola. 2017. *Birth Characteristics in England and Wales: 2016*. London: Office for National Statistics.
- Hamm, Jill V., B. Bradford Brown, and Daniel J. Heck. 2005. "Bridging the Ethnic Divide: Student and School Characteristics in African American, Asian-Descent, Latino, and White Adolescents' Cross-Ethnic Friend Nominations." *Journal of Research on Adolescence* 15: 21–46.
- Hewstone, M., and B. Brown. 1986. "Contact is not Enough: an Intergroup Perspective on the 'Contact Hypothesis'." In *Contact and Conflict in Intergroup Encounters*, edited by M. Hewstone and R. Brown, 1–44. New York: Basil Blackwell.
- Holsapple, Matthew A. 2012. "Service-Learning and Student Diversity Outcomes: Existing Evidence and Directions for Future Research." *Michigan Journal of Community Service Learning* 18: 5–18.

- Houlette, Melissa A., Samuel L. Gaertner, Kelly M. Johnson, Brenda S. Banker, Blake M. Riek, and John F. Dovidio. 2004. "Developing a More Inclusive Social Identity: An Elementary School Intervention." *Journal of Social Issues* 60: 35–55.
- Jessiman, Patricia, and Emma Drever. 2010. *The Youth of Today Evaluation: Technical Report*. London: Department for Education.
- Keating, A., D. Kerr, T. Benton, E. Mundy, and J. Lopes. 2010. *Citizenship Education in England 2001-2010: Young People's Practices and Prospects for the Future: the Eighth and Final Report From the Citizenship Education Longitudinal Study (CELS)*. London: Department for Education.
- Knifsend, Casey A., and Jaana Juvonen. 2017. "Extracurricular Activities in Multiethnic Middle Schools: Ideal Context for Positive Intergroup Attitudes?" *Journal of Research on Adolescence: The Official Journal of the Society for Research on Adolescence* 27: 407–422.
- Koopmans, Ruud, and Susanne Veit. 2014. "Ethnic Diversity, Trust, and the Mediating Role of Positive and Negative Interethnic Contact: A Priming Experiment." *Social Science Research* 47: 91–107.
- Larson, R. W., and S. Verma. 1999. "How Children and Adolescents Spend Time Across the World: Work, Play, and Developmental Opportunities." *Psychological Bulletin* 125: 701–736.
- Laurence, James. 2019a. "Community Disadvantage, Inequalities in Adolescent Subjective Well-being, and Local Social Relations: The Role of Positive and Negative Social Interactions." *Social Science & Medicine* 237: 112442.
- Laurence, James. 2019b. "Youth Engagement, Positive Interethnic Contact, and 'Associational Bridges': A Quasi-Experimental Investigation of a UK National Youth Engagement Scheme." *Journal of Youth and Adolescence* 48: 1264–1280.
- Laurence, James, Katharina Schmid, and Miles Hewstone. 2017. "Ethnic Diversity, Inter-group Attitudes and Countervailing Pathways of Positive and Negative Inter-group Contact: An Analysis Across Workplaces and Neighbourhoods." *Social Indicators Research* 136: 719–749.
- Laurence, James, Katharina Schmid, James R. Rae, and Miles Hewstone. 2018. "Prejudice, Contact, and Threat at the Diversity-Segregation Nexus: A Cross-Sectional and Longitudinal Analysis of How Ethnic Out-Group Size and Segregation Interrelate for Inter-Group Relations." *Social Forces* 97: 1029–1066.
- Lee, Joanna M., Lauren J. Germain, Edith C. Lawrence, and Jenna H. Marshall. 2010. "'It Opened my Mind, My Eyes. It was Good.'" Supporting College Students' Navigation of Difference in a Youth Mentoring Program." *Educational Horizons* 89: 33–46.
- Leuven, E., and B. Sianesi. 2018. *Psmatch2: Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing* [Computer Software]. Boston, MA: Boston College.
- Levy, Sheri R, and Julie Milligan Hughes. 2009. "Development of Racial and Ethnic Prejudice among Children." In *Handbook of Prejudice, Stereotyping, and Discrimination*, edited by T. D. Nelson, 23–42. New York: Psychology Press.
- Lolliot, Simon, Benjamin Fell, Katharina Schmid, Ralf Wölfer, Hermann Swart, Alberto Voci, Oliver Christ, Rachel New, and Miles Hewstone. 2015. "Measures of Intergroup Contact." In *Personality and Social Psychological Constructs*, edited by Gregory J. Boyle, Donald H. Saklofske, and Gerald Matthews, 652–683. London: Academic Press.
- Maoz, Ifat. 2002. "Is There Contact at all? Intergroup Interaction in Planned Contact Interventions Between Jews and Arabs in Israel." *International Journal of Intercultural Relations* 26: 185–197.
- McAdam, Doug. 1986. "Recruitment to High-Risk Activism: The Case of Freedom Summer." *American Journal of Sociology* 92: 64–90.
- McGee, Rob, Sheila Williams, Philippa Howden-Chapman, Jennifer Martin, and Ichiro Kawachi. 2006. "Participation in Clubs and Groups From Childhood to Adolescence and its Effects on Attachment and Self-esteem." *Journal of Adolescence* 29: 1–17.
- Melchior, Alan, Joseph Frees, Lisa LaCava, Chris Kingsley, Jennifer Nahas, Jennifer Power, Gus Baker, et al. 1999. "Summary Report: National Evaluation of Learn and Serve America".
- Mellor, David, Yoko Hayashi, Lucy Firth, Mark Stokes, Sue Chambers, and Robert Cummins. 2008. "Volunteering and Well-Being: Do Self-Esteem, Optimism, and Perceived Control Mediate the Relationship?" *Journal of Social Service Research* 34: 61–70.

- MHCLG. 2018. *Integrated Communities Strategy Green Paper*. London: Ministry of Housing, Communities and Local Government.
- Moody, James. 2001. "Race, School Integration, and Friendship Segregation in America." *American Journal of Sociology* 107: 679–716.
- Morgan, William, and Matthew Streb. 2001. "Building Citizenship: How Student Voice in Service-Learning Develops Civic Values." *Social Science Quarterly* 82: 154–169.
- National Audit Office. 2017. *National Citizen Service*. London: Cabinet Office and Department for Culture, Media & Sport.
- National Citizen Service Trust. 2018. NCS Trust Announces NCS 2.0, Seeking More Partnerships with Local Organisations.
- Pancer, S. Mark. 2015. *The Psychology of Citizenship and Civic Engagement*. New York, NY: Oxford University Press.
- Park, Nansook. 2004. "The Role of Subjective Well-Being in Positive Youth Development." *The ANNALS of the American Academy of Political and Social Science* 591: 25–39.
- Pettigrew, Thomas F. 2008. "Future Directions for Intergroup Contact Theory and Research." *International Journal of Intercultural Relations* 32: 187–199.
- Pettigrew, Thomas F., and Linda R. Tropp. 2006. "A Meta-analytic Test of Intergroup Contact Theory." *Journal of Personality and Social Psychology* 90: 751.
- Plant, E. Ashby, and Patricia G. Devine. 2003. "The Antecedents and Implications of Interracial Anxiety." *Personality and Social Psychology Bulletin* 29: 790–801.
- Putnam, Robert D. 2000. *Bowling Alone: The Collapse and Revival of American Community*. New York: Simon & Schuster.
- Reynolds, Tracey. 2007. "Friendship Networks, Social Capital and Ethnic Identity: Researching the Perspectives of Caribbean Young People in Britain." *Journal of Youth Studies* 10: 383–398.
- Simpson, Ludi. 2012. *More Segregation or More Mixing?* Manchester: Centre on the Dynamics of Ethnicity.
- Smith, Elizabeth S. 1999. "The Effects of Investments in the Social Capital of Youth on Political and Civic Behavior in Young Adulthood: A Longitudinal Analysis." *Political Psychology* 20: 553–580.
- Stephan, Walter G., and Cookie White Stephan. 1985. "Intergroup Anxiety." *Journal of Social Issues* 41: 157–175.
- Storm, Ingrid, Maria Sobolewska, and Robert Ford. 2017. "Is Ethnic Prejudice Declining in Britain? Change in Social Distance Attitudes among Ethnic Majority and Minority Britons." *The British Journal of Sociology* 68: 410–434.
- ten Berge, Jannes Beer, Bram Lancee, and Eva Jaspers. 2017. "Can Interethnic Friends Buffer for the Prejudice Increasing Effect of Negative Interethnic Contact? A Longitudinal Study of Adolescents in the Netherlands." *European Sociological Review* 33: 423–435.
- van der Meer, Tom, and Jochem Tolsma. 2014. "Ethnic Diversity and Its Effects on Social Cohesion." *Annual Review of Sociology* 40: 459–478.
- Van Zomeren, Martijn, Agneta H. Fischer, and Russell Spears. 2007. "Testing the Limits of Tolerance: How Intergroup Anxiety Amplifies Negative and Offensive Responses to Out-Group-Initiated Contact." *Personality and Social Psychology Bulletin* 33: 1686–1699.
- Watkins, Natasha D., Reed W. Larson, and Patrick J. Sullivan. 2007. "Bridging Intergroup Difference in a Community Youth Program." *American Behavioral Scientist* 51: 380–402.

## Appendix 1. Pre-matched and Post-Matched Descriptive Statistics for Participant and Control-group Samples.

		Treated				Control			
		Mean	SD	N	Range	Mean	SD	N	Range
Civic participation in last 3 months	<i>Unmatched</i>	0.68	0.47	1589	0–1	0.63	0.48	2024	0–1
	<i>Matched</i>	0.67	0.47	1379	0–1	0.68	0.47	1910	0–1
Informal helping in last 3 months	<i>Unmatched</i>	0.74	0.44	1569	0–1	0.73	0.45	2015	0–1
	<i>Matched</i>	0.73	0.45	1379	0–1	0.72	0.45	1910	0–1
Not studying for A-levels qualification	<i>Unmatched</i>	0.54	0.5	1589	0–1	0.48	0.5	2034	0–1
	<i>Matched</i>	0.54	0.5	1379	0–1	0.53	0.5	1910	0–1
Index of social confidence	<i>Unmatched</i>	–0.16	0.88	1585	–3.18–1.34	–0.07	0.89	2028	–3.18–1.34
	<i>Matched</i>	–0.17	0.87	1379	–3.18–1.34	–0.14	0.91	1910	–3.18–1.34
<i>Gen. trust cf. Can't be too careful</i>									
Depends	<i>Unmatched</i>	0.53	0.5	1597	0–1	0.51	0.5	2035	0–1
	<i>Matched</i>	0.54	0.5	1379	0–1	0.55	0.5	1910	0–1
Most can be trusted	<i>Unmatched</i>	0.23	0.42	1597	0–1	0.23	0.42	2035	0–1
	<i>Matched</i>	0.23	0.42	1379	0–1	0.23	0.42	1910	0–1
Free school meals eligible (cf. not)	<i>Unmatched</i>	0.21	0.41	1605	0–1	0.16	0.37	2038	0–1
	<i>Matched</i>	0.2	0.4	1379	0–1	0.2	0.4	1910	0–1
Female (cf. male)	<i>Unmatched</i>	0.62	0.45	1602	0–1	0.68	0.42	2039	0–1
	<i>Matched</i>	0.63	0.44	1379	0–1	0.63	0.44	1910	0–1
non-White (cf. white)	<i>Unmatched</i>	0.4	0.49	1608	0–1	0.16	0.37	2040	0–1
	<i>Matched</i>	0.39	0.49	1379	0–1	0.39	0.49	1910	0–1
Frequency of positive contact	<i>Unmatched</i>	4.14	0.8	1593	1–5	4.11	0.85	2035	1–5
	<i>Matched</i>	4.16	0.79	1379	1–5	4.17	0.81	1910	1–5
Frequency of negative contact	<i>Unmatched</i>	2.29	0.84	1586	1–5	2.19	0.85	2028	1–5
	<i>Matched</i>	2.27	0.82	1379	1–5	2.28	0.83	1910	1–5
Outgroup Feeling Thermometer	<i>Unmatched</i>	74	22.18	1585	0–100	72	22.26	2020	0–100
	<i>Matched</i>	74	22.06	1379	0–100	74	22	1910	0–100
Hours civic/informal activities (typical month)	<i>Unmatched</i>	9.86	17.39	1608	0–100	11.75	18.3	2041	0–100
	<i>Matched</i>	9.9	17.01	1379	0–100	10.54	15.66	1910	0–100
Index of Bridging Ties	<i>Unmatched</i>	–0.01	0.83	1541	–2.53–1.39	–0.25	0.85	2020	–2.53–1.39
	<i>Matched</i>	0	0.84	1379	–2.53–1.39	0	0.82	1910	–2.53–1.39
<i>Cf. 15–16 years old</i>									
16.5 years old	<i>Unmatched</i>	0.15	0.35	1580	0–1	0.16	0.37	2019	0–1
	<i>Matched</i>	0.15	0.36	1379	0–1	0.15	0.36	1910	0–1
17 years old	<i>Unmatched</i>	0.09	0.28	1580	0–1	0.08	0.28	2019	0–1
	<i>Matched</i>	0.09	0.29	1379	0–1	0.09	0.28	1910	0–1
17.5 years old	<i>Unmatched</i>	0.29	0.46	1580	0–1	0.29	0.46	2019	0–1
	<i>Matched</i>	0.3	0.46	1379	0–1	0.29	0.45	1910	0–1
18 years old	<i>Unmatched</i>	0.08	0.27	1580	0–1	0.07	0.25	2019	0–1
	<i>Matched</i>	0.08	0.27	1379	0–1	0.07	0.26	1910	0–1
<i>cf. North East/North West/Yorkshire and the Humber</i>									
East Midlands/West Midlands	<i>Unmatched</i>	0.13	0.34	1608	0–1	0.23	0.42	2041	0–1
	<i>Matched</i>	0.13	0.34	1379	0–1	0.15	0.36	1910	0–1
East of England/London/ South East/South	<i>Unmatched</i>	0.48	0.5	1608	0–1	0.47	0.5	2041	0–1
	<i>Matched</i>	0.48	0.5	1379	0–1	0.44	0.5	1910	0–1