Ada Sonnenfeld
Jennifer Doherty
Miriam Berretta
Shannon Shisler
Birte Snilstveit
John Eyers
et al.

Strengthening intergroup social cohesion in fragile situations

February 2021

Systematic Review 46

Social protection

3ie **International Initiative for Impact Evaluation**

About 3ie

The International Initiative for Impact Evaluation (3ie) promotes evidence-informed equitable, inclusive and sustainable development. We support the generation and effective use of high-quality evidence to inform decision-making and improve the lives of people living in poverty in low- and middle-income countries. We provide guidance and support to produce, synthesise and quality assure evidence of what works, for whom, how, why and at what cost.

3ie systematic reviews

3ie systematic reviews appraise and synthesise the available high-quality evidence on the effectiveness of social and economic development interventions in low- and middle-income countries. These reviews follow scientifically recognised review methods, and are peer-reviewed and quality assured according to internationally accepted standards. 3ie is providing leadership in demonstrating rigorous and innovative review methodologies, such as using theory-based approaches suited to inform policy and programming in the dynamic contexts and challenges of low- and middle-income countries.

About this review

Strengthening intergroup social cohesion in fragile situations

Ada Sonnenfeld
International Initiative for Impact Evaluation (3ie)

Jennifer Doherty
3ie

Miriam Berretta
3ie

Shannon Shisler
3ie

Birte Snilstveit
3ie

John Eyers
3ie

Kiara Castaman
Independent consultant

Romanshi Gupta
Independent consultant

Maria Daniela Anda Leon
Independent consultant

Amber Franich
Independent consultant

Cem Yavuz
Independent consultant

Ashton Baafi
Independent consultant

Rahema Obaid
Independent consultant

3ie Systematic Review 46

February 2021

Summary

Background

Nearly a quarter of people in the world live in a fragile context. They lag behind global averages on a range of socio-economic outcomes, in addition to being at a greater risk of violence and its many consequences. UN Sustainable Development Goal 16, on promoting peaceful and inclusive societies, shows the importance of addressing this population's needs. The interventions in this systematic review seek to create sustainable peace by promoting social cohesion within low- and middle-income countries (L&MICs) where populations live in fragile contexts. Specifically, we focus on interventions that aim to address intergroup social cohesion – the relationships between different social groups within society.

Objectives

The main objective of this review is to identify, appraise and synthesise evidence that answers the question: to what extent are programmes that aim to promote intergroup cohesion for sustainable peace in fragile communities effective in achieving their objectives, as compared with similar communities where such support is not provided? Additional objectives include: (1) understanding the extent to which effects vary by geography or population group; and (2) identifying the key mechanisms, and barriers and facilitators that influence how the interventions work, for whom, under what conditions, and at what cost.

Methods

We conducted a systematic and transparent search to identify all potentially relevant published and unpublished studies. We developed a search strategy in collaboration with an information specialist and carried out a systematic search of key academic databases, donor and practitioner websites, and portals of impact evaluations. We also conducted forwards and backwards citation searches for included studies. We imported all search results and managed them using Eppi reviewer. The searches were carried out between January and April 2020. We included studies reported in any language.

Selection criteria: We included impact evaluations of interventions aiming to build intergroup social cohesion that report at least one measure of social cohesion. These measures could include: trust, sense of belonging, willingness to participate, willingness to help, and acceptance of diversity. Since our inclusion criteria focused on outcomes, rather than the more typical focus on intervention type, our search yielded a variety of different interventions, aiming to promote various dimensions of social cohesion. Interventions must have been implemented in situations of fragility in L&MICs, with target populations made up of at least two social groups between whom tensions are present, and where violence is either ongoing, or there is a risk of violence, or there was a recent violent outbreak. We included studies that used an experimental or quasi-experimental study design, with appropriate methods undertaken to address potential confounding and selection bias.

Data collection and analysis: We screened the titles and abstracts produced by the database and grey literature searches against the inclusion criteria, using double

screening, where two reviewers reviewed each title and abstract. A core team member reconciled any disagreements between the coders' decisions. Two reviewers then screened the full texts of the studies included at the title and abstract stage. We further carried out double data extraction. For each study, two reviewers extracted descriptive and quantitative data, and two reviewers independently carried out risk of bias assessments.

We analysed and grouped the interventions according to their common components. We undertook detailed analysis of all outcomes reported, and built a typology to group outcomes based on similar constructs. Where possible, we calculated standardised effect sizes for each outcome reported in the included studies. Where the data were not available, authors were contacted for additional details. A meta-analysis was conducted using a random effects model with a restricted maximum likelihood estimator when there were enough studies reporting similar outcome constructs and intervention types. To address issues caused by dependent effects, we included a single outcome per study in each model. We examined heterogeneity and conducted a sensitivity analysis and a check for reporting bias when possible.

When the studies available in each intervention-outcome pairing did not allow for meta-analysis, we made a narrative synthesis of results, together with standardised effect sizes. To identify trends and understand relationships between possible moderating variables, we coded descriptive and qualitative data and conducted horizontal and vertical syntheses. Framework synthesis was used to identify and investigate possible factors related to the context, programme design and implementation process that influenced how and in what ways the interventions worked. Cost evidence was consolidated and analysed, although minimal reporting of cost information limited what analyses could be conducted.

Main results

The search returned 77,017 records for screening, from which we included 37 impact evaluation papers. These papers represent 24 different studies evaluating 31 unique interventions or intervention arms, which aimed to build intergroup social cohesion in fragile contexts.

Most of the studies were conducted in Sub-Saharan Africa, but we also identified studies carried out in the Middle East and North Africa, Europe and Central Asia, and East Asia and Pacific. No studies aiming to improve intergroup cohesion between conflicting groups were identified in Latin America and the Caribbean. All of the studies included in the review have some methodological weaknesses, but in the majority of cases, these do not induce high risk of bias. Six studies are identified as having a high risk of bias, the majority of which are of non-randomised study designs.

Intervention programming differed substantially. Based on an analysis of the different intervention components, we identified five distinct types of interventions: (1) school-based peace education interventions; (2) collaborative contact interventions; (3) intergroup dialogue interventions; (4) workshop-based peace education with intergroup contact and economic support; and (5) media for peace interventions.

While there are some exceptions, the overall pattern across intervention types and different measures of social cohesion suggests small positive effects. This trend is repeated when the analysis is disaggregated by intervention types and dimensions of social cohesion, although due to the low number of studies included for each analysis, the majority of these estimates remain imprecise and we also observe high heterogeneity across many of the analyses.

We identified four studies of school-based peace education interventions, which aimed to improve social cohesion by triggering mechanisms of 'seeing the other' through peace education and, in three cases, 'talking with the other' and 'working with the other', where intergroup activities were undertaken. Each programme targeted between 1 and 40 different schools, and was typically held over a single academic year, although one lasted three years. The results suggest small positive effects on trust, willingness to participate and willingness to help, and weakly positive effects on acceptance of diversity that cross the line of no effect. Effects on sense of belonging were null or negative. One school-based peace education intervention measured indirect programme effects on parents who did not participate in the activities, and found null or negative effects across all outcomes. Few of the studies reported on intermediate social cohesion outcomes, although two studies found null effects on knowledge of peace and conflict concepts. One study identified negative effects on self-efficacy, which it connected to the negative effects on sense of belonging that had been identified, suggesting that perhaps the intervention had opened students' eyes to the dangers and consequences of discrimination without sufficiently supporting them to feel empowered to address it.

Four further studies evaluated collaborative contact interventions, which included mechanisms of 'working with the other' that brought different groups together to work collaboratively as a team or in groups on shared projects. The interventions lasted from two months to a year, but varied widely in scope, from targeting classes at a single vocational training institute to a policy with national coverage. Our meta-analyses of the effects of these interventions found a small positive effect on willingness to participate, and a very small but imprecise effect on acceptance of diversity. Positive and significant effects were identified on sense of belonging, among the only two studies to report such measures for this intervention group. Effects on trust and willingness to help were typically mixed, although three studies identified at least one measure on willingness to help that was positive and significant, primarily where the measure related directly to the intervention activities. Few studies measured intermediate social cohesion outcomes. With one exception, all reported measures of sociocultural awareness, social and emotional skills suggest no effect.

Six studies measured the effects of intergroup dialogue interventions, which were primarily based on mechanisms of 'talking with the other', although a few incorporated elements of 'seeing the other' through peace education components and 'working with the other' through collaborative projects. These interventions were typically of a short duration, with five programmes ranging between four days and three months, and a sixth lasting seven months. Although two meta-analyses identified effects in a negative direction, these were primarily driven by a study with a high risk of bias in its design and analysis, which reported 13 of the 14 outlier effects identified. With one other exception, however, effects across interventions in this category were mixed or null. The exception was a truth and reconciliation intervention in Sierra Leone, which found positive effects

on trust, willingness to participate and willingness to help. The few studies that reported intermediate social cohesion outcomes found no effects, apart from a positive effect on forgiveness in the truth and reconciliation intervention in Sierra Leone. However, that study also identified a worrisome negative effect of participation in the truth and reconciliation process on measures of depression and post-traumatic stress disorder (PTSD).

Five studies evaluated workshop-based peace education interventions with intergroup contact and economic support (workshops-contact-econ), relying on the mechanisms of 'seeing the other', 'talking with the other' and 'working with the other'. Four of these programmes were of a relatively long duration of four to five years, with the remainder lasting one year. They tended to be larger in scope as well, both in geographic coverage and in intensity of the interventions. The meta-analyses suggest a positive and significant effect on trust, while effects on sense of belonging and willingness to participate were positive but imprecise. Few studies measured effects on willingness to help or acceptance of diversity, but effects were mixed where reported. As with other intervention groups, few studies measured effects on intermediate outcomes. Where reported, however, effects on outcomes like dispute resolution and support for violence and extremism were nil.

Finally, the five studies of media for peace interventions aimed to trigger mechanisms of 'seeing the other' through radio dramas or messaging campaigns. There was a key distinction, however, between the four studies of radio dramas, which focused on triggering social and behavioural changes through 'edutainment', and the fifth study of a messaging campaign promoting counter-narratives to violence. The radio drama interventions were longer lasting, ranging from one to roughly five years, and were implemented as nationally broadcast radio programmes. The strategic communications intervention was a one-month media campaign. Due to this distinction, our analysis focuses primarily on the radio dramas. The meta-analysis of the effect of radio drama interventions on outcomes of trust suggests a small positive effect. We find no effect overall of radio dramas on acceptance of diversity, yet there was substantial heterogeneity in the models, which was driven by a single study. The study that identified primarily null or negative effects was implemented in a context of ongoing conflict; by contrast, studies reporting positive effects were implemented in post-conflict contexts. The radio drama interventions in contexts of latent conflict appeared particularly effective at changing perceptions of norms, including the value of talking about trauma. Across intermediate outcomes, two of the radio drama studies identified positive effects on sociocultural awareness outcomes, while two others identified effects in opposite directions on knowledge of peace and conflict concepts. The study of the media campaign found no effect of the intervention on support for violence and extremism.

We conducted a qualitative synthesis of barriers and facilitators that may help to explain these results. The factors identified through that synthesis are summarised here:
- **Programmes that accurately identified local bottlenecks to intergroup social cohesion tended to have larger and more positive effects.** We identified multiple instances in which the bottlenecks to social cohesion targeted by the interventions appear to have been misaligned either with the context or with the population.

- **A lack of conflict assessments may be a barrier to better targeting of programme participants and key intervention strategies.** While conflict assessments are regularly used to inform the design and targeting of programming in fragile contexts, this was only explicitly mentioned in a couple of contexts. A lack of such analysis, and the associated targeting that follows, represents a possible barrier to larger effects from the programmes included in this review.

- **A lack of substantive changes in intermediate social cohesion outcomes may be a barrier to larger improvements in final social cohesion outcomes.** The effects of intergroup social cohesion interventions on intermediate outcomes, such as sociocultural awareness, social and emotional skills, and dispute resolution practices, are limited, and small at best, with some examples of negative effects. A breakdown in the theorised causal chain, from intervention activities to improved social cohesion outcomes, may explain small effects on these final outcomes. However, as few studies report intermediate outcomes, it is not clear if this 'breakdown' is due to limitations in programme design (theory) or implementation (practice).

- **Intergroup social cohesion interventions may not be sufficient for sustainable social cohesion without structural changes addressing threats to human security outcomes.** Ultimately, when it comes to building sustainable social cohesion and peace, without broader structural changes that address structural drivers of conflict there may be limits to the effects that can be expected from intergroup social cohesion interventions.

- **Smaller-scale interventions may not provide sufficient intensity of treatment to have effects beyond direct participants.** Many of the studies in this review that found nil or negative effects were those evaluating the effects of smaller-scale intergroup social cohesion interventions on indirect participants. The diffusion of effects to those who do not participate directly in any of the intervention activities tends to be smaller than the effects on people who receive a higher 'dosage' of the intervention.

- **Impacts on intergroup relationships among participants may emerge sooner than impacts on wider intergroup relationships, which may need more time to shift.** Smaller-scale interventions that incorporate elements of intergroup contact in particular may be better able to identify effects on changes in relationships among participants from different groups who have interacted with each other through the intervention. Effects on intergroup relationships more broadly may require more intensive interventions and/or time to emerge.

- **Long and non-linear causal chains may be a barrier to substantive improvements in social cohesion.** Social cohesion is a complex outcome, both to shift and to measure. Effects are likely to take time to materialise, be difficult to measure and may also result from non-linear causal chains. Local conditions at baseline may influence the type and scale of possible impacts. Particularly, effects on acceptance of diversity may be highly context-dependent.

Conclusion and implications

While there are some exceptions, the overall pattern across intervention types and different measures of social cohesion suggests small positive effects. This trend is repeated when the analysis is disaggregated by intervention types and dimensions of social cohesion, although, due to the low number of studies included for each analysis, the majority of these estimates remain imprecise and we also observe high heterogeneity across many of the analyses. The results, and in particular the effects of specific interventions on specific outcomes, should be interpreted with some caution as they are often based on relatively few observations and many rely on studies with high risk of bias.

Implications for policy and practice
Based on the findings from both our qualitative and quantitative analysis, we identify the following implications for policy and practice:
- Although the isolated effects are small, it is possible to improve social cohesion outcomes through targeted intergroup interventions – but they are only one piece of the puzzle to build sustainable peace.
- However, intergroup social cohesion interventions are not sufficient on their own to address the underlying drivers of intergroup tensions. For larger improvements in social cohesion outcomes, more complex interventions with complementary strategies to address intergroup social cohesion and structural barriers to inequality and human security outcomes may be needed. However, more theory building work is required to understand how the different drivers and corresponding strategies to address them interact.
- Realistic timeframes are needed to allow substantive changes to social cohesion to materialise, and for the potential for non-linear effects, taking into account the duration and intensity of the intervention. Where higher-level impacts may not be expected, evaluations should focus on identifying and measuring effects on intermediate outcomes.
- If the intervention aims to improve social cohesion outcomes at a population level, and beyond direct participants, more intensive intervention strategies are likely to be required.
- Conducting a rigorous diagnostic conflict assessment during intervention design may facilitate effectiveness, by identifying the relevant local bottlenecks to social cohesion and sustainable peace in the intervention context, including identifying key stakeholders from all groups who may need to be engaged in the intervention to ensure effectiveness. This can ensure intervention activities are aligned with the local context and respond to target groups" needs.
- Theories of change for intergroup social cohesion interventions should receive much more attention. Middle-range theories should be developed that provide a common framework onto which locally relevant indicators can be mapped, including both intermediate and final outcomes, and comprising measures of all five dimensions of social cohesion: trust, sense of belonging, willingness to participate, willingness to help, and acceptance of diversity. The risk of negative effects also highlights the need to identify and mitigate any potential negative unintended effects.
- When designing new programmes or policies, consulting this report and relevant primary studies can help to improve outcomes by informing the design according to the best available evidence.

- Finally, when developing and implementing interventions where there is a limited evidence base, consider whether it would be possible to include an impact evaluation within the new programme.

Implications for research

This review has identified a modest and growing body of literature assessing the effects of programmes aiming to improve intergroup social cohesion in fragile contexts. The existing evidence allows us to identify some preliminary findings, but our ability to identify strong and generalisable findings is somewhat limited by the characteristics of the available evidence. First, while the growth in impact evaluations is encouraging, the field is highly fragmented. This fragmentation spans the types of interventions being evaluated, how authors describe them, the outcome constructs being measured and how they are measured. As a result, the number of observations for any given combination of intervention and outcome constructs is generally small, limiting our ability to compare and synthesise findings across studies to identify generalisable and context-specific findings.

Second, the geographical distribution of studies is uneven, with the majority of studies being conducted in Sub-Saharan Africa, with a handful of studies from the Middle East and North Africa, Europe and Central Asia, and East Asia and Pacific, and no studies from Latin America and the Caribbean identified. Third, few of the existing studies adopt mixed-method designs with causal chain analysis, limiting in particular the extent to which the existing evidence addresses questions related to programme design, implementation and context. Fourth, while the included studies meet a threshold for inclusion based on study design and analysis criteria, a substantial share of studies suffer from significant risks of bias, due to limitations in their study design and analytical approach.

Fifth, there are some important gaps in the way studies are conducted and the questions they address. Notably, over half of the included studies do not include any mention of ethics. For an area where there is also evidence to suggest that interventions can do harm, appropriate procedures for addressing ethics, including through formal review and ethics approval, is essential. Finally, relatively few studies address intermediate outcomes, costs and equity. Analyses of costs are especially important to ensure that study results are relevant to a policy and practice audience. We had intended to conduct cost-effectiveness analysis for this review, but we had to abandon these plans due to a lack of data and transparent reporting for the studies that do provide data on this.

Funders and researchers should consider the following when commissioning and designing new studies:
- Developing a standardised intervention taxonomy to facilitate the use of common terminology to describe the same interventions;
- Adopting a common framework across studies, including both intermediate and impact outcomes to which locally relevant indicators can be mapped, to enhance the value and potential for cross-study lessons and evidence synthesis – this should include measures of all five dimensions of social cohesion: trust, sense of belonging, willingness to participate, willingness to help, and acceptance of diversity;
- Recognising that effects on acceptance of diversity may be highly context-dependent, and ensuring that efforts are made to identify appropriate measures of change for the local context;

- Where participants knowingly self-select into intergroup social cohesion interventions, evaluating effects on the nature of intergroup interactions may be more appropriate than measuring frequency or openness to interactions;
- Ensuring analysis is structured along the causal chain, including identifying and evaluating effects on outcomes earlier in the chain before social cohesion outcomes, particularly for evaluations based on shorter follow-up periods;
- Ensuring new studies include data on costs, based on a clear and transparently reported approach;
- Employing study designs informed by a mixed-methods, theory-based approach to impact evaluations that considers a range of questions relevant to policy and practice, including intervention design, implementation, contextual factors and intermediate outcomes;
- Ensuring research designs and methods are sensitive to inequalities across different population groups – taking into account diverse experiences, power dynamics and gendered inequality in study design and conduct will ensure new studies are sensitive to the needs and effects of programmes with regard to vulnerable groups;
- Adopting best practice for ethical research conduct and protection of research participants, including undertaking and reporting review and approval of study protocols and procedures by relevant review boards to ensure that interventions and research do no harm; and
- Adhering to commonly accepted standards for research transparency and reporting, including pre-registration of all new studies (experimental and quasi-experimental, e.g. in the 3ie Research Transparency Policy 2018).

How up to date is this review?

We undertook the systematic searches of academic databases and grey literature in January 2020, and completed backwards and forwards citation searches of included studies in April 2020.

Contents

List of figures and tables

Abbreviations

AEA	American Economic Association
BMZ	German Federal Ministry for Economic Cooperation and Development
CDD	Community-driven development
DAC	Development Assistance Committee (OECD)
DRC	Democratic Republic of the Congo
EGM	Evidence gap map
FSI	Fragile States Index
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit (German Corporation for International Cooperation – the German development agency)
ITS	Interrupted time series
L&MIC	Low- and middle-income countries
ODA	Official development assistance
PAP	Pre-analysis plan
PBF	Peacebuilding Fund (United Nations)
PTSD	Post-traumatic stress disorder
RCT	Randomised controlled trials
RIDIE	Registry for International Development Impact Evaluations (3ie)
SD	Standard deviation
SMD	Standardised mean difference
SPF	State and Peacebuilding Fund (World Bank Group)
TDA	Transitional development assistance

1. Background

1.1 The problem, condition or issue

In 2016, 1.8 billion people – nearly a quarter of the world's population – were living in situations of fragility. This figure is expected to grow to 3.3 billion by 2050 (OECD 2018). The increasingly protracted nature of many conflicts, alongside rising threats from the global pandemic, climate change and natural disasters, risks exacerbating the instability of already fragile places. Populations living in fragile contexts are at risk of violence and psychological distress, which can have long-term impacts on affected communities and future generations (Sangalang and Vang 2017). Moreover, fragile contexts tend to lag behind global averages on a range of socio-economic outcomes included in the 2030 Sustainable Development Goals (OECD 2018). Contexts of fragility are characterised by challenges across multiple sectors for already vulnerable groups (OCHA 2019). Individuals and households in fragile contexts are vulnerable along multiple dimensions of human security, facing challenges to their physical security, their ability to earn income, their health, access to education, and political rights (Giessmann et al. 2019). Fostering sustainable peace and addressing the drivers of threats to human security in such contexts could therefore produce a range of positive impacts.

Social cohesion is widely considered important in building and sustaining peaceful societies in fragile contexts. In many contexts, multidimensional threats to human security exacerbate tensions across social cleavages. Strong social cohesion may help safeguard against outbreaks of violence when communities are faced with shocks, and thus fostering social cohesion is often a core goal of interventions aiming to prevent violence (UN and World Bank 2018). The literature suggests that higher levels of trust, a shared sense of belonging, and norms around community participation – key dimensions of social cohesion – can create mechanisms to mediate or manage potential conflicts between groups across social cleavages. These cleavages can be between different ethnic or religious groups, or along other dimensions, such as between host communities and displaced populations (Colletta and Cullen 2000; Chan et al. 2006; De Berry and Roberts 2018; Kim et al. 2020). Improved social cohesion can also lead to some convergence on the needs and interests held by different groups, which can make working together constructively easier (Marc et al. 2013). External shocks to human security are a part of nature, and conflicts are often seen as inevitable, but it is the ways in which people respond to shocks and conflict that are responsive to change (Barron et al. 2006). Building social cohesion over a sustained period may therefore make a society more resilient, as people develop the skills, relationships and platforms to navigate conflicts and work together to adapt to threats. Efforts to build social cohesion are not only necessary in active- or post-conflict situations, but also to prevent future conflicts (OECD 2012) or to help address rifts caused by social or economic change (Marc et al. 2013). As such, a stronger understanding of the different options for building social cohesion, and their relative effectiveness, is crucially important.

The UN Sustainable Development Goals highlight the importance of building sustainable peace. Goal 16 aims to 'promote peaceful and inclusive societies for sustainable development, provide access to justice for all and build effective, accountable and inclusive institutions at all levels' (UNDESA n.d.). Within international policy debates on

peace, there has been a shift in focus from an immediate objective of ending active conflicts to a longer-term goal of preventing the outbreak of conflict in the future by promoting sustainable peace (Sönsken et al. 2019). In this way, the UN General Assembly (2016) refers to sustaining peace as both a goal and a process.

1.2 Defining the 'universe' of social cohesion

The term 'social cohesion' has its roots in the classical work of Durkheim (1897), which introduced the concept of 'organic solidarity' of modern society. In this early work, social cohesion was conceived of as the inter-dependence that exists between different members of society. Subsequent authors further developed this concept, offering different definitions focused on specific elements of social cohesion, depending on the author's area of interest. There remain many different definitions of social cohesion cited within recent literature, and there is more theoretical work needed to articulate and clarify the boundaries and relationships between social cohesion and related terms such as social capital (Kim et al. 2020) and social inclusion (Dugarova 2015).

For the purposes of this review, we adopt a modified version of the definition of social cohesion developed by Chan et al. (2006), drawing also on Kim et al. (2020) and Colletta and Cullen (2000). We start by defining a framework for understanding the 'universe' of social cohesion relevant to building sustainable peace, before outlining the elements of social cohesion that will be the focus of this review. We first distinguish between two broad types of social cohesion, vertical and horizontal. Vertical social cohesion refers to the relationships between the state and different groups within society (Kaplan 2009). Horizontal cohesion refers to the relationships between individuals and groups within a civil society (Chan et al. 2006).

Both of these broad types of social cohesion are shaped by social ties. These social ties can have one of two properties: bonding or bridging. *Bonding ties* refer to the relationships between members of a single societal group, i.e. intragroup relationships. These ties relate to the strength of cohesion within a group and also the extent of inclusion or exclusion of individuals in relation to their primary identity group (Colletta and Cullen 2000). Bonding is also sometimes referred to as interpersonal social cohesion (King et al. 2010a). *Bridging ties* refer to the relationships between different groups, or between members of different groups across a social cleavage (Colletta and Cullen 2000). Social identities may be formed around many different factors, such as ethnicity, religion, political party, nationality or region. Bridging is also sometimes referred to as intergroup social cohesion (King et al. 2010a).

Within this framework, social cohesion is a broad concept that can relate to both different levels of society (state, institutions, social groups and individuals) and different types of existing relationships (intragroup or intergroup). Social cohesion is also often defined and measured by a series of dimensions, which characterise the state of social cohesion within any level or relationship type. We focus on five key dimensions that emerge frequently in the literature, which we define below.

Trust: Trust refers to an individual's 'confidence in the reliability of a person or system, regarding a given set of outcomes or events' (Giddens 1990, p.34). The outcomes of interest are unspecified in this definition and could apply to a broad range of situations.

Chan and colleagues (2006) argue that trust is a core aspect of social cohesion because it is difficult to conceive of members of a society forming an effective and meaningful whole without trust among them. For example, trust could relate to the confidence that an individual has that another person or institution will not harm them, even if they were able to do so (Kim et al. 2020).

Sense of belonging: Sense of belonging refers to individuals' or groups' sense of shared identity. This sense of belonging is a critical aspect of social cohesion because social cohesion is formed through repeated interactions (Chan et al. 2006). These ongoing interactions are a contrast to situations where individuals may work together and trust each other in the short term in an emergency context, without the experience having an impact upon each individual's sense of identity (ibid.). It is important to recognise that people often have multiple identities across which they may form a shared sense of belonging, such as national, religious, ethnic or political identities. While the relative salience of different identities with which an individual feels associated may shift over time, building a shared sense of belonging along one shared identity rarely displaces other identities entirely.

Willingness to participate: This concept refers to individuals' willingness to participate in political or civil society (Chan et al. 2006). Willingness to participate is an important dimension of social cohesion because it brings individuals/groups or institutions together in interactions. Based on the recognition that participation does not always stem from altruistic motives, we differentiate it from 'willingness to help', which we describe below.[1]

Willingness to help: Willingness to help refers to individuals' willingness to engage in actions that benefit others based on altruistic motives (Chan et al. 2006; Lockwood 1999). Such actions might include community service, volunteering or donating to causes that benefit others (Moayed 2019). Chan and colleagues (2006) conceive of this as a core aspect of social cohesion because it is difficult to imagine members of a society being able to form an effective and meaningful whole without them being willing to help each other.[2]

---

[1] We also avoid reference to specifically 'civic' participation for similar reasons. Although civic participation and civic engagement are used in some discussions of social cohesion, these terms have been associated in some cases with specifically political forms of participation or with participation to create benefits for others (Dragolov et al. 2013; Kim et al. 2020).

[2] We prefer to retain the 'willingness to help' terminology of Chan and colleagues (2006), rather than adopt the term 'collective action norms' used by Kim and colleagues (2020) to describe moral codes that 'encourage costly actions that primarily benefit others' (p.17). Collective action terminology is sometimes linked to political action and also used frequently in debates about contributions and free riding related to public goods. These choices are not necessarily altruistic. Retaining and separating the two terms 'willingness to participate' and 'willingness to help' allows us to explicitly recognise that some actions may benefit one's self as well as society, while others are more altruistic. Both enable society to function more cohesively.

Acceptance of diversity: Acceptance of diversity refers to the extent to which individuals or groups recognise others' rights to belong, be trusted and/or be helped, even where there are differences in values, identities or lifestyles. Acceptance of diversity may refer to people within one's own group as well as across groups (Kim et al. 2020).[3]

As we elaborate in subsequent sections of the review, each of the five dimensions of social cohesion described above can manifest themselves, and thus be measured, through both attitudes and norms (somewhat subjective measures that tend to be self-reported) and through behaviours (somewhat more objective measures that tend to be observable).

This definition of social cohesion and the different constituent parts that we have elaborated above encompass a wide 'universe' of social cohesion. In the next section of this report, we define the specific portion of the social cohesion universe upon which we have focused in this systematic review.

1.3 Focusing on intergroup social cohesion

The definition of social cohesion we present above is comprehensive, and the range of different types of interventions that could influence elements of social cohesion defined in this way is too broad to cover effectively within a single review. Considering practical aspects of scope, as well as the policy interest and guidance of our advisory group, this review is focused on **the subset of interventions that aim to improve intergroup social cohesion by strengthening bridging ties between different groups**.

Interventions aiming to create sustainable peace by increasing social cohesion typically focus on strengthening social ties across individuals and groups located on different sides of societal cleavages, rather than within more homogenous communities, according to our advisory group. Given that conflict tends to occur along the lines of societal cleavages, building bridging relationships between groups on either side of that cleavage is likely a key objective of programmes seeking to foster sustainable peace (Siddique 2001). Due to investments in programming predominantly being focused on those interventions, our focus on this portion of social cohesion is likely to provide a broad base of evidence for our review.

Figure 1 depicts the choice we have made to focus first upon intergroup (horizontal) social cohesion and then on bridging ties within that sphere. We conceptualise that strengthening bridging ties may work through any or all five dimensions of social cohesion as defined above (trust, sense of belonging, willingness to participate, willingness to help, and acceptance of diversity). In the figure, we show the choices made at each level by marking the path not chosen at each level with a red border.

---

[3] Chan and colleagues (2006) explicitly rejected the notion of including aspects of diversity in their definition, arguing that values of diversity were not required for social cohesion. However, we consider that questions of diversity and interaction between individuals and groups with differences are particularly important when considering how social cohesion relates to sustainable peace (Kim et al. 2020; Jensen 2019).

Figure 1: Scope of the review on social cohesion
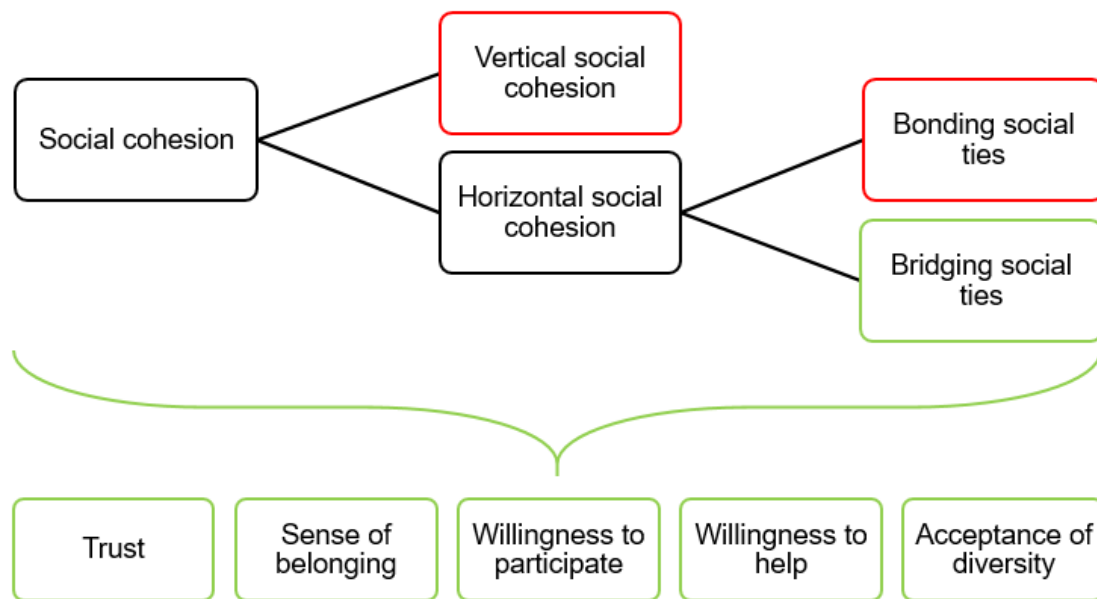
Source: Authors.

1.4 The interventions

There are many ways in which policymakers and practitioners seek to increase intergroup social cohesion for sustainable peace, including many that focus on strengthening bridging social ties between groups. However, there is no established typology of intervention types. Through our search process and literature review, we developed a typology of components frequently included in different interventions aiming to build intergroup relationships. Most interventions comprised more than one specific component.

A substantial body of scholarship exists testing the effectiveness of intergroup contact on social cohesion (Allport 1954; Pettigrew and Tropp 2005; Pettigrew and Tropp 2008; Hewstone and Swart 2011; Pettigrew and Hewstone 2017). Initially, four conditions were hypothesised as being essential for intergroup contact to be effective at improving social cohesion: equal status between groups within the intervention; common goals; intergroup cooperation; and support from authorities (Allport 1954). However, more recent meta-analyses have found that, while these conditions may facilitate the effectiveness, they are not required (Pettigrew and Tropp 2008). Additional facilitating conditions were identified: building knowledge of the 'other', reducing anxiety surrounding contact with the 'other', and empathy and perspective-taking (ibid.). Many of the interventions in this review introduce intergroup contact, and we note below the different ways in which that is addressed.

The five intervention types presented below were created in order to structure this review effectively. This is not an exhaustive list of all interventions that could, in theory, be conducted to build social cohesion. Rather, these categories were designed to facilitate analysis of the studies we found. This section presents the intervention types at a theoretical level. More details on the specific interventions we identified appear in Section 4.2.1.

1.4.1 School-based peace education

Peace education is often defined as a philosophy and process of building social, emotional and behavioural skills that facilitate non-violent conflict transformation (Harris and Morrison 2013). School-based peace education programmes are typically based on curricula designed to foster social and emotional learning (Hymel and Darwich 2018). While the curricula vary by context and age group, school-based peace education programmes typically aim to build skills such as perspective-taking, empathy and self-reflection, often seen as critical to building intergroup social cohesion. Peace education programmes are often based in a concept of 'positive peace', which aims to build the conditions for a peace that extends beyond the absence of physical violence, to include the absence of structural violence as well (Zembylas et al. 2016). School-based peace education programmes that target intergroup cohesion may or may not be implemented in schools where children from both groups are already in classes together. They distinguish themselves from other interventions in this review by nature of the age of the target participants (under 18 years) and in the implementation delivery modality, which is delivered by students' regular teachers with whom they already have a meaningful relationship.

These interventions may increase social cohesion by increasing understanding between different groups and promoting an acceptance of diversity. The training may focus on teaching values that promote positive peace (such as tolerance, solidarity and citizenship) or building awareness and understanding about other groups (such as other languages and cultures or alternative histories of the state). For example, in Turkey, a programme studied by Alan et al. (2020) focused on building perspective-taking skills in children in schools that had received large numbers of Syrian refugee children. The programme's goal was to encourage the children to build awareness of, and curiosity in, the experiences and perspectives of people different from themselves. These interventions are based on a long-term theory of change for sustainable peace, which recognises the importance of instilling skills and values for positive peace in children (Harris and Morrison 2013).

1.4.2 Collaborative contact

This group of interventions aims to build social cohesion by facilitating opportunities for collaborative contact for people from different groups, wherein they are required to work together on a shared project or team, such as a sports team. Whether implicitly or explicitly, they provide exposure to other groups and opportunities to collaborate, reducing anxiety around interacting with members of the other group in line with the literature on positive contact (Pettigrew and Tropp 2008). As such, these interventions may increase social cohesion by supporting the participation of different groups and increasing the willingness to help each other. For example, a soccer league established by Mousa (2019) ensured that certain teams had a mixture of Christian and Muslim players, and measured whether participants playing on mixed teams changed their perspective of the other group. Unlike the other interventions in this review, participants in collaborative contact interventions were frequently blind to the intergroup cohesion objective of the programme. For example, in Nigeria, a programme designed to facilitate collaborative contact between high-risk young men from different religious backgrounds masked itself as a computer literacy course, with a strong focus on learning through collaborative group work (Scacco and Warren 2018).

1.4.3 Intergroup dialogues

Interventions based on intergroup dialogues endeavour to create a safe space for people from different groups to engage in structured, deliberative discussions regarding the nature or drivers of tensions and ways to address them. As such, these interventions may increase social cohesion by supporting the participation of different groups. They may also increase trust and acceptance of diversity through better mutual understanding. Interventions incorporating dialogue sessions often include components of peace education in order to facilitate the dialogue. This peace education often focuses on strengthening the skills for conflict resolution and conflict transformation.

Conflict resolution training aims to enable parties to a conflict to work through their differences in a non-violent and productive way. Participants analyse their personal perspectives, analyse the other side's perspective and think about how their position could be more effectively framed to facilitate resolution (Berghof Foundation 2019). Conflict transformation is similar, but has a specific focus on changing the underlying conditions that have led or could lead to violence (ibid.). For example, in Ethiopia, Svensson and Brouneus (2013) evaluated the effects of a programme of 'sustained dialogue', comprising weekly dialogue sessions between students on a campus on which violence had broken out. The students first learned about relevant social and emotional skills, and then were led in a series of mediated discussions to try to build understanding of the drivers of the conflicts and develop action plans to address them.

Some of these interventions incorporate a specific focus on promoting reconciliation. Reconciliation dialogues work to repair relationships between groups, such as by bringing together ex-combatants and victims, or stakeholders from both sides of a conflict. One well-known method through which this is undertaken is truth and reconciliation processes. These processes seek to create open discussions of grievances between victims and perpetrators of violence. They are theorised to lead to both individual and societal healing through a process of recognising and acknowledging past crimes or injustices. As such, these interventions may increase social cohesion by supporting the participation of different groups and enabling greater trust in the future by reconciling past grievances. One example is the truth and reconciliation programme Fambul Tok in Sierra Leone, which brought together perpetrators and victims in a bonfire ceremony to testify to their experiences, and then a 'cleansing' ceremony in which perpetrators sought and were granted forgiveness (Cilliers et al. 2018).

1.4.4 Workshop-based peace education with intergroup contact and economic support

This group of interventions (which we refer to as 'workshops-contact-econ' for short) combined multiple approaches to building intergroup social cohesion in fragile contexts. They start with a peace education component, which typically focuses on some of the social, emotional and conflict resolution skills noted above. Different elements of the peace education component may target different populations and be delivered through different mechanisms. The peace education was typically delivered through workshops or training sessions. In some cases, participatory theatre is used to engage the audience in thinking critically about a situation. For example, in Finkel and colleagues (2018), communities came together to watch a play that highlighted a particular message of social cohesion or peace, and were then led in a discussion about what the play had

shown and how it related to their lives.[4] There are also cases in which the interventions aimed to build the capacity of local media for social cohesion by training journalists and media broadcasters to use their platforms to promote peace.

As participants build skills designed to strengthen intergroup cohesion, the interventions then provide platforms through which intergroup contact may be facilitated. In some cases, the platforms are specifically related to the conflict between the groups, such as community conflict prevention forums. In other cases, the platforms for contact are more social, such as community theatre productions. The expectation is that, ultimately, these interventions will build trust and willingness to participate across groups. For example, in Jordan, Mercy Corps organised community activities designed to facilitate casual contact between Jordanian host families and Syrian refugee families, including sports days, film screenings and handicraft workshops (Ferguson 2019).

Finally, the interventions further comprise an element of economic support, in which both groups are given the opportunity to collaboratively design and implement projects that will support both communities, such as small infrastructure projects. In this way, these interventions aim to address a key facilitating factor for intergroup contact identified in the literature, which is to identify and support the pursuit of common goals (Allport 1954). For example, Mercy Corps facilitated sessions with groups of farmers and pastoralists to discuss conflicts over land use and develop resource management plans supported by both groups, as a way of mitigating the likelihood of future conflicts (Dawop et al. 2019). As such, skills developed through interventions that progress from building skills to facilitating cooperative contact may increase social cohesion by supporting the participation of different groups in inclusive community development. These interventions typically have broader target groups, are often larger in scale and of longer duration than other interventions included in this review. These multi-component interventions often also include further context-specific components. However, since the peace education, contact and support for economic projects are their common components, we refer to these interventions throughout the remainder of this review as workshops-contact-econ interventions.

1.4.5 Media for peace

These interventions use a range of media to provide messages to target audiences that promote tolerance, non-violence, cooperation, reconciliation, collective action in problem-solving and dispute resolution (La Ferrara 2016). Interventions that primarily consisted of media components to promote peace and social cohesion were identified in two forms: peace messaging campaigns and 'edutainment,' typically in the form of radio dramas.

Peace messaging campaigns aim to distribute messages of positive peace, advocating for non-violence and tolerance by providing exposure to counter-narratives (Ferguson 2016). These interventions may use TV adverts, radio broadcasts, posters or articles in the print and electronic press. For example, in an intervention evaluated by Vicente and Vilela (2020) in Mozambique, faith leaders developed a campaign to promote tolerance and prevent violent extremism based on religious teachings.

---

[4] We recognise the overlap between this element of the workshops-contact-econ interventions and media for peace interventions. However, because in the context of workshops-contact-econ interventions, the media component is typically a small one within a larger intervention, we include these kinds of studies in this group.

Edutainment programmes aim to engage participants in thinking about social cohesion by embedding messages within a story. To this end, radio dramas aim to encourage listeners to see members of the other group as sympathetic characters with whom they may share commonalities, or with whom they can empathise (La Ferrera 2016). In this way, these interventions aim to improve social cohesion by listening to positive examples of intergroup cohesion, and building comfort with the idea of trusting or interacting with members of the other group. In the case of the Radio La Benevolencija programmes, the campaign took the form of a radio drama that presented a fictionalised parable of a local conflict (Bilali et al. 2016). The story of the drama showed sympathetic characters on both sides getting to know each other and overcoming their prejudices, fears of each other and history of violence.

1.5 How the intervention might work

While all included intervention types seek to have an effect on the same outcome, the pathways through which they aim to affect that outcome are varied. Due to the wide range of approaches undertaken with the aim of influencing intergroup social cohesion, there is no single theory of change that may help explain each of the interventions included in this review. Here we identify and outline four key pathways through which the interventions may work. These pathways combine a more traditional model of behaviour change that focuses on altering individuals' *knowledge*, *attitudes* and/or *behaviours* (Schrader and Lawless 2004), and a more recent model that focuses on shifting *social norms* in order to change individual attitudes and behaviours (Tankard and Paluck 2016). While each pathway starts with one of these four components, the others may follow.

1.5.1 Pathway 1: Knowledge
There are three generally accepted forms of knowledge: declarative, procedural and conditional (Schrader and Lawless 2004). These are defined as knowing what, knowing how and knowing when and why, respectively (ibid.).

The pathway that leads from creation of knowledge to improved social cohesion may work in several ways. For example, interventions that use campaigns, education and interaction may help different groups to learn about other group's culture, history and role in past conflicts as either victims or perpetrators. Creating a higher level of knowledge about these factors may help to increase understanding between different groups, reduce fears about the unknown or lead to individuals acknowledging their own culpability in past incidents. Quinn refers to these 'basic truths and knowledge among these groups' as 'thin sympathy' (2020, p.27). This basic knowledge could then provide a foundation for changes in one or more of the five dimensions of social cohesion. For example, knowledge about another group may make their actions and intentions more comprehensible and therefore increase levels of trust. In this way, increased knowledge may lead to attitudinal change.

Social cohesion interventions whose pathway begins with knowledge may also focus on teaching specific skills that help to foster peaceful interactions in the future. These could lead to changes in procedural forms of knowledge. Through increasing individuals' knowledge of how to negotiate and resolve conflict, for example, interventions may foster more peaceful interactions between groups and avoid conflict in the future. In this manner, interventions that work through increasing knowledge may also have an influence on behaviours.

1.5.2 Pathway 2: Attitudes

Attitudes can be defined as the emotions, feelings and reactions stimulated by a particular object, action or thought (Schrader and Lawless 2004). These emotions, feelings and reactions tend to have either positive or negative connotations (ibid.). As such, attitudes have more of an affective character than knowledge.

Interventions that seek to work through a pathway of attitude change can contribute to social cohesion by shifting negative attitudes that individuals have towards another group or members of that group to more positive attitudes. For example, campaigns and media images may seek to portray members of another group in a positive light and to depict peaceful interactions between groups. In this way, individuals may shift from attitudes of distrust and antipathy to attitudes of trust and sympathy. These shifts in attitudes may be manifested in some or all of the five dimensions of social cohesion, leading to positive and peaceful interactions in the future. In this manner, interventions that work through changing attitudes may also have an influence on behaviours. If one or more interventions manages to shift a significant share of a population's attitudes, this may, with time, also influence social norms.

1.5.3 Pathway 3: Behaviours

Behaviours can be defined as the observable 'way in which a person, organism, or group responds to a set of conditions' (Schrader and Lawless 2004, p.11). Interventions may directly aim to alter people's behaviour. For example, interventions may bring members of different groups together to interact in safe spaces when they previously had no direct interaction. Interventions that work through a behavioural pathway seek to create observable manifestations of one or more of the five dimensions of social cohesion.

Interventions that seek to change behaviours can alternatively operate through indirect routes by first creating shifts in knowledge or attitudes, which in turn alter people's observable behaviour. Additionally, if a significant amount of a particular population shifts their behaviour in a particular direction, interventions working through the pathway of behaviours may have an impact on social norms.

1.5.4 Pathway 4: Social norms

Social norms can be defined as 'our perceptions of what is typical or desirable in a group or in a situation' (Tankard and Paluck 2016, p.184). Some interventions aiming at social cohesion aim to change people's *perceptions* of norms in order to encourage an individual to conform to those norms through their attitudes and behaviour (ibid.).

Research suggests than an individual is likely to shift their attitudes and/or behaviour if they believe that most people in their society would deem that attitude or behaviour typical or desirable (Tankand and Paluck 2016). It is not necessarily the case that people in their society actually do think or behave in that way. For example, a media intervention may broadcast a soap opera that portrays intergroup marriage as accepted by relevant communities, even though it is not yet an accepted norm. The viewers may therefore perceive that the social norm is to be accepting of intergroup marriages and may then begin to shift their own attitudes and behaviours in line with that perceived norm. Over the longer term, this could influence social norms themselves, rather than people's perceptions of them. In this way, an intervention that seeks to have an effect on social cohesion through social norms may also work through a knowledge pathway

by providing information and will also seek to have subsequent effects on behaviour and attitudes.

1.5.5 Pathway linkages and progression

Rather than four distinct pathways, there are clear linkages between them. 'Head-first' approaches, which aim to change beliefs in order to then alter behaviour, can be contrasted with 'feet-first' approaches, which aim to alter behaviour in order to then change beliefs (McCauley 2002). For example, increased knowledge may alter both the way people feel about a situation and the way they act within it – affecting the pathways of attitudes and behaviours (Schrader and Lawless 2004). Interventions that change people's behaviours by creating new interactions may lead the individuals involved to better understand about each or alter their feelings towards each other – affecting pathways of knowledge and attitudes. While models of behaviour change often assume that changes in knowledge and attitudes produce changes in behaviours, there is some evidence that norms and behaviours may change before attitudes or personally held beliefs change (Paluck 2009). In this sense, the pathways of change move in different directions. In some situations, interventions may even need to work on different pathways simultaneously: a reform in school structure may be insufficient to build social cohesion, for example, without parallel reforms in curriculum (King 2014).

To understand the mechanisms through which different pathways may be triggered, we adopt and expand upon a typology of three mechanisms through which direct and indirect exposure to 'the other' may improve intergroup social cohesion, developed by Amanda Blewitt (n.d.). The first, 'seeing the other', refers to approaches that introduce participants to others' perspectives through education or media, but do not necessarily involve any direct, personal interaction between groups. This type of indirect contact aims to build knowledge of the other, a facilitating factor identified in the contact (Allport 1954). The second type, 'meeting with the other', comprises approaches that bring people from different groups together in person. The third type, 'talking with the other', is based on facilitated, purposeful dialogue between people from different groups regarding the nature of their differences and opportunities for building bridging ties.

Building on this tripartite typology, within the 'seeing the other' category we further differentiate between approaches that focus on respecting individual differences and those that focus on group-level differences. As all subsequent approaches necessarily involve meeting each other as well, we use 'meeting with the other' to refer to approaches where this is the only change in contact triggered by the intervention. Finally, we add a fourth category, 'collaborating with the other', which comprises approaches wherein participants from different groups are given opportunities to work together as a team. This collaboration may be on an actual sports team, or through opportunities to develop and implement joint projects, which may or may not be related to the source of tensions. This approach assumes that, by working collaboratively towards a shared goal or project, participants will get to know each other beyond the group-based identities, and thereby strengthen intergroup social cohesion (Allport 1954). In practice, the pathways are unlikely to work in a linear manner.

The process of change within these pathways may be incremental and slow. Over time, changes that occur through these four pathways may create a virtuous circle of social cohesion, whereby increases in knowledge and shifts in social norms create positive

changes in behaviours, which then shift people's core attitudes and beliefs. However, the pathways are unlikely to work in a linear manner. In the short term, and particularly in fragile, violence- and conflict-affected situations, there may be more of a pendulum of progress and backlash before long-term positive results emerge (Kim et al. 2020). This pattern means that there may be positive progress along one pathway but limited or negative progress through another. For example, an intervention that works through the pathway of knowledge may raise people's awareness about a history of conflict, but this could increase negative attitudes towards another group through learning about specific atrocities that were perpetrated.

1.5.6 Sustainable peace as final outcome

Through these four pathways, varied intervention types may help to increase levels of bridging social ties across different groups in society. By increasing one or more of the key dimensions of social cohesion – trust, sense of belonging, willingness to participate, willingness to help, and acceptance of diversity – these interventions seek to strengthen the ties between different social groups. Strengthened social cohesion should in turn increase the likelihood that peace will be sustained. Strengthened social ties between different groups should also make these societies more resilient in the face of future shocks that could lead to tensions and conflict.

Even if interventions are successful in their goal, however, horizontal social cohesion between different groups is unlikely to be sufficient for the ultimate goal of sustainable peace. There are other elements of social cohesion that are important for creating sustainable peace, including vertical relationships between society and the state. Interventions that focus on bridging horizontal social cohesion may have an influence on other parts of a broader conception of social cohesion that may be positive or negative. For example, increasing the strength of horizontal ties by increasing people's propensity to participate in societal interactions may also increase their likelihood to participate in political processes, thereby strengthening vertical cohesion. Alternatively, interventions that create bridges between individuals of different groups may reduce the previous level of in-group solidarity within those specific groups, thereby weakening bonding ties.

Moreover, there are multiple other factors in society beyond social cohesion that can affect prospects for peace. Shifts in those conditions are also likely to influence the prospects for social cohesion (Anderson 2004). As such, none of the included interventions is expected to create sufficient conditions for sustainable peace on their own. Each is, however, hypothesised to strengthen the conditions for peace.

In analysing the pathways of change of our included interventions, we need to consider risks and assumptions associated with their implementation. Some of these risks and assumptions, as noted above, will relate to how changes in one pathway may affect other pathways or may have spillover effects into other portions of the social cohesion universe. Others will relate to the challenges inherent in working on complex social change in fragile contexts. For example, it may be difficult for external policymakers to design and implement appropriate interventions to strengthen cohesion in complicated political and social contexts (Colletta and Cullen 2000; Fearon et al. 2009; Marc et al. 2013). Some of the interventions may also create risks for participants based on the sensitive nature of topics related to peace and fragility. For example, a recent non-systematic review found that some interventions designed to improve intergroup relations

increased symptoms of post-traumatic stress disorder (PTSD) for victims of violence (Nolan and Knox 2019). We discuss issues like these in our analysis of the barriers and facilitators along the interventions' causal chains.

1.6 Why this review is needed: relevance to policy and practice

In recent decades, substantial funding has been directed to fragile states, especially towards social protection and peacebuilding activities. In 2016, approximately 11 per cent of total official development assistance (ODA) was allocated to peacebuilding activities in extremely fragile contexts, and approximately 10 per cent in fragile contexts (OECD 2018). Overall, 20 per cent of total ODA was dedicated to basic safety and security, core government functions and inclusive political processes activities (OECD 2018). In 2017, US$56.3 billion of ODA was provided to fragile states (OECD 2019; Fund for Peace 2019).

Many countries have identified supporting fragile states as a key priority. For example, the UK has committed to spending at least 50 per cent of the DFID budget in fragile states (UK Aid 2015); in practice, this has amounted to 53 per cent in 2015, and 57 per cent in 2016 and 2017 (DFID 2019). In Germany, ODA dedicated to conflict, peace and security has increased rapidly, from US$127 million in 2007, to 476 million in 2016, of which a significant share was distributed to civilian peacebuilding, conflict prevention and resolution (Deneckere and Hauck 2018). Notably, the German Federal Ministry for Economic Cooperation and Development (BMZ) commissioned and executed the German development agency GIZ's transitional development assistance (TDA) from 2016 to 2022, as an instrument to provide assistance for people in crises, disasters and conflicts and to enhance the resilience of people and institutions. Co-funded with other institutions, including the EU, TDA has been used in countries such as Turkey, Chad (GIZ 2020) and Burundi (BMZ 2015).

Multilateral donors have also invested heavily in initiatives to build peaceful societies in fragile contexts. For example, the State and Peacebuilding Fund (SPF) is the World Bank Group's global multi-donor trust fund created in 2008 to foster innovative strategies for state- and peacebuilding in areas affected by fragility, conflict and violence. It aims to boost resilience and support socio-economic conditions for peaceful, stable and sustainable development. It is now worth over US$342.9 million, and has supported over 200 interventions in 57 countries (SPF 2018). Similarly, the United Nations Peacebuilding Fund (PBF) was launched in 2006 to support programmes, activities, actions and organisations committed to restoring peace by promoting a peaceful resolution of conflict, revitalising the economy and re-establishing essential administrative services. From 2006 to 2017, it approved US$772 million to fund peacebuilding projects in 41 countries, and from 2017 to 2019, this was scaled up to an additional US$531 million for 51 countries (PBF 2020). By 2018, this corresponded to 338 peacebuilding projects in 50 countries emerging from or dealing with the impact of conflict (DFID 2019).

While it is clear that significant resources go towards efforts to improve social cohesion, no global statistics present a clear total. Varied programme classification, as well as the definition and measurement of social cohesion, are largely inconsistent. As a result, gathering aggregate information on the level of investment in efforts to build social cohesion is challenging.

Notably, however, GIZ has set up a Fund for Social Cohesion specifically dedicated to financing numerous projects in low- and middle-income countries (L&MICs) to foster social cohesion, peace and conflict resolution. Examples of programmes supported through this fund include a four-year Civil Peace Service programme in Burundi with a total €4.3 million financial commitment (GIZ n.d.).

Given the large amount of funding from a broad range of donors, it is important to know which type of social cohesion interventions are effective.

1.6.1 Review of the literature

In some ways, the evidence available on social cohesion is relatively extensive. In the recent evidence gap map (EGM), *Building peaceful societies* (Sonnenfeld et al. 2020), approximately 50 impact evaluations measured the effects of interventions on social cohesion outcomes. Yet, there remains much to do to understand this evidence.

First, this evidence base is fragmented, spanning a broad range of very different interventions. Previous syntheses have tended to focus on just one type of intervention. For example, there has been a substantial amount of work dedicated to understanding the effect of community-driven development (CDD) interventions on social cohesion (e.g. White et al. 2018). They did not, however, examine the broad range of other interventions that may affect social cohesion outcomes.

Second, existing studies often look at social cohesion outcomes as secondary outcomes, while the primary focus and theory of change relate to outcomes of particular dimensions of human security, such as strengthening education or food security. When social cohesion is not the primary outcome of focus in a review, it is unlikely to capture the different dimensions of social cohesion that we have outlined in Section 1.2. For example, while the EGM above identified five systematic reviews that included reference to social cohesion outcomes, only one (King et al. 2010b) contained a focus on social cohesion outcomes. For example, Czuba et al. (2017) included social cohesion outcomes in their examination of in-kind food assistance on pastoralists in humanitarian crises, but the article's focus was elsewhere.

Third, the reviews that do exist are restricted to specific contexts or out of date. For example, in 2010 King and colleagues published a systematic review of social cohesion projects in Sub-Saharan Africa (King et al. 2010b). They found only eight includable studies, which evaluated CDD programmes and curriculum interventions. This review was rigorous, following Cochrane requirements; however, the geographic evidence base was limited to one region and the evidence base must now be updated with a decade worth of studies. In another example, Paluck et al. (2019) evaluated 27 intergroup contact studies, based on a previous meta-analysis by Pettigrew and Tropp (2006). In their research, there are some inclusion criteria applied – for example, the study design must include random assignment and delayed outcome measures. However, more than half of the reviewed studies were implemented in high-income contexts, such as the United States, and we identified at least three studies from L&MIC contexts that were published before their search was finalised and would meet their criteria but were not identified by their search process. This suggests that the generalisability of their findings may be biased towards high-income contexts. Further, as the search was completed in 2016, the findings do not reflect the substantial growth in the literature in the years since.

Fourth, not all thematically relevant reviews follow systematic procedures. For example, Idris (2016) examined the relationship between social cohesion and a range of interventions, including CDD programmes, education, social protection, jobs creation and support to civil society in post-conflict societies. While the range of interventions and focus on social cohesion is relevant to our review, this rapid literature review did not use transparent and systematic methods to identify and analyse the evidence that feeds into its findings.

As such, there is no existing, up-to-date systematic review focused on interventions to improve social cohesion in fragile L&MIC contexts. This systematic review addresses this gap and, in so doing, provides a resource to inform decisions about the types of programmes and intervention features that may be most effective in fostering social cohesion in such contexts.

2. Objectives

The main objective of this review is to identify, appraise and synthesise evidence on the effects of interventions that aim to promote intergroup cohesion for sustainable peace in fragile communities.

We compare the effects of different types of programmes that aim to promote intergroup cohesion. In doing so, we address the following review questions:
1. What are the effects of interventions that aim to promote intergroup cohesion in fragile communities on social cohesion outcomes? (*Review Question 1*).
2. What evidence is available on sustainable peace outcomes, including resilience and human security for participants? (*Review Question 2*).
3. To what extent do effects vary by population group? (*Review Question 3*).
4. What factors relating to programme design, implementation, context and mechanism are associated with better or worse outcomes? (*Review Question 4*).
5. What evidence is available on programme costs and incremental cost-effectiveness in included studies of effects? (*Review Question 5*).

3. Methods

3.1 Criteria for considering studies for this review

3.1.1 Types of studies
To answer Review Questions (RQ) 1–3, we included studies that use experimental or quasi-experimental study designs to measure the net change in outcomes that are attributed to an intervention. Specifically, we included the following study types:
1. Randomised controlled trials (RCTs), with assignment at individual, household, community or other cluster level, and quasi-RCTs using prospective methods of assignment such as alternation;
2. Non-randomised studies with selection on unobservables:
   a. Regression discontinuity designs, where assignment is done on a threshold measured at pre-test, and the study uses prospective or retrospective approaches of analysis to control for unobservable confounding;
   b. Studies using design or methods to control for unobservable confounding, such as natural experiments with clearly defined intervention and comparison

groups, which exploit natural randomness in implementation assignment by
decision-makers (e.g. public lottery) or random errors in implementation, and
instrumental variables estimation;

3. Non-randomised studies with pre-intervention and post-intervention outcomes
data in intervention and comparisons groups, where data are individual level
panel or pseudo-panels (repeated cross-sections), which use the following
methods to control for confounding:
   a. Studies controlling for time-invariant unobservable confounding, including
   difference in differences, or fixed- or random-effects models with an interaction
   term between time and intervention for pre-intervention and post-intervention
   observations;
   b. Studies assessing changes in trends in outcomes over a series of time points
   (interrupted time series, ITS), with or without contemporaneous comparison
   (controlled ITS), with sufficient observations to establish a trend and control
   for effects on outcomes due to factors other than the intervention (e.g.
   seasonality); and

4. Non-randomised studies with control for observable confounding, including non-
parametric approaches (e.g. statistical matching, covariate matching, coarsened-
exact matching, propensity score matching) and parametric approaches (e.g.
propensity-weighted multiple regression analysis).

To address Review Question 4 relating to programme design, implementation, context
and mechanisms, we included a broader range of evidence on programme design and
implementation. In addition to drawing on descriptive information in included experimental
and quasi-experimental studies, we also included process evaluations, qualitative
studies, programme and project design and implementation documents, where available
to contextualise and better understand the included impact evaluations.

To address Review Question 5 relating to unit cost, cost-efficiency, cost-effectiveness or
cost–benefit, we included economic evaluations, project documents and process
evaluations of included programmes and cost information in included impact evaluations,
where available.

3.1.2 Types of participants
We included studies from fragile communities in L&MICs, including participants from the
general population and those from specific population subgroups, such as displaced
populations, refugees, women or youth. Specifically, we included studies that targeted at
least two different social groups between whom tensions persisted that were at risk of
becoming, had recently become or were currently violent. As discussed above, these
groups tend to fall into different identity groups formed along social cleavages.

3.1.3 Types of interventions
We included studies of interventions that aimed to improve intergroup social cohesion by
strengthening bridging social ties between two or more social groups divided by
cleavages as discussed above. The interventions could be implemented by governments,
aid agencies or other non-state actors.

Beyond this broad intervention criteria, inclusion decisions were based primarily on
whether studies measured at least one dimension of social cohesion – the primary

outcome of interest in this study (defined in detail above). The reason for this approach to inclusion is largely that the field is fairly fragmented, with a lack of clear intervention typologies or standardised programmes. Given the open questions regarding boundaries between social cohesion, social capital and social inclusion, we did not require studies to include the term "social cohesion." Instead, we adopted a nuanced approach that aimed to identify whether the core focus of the intervention was to strengthen relationships across social cleavages, which is at the heart of the concept of "bridging" intergroup social cohesion that forms the focus of this review. As we describe above, examples of the types of programmes we included are: school-based peace education; collaborative contact; intergroup dialogues; workshop-based peace education with intergroup contact and economic support; and media for peace.

When assessing studies for inclusion, one of the first criteria is that the study must focus on an intervention. We defined intervention as an organised set of activities implemented under a given policy, programme, project or experiment, with the aim of creating positive impacts. We thus excluded studies that, for example, measured the effects of exposure to violence on social cohesion. Within media studies particularly, this meant that we excluded studies that measured the impact of "naturally occurring media content" wherein the content of the media to which participants in the research were exposed was outside the control of the stakeholders in the study (Moehler 2014, 11).

Many interventions in this field are made up of multiple components, and such interventions were included if at least one of the primary components met our inclusion criteria.

To help more clearly delimit the scope of our review, we also adopted a number of intervention exclusion criteria.

First, we exclude community-driven development or reconstruction interventions. Previous systematic reviews have found that, in practice, the majority of CDD interventions focused on increasing bonding or interpersonal forms of social cohesion, rather than working across groups (King et al. 2010b). This is partially because the size of the community targeted with CDD interventions often only includes one identity group. Further, there have been recent systematic reviews published that examined this intervention category in depth (e.g. White et al. 2018). It is therefore unlikely that our review would be examining much new evidence on those interventions. As such, by excluding those interventions from our review, we avoid duplication of effort.

Second, we exclude interventions that aim to promote sustainable peace by providing economic support only, such as cash transfers or livelihoods training. The intergroup social cohesion component of such interventions aims either to reduce the economic inequality between two groups, or to mitigate tensions over scarce resources where both groups are vulnerable, such as when refugees settle in vulnerable host communities. While this could be viewed as a route to intergroup social cohesion, we primarily exclude this intervention type due to the findings of a recent 3ie EGM (Sonnenfeld et al. 2020). The EGM examined interventions aiming to build peaceful societies in fragile contexts and identified numerous ongoing studies on cash transfer, subsidies and livelihood programmes that planned to assess social cohesion outcomes. With so many studies in progress, a synthesis of this portion of the literature may thus be premature. This exclusion is limited to interventions that are only or primarily economic in nature. Where

some form of economic support is provided as a component of a larger programme alongside components such as the workshops-contact-econ interventions described above, it is included.

Third, we exclude mental health and psychosocial support interventions that are focused predominantly on individuals and do not have a component that focuses on interaction between groups. Psychological components of reintegration programmes for specific groups of non-combatant and gang members are, however, includable. We also made this decision based predominantly on the findings of the EGM mentioned above (Sonnenfeld et al. 2020), which found a substantial evidence base that would be better synthesised by a separate, more targeted review.

Fourth, we have excluded interventions that focus on strengthening bridging ties between groups that are not past or likely parties to violent conflict. For example, we exclude interventions focused on strengthening bridging ties between people affected by HIV/AIDS and wider society. Our interest is in social cohesion as a means to foster sustainable peace in fragile contexts.

Finally, we did not include interventions focused on gender, despite recognising that positive gender relations are important for sustaining peaceful societies. This is partly for the similar reason that, while there may be situations of violence between men and women, societal-level violent conflict does not usually unfold along these group lines. 3ie is also undertaking a separate systematic review to examine gender in fragile contexts.

Table 1 in Online appendix A.3 presents examples of studies that did and did not meet our inclusion criteria, with explanations as to why, in order to illustrate these nuanced criteria.

3.1.4 Types of comparisons
We included studies that compared the effects of an intervention aiming to promote stronger bridging ties between different groups in society against similar situations where communities either received an unrelated intervention (such as a health intervention) or did not receive an intervention and continued under business-as-usual conditions. We excluded studies where there was no control group.

3.1.5 Types of outcome measures
To be included in the review, studies had to report at least one measure of a social cohesion outcome. This measure could relate to any of the five dimensions of social cohesion: trust, sense of belonging, willingness to participate, willingness to help, and acceptance of diversity. For studies with measures of our primary outcomes, the social cohesion outcomes, we also included measures of sustainable peace outcomes, as defined below. However, studies only reporting measures of peace, resilience or human security, without a focus on social cohesion, were excluded.

Primary outcomes
Our primary outcomes of interest are social cohesion outcomes related to bridging social ties across groups. These measures include one or more of the social cohesion dimensions: trust, sense of belonging, willingness to participate, willingness to help, and acceptance of diversity.

Studies could use different measures to assess these outcomes. These may include self-reported attitudes and norms, observable behavioural outcomes from games played within the context of the intervention, or 'natural' games designed to mimic real-world conditions.

Below we list several social cohesion outcomes that are relevant to the five dimensions of social cohesion. The list draws on work on existing social cohesion outcomes and measures (Baumgardner-Zuzik et al. 2020; Dragalov et al. 2013; Kim et al. 2020) and a preliminary survey of the literature. In many cases, these are broken down further by whether they refer to a respondent's personal knowledge, attitudes or behaviours ('self') or a respondent's perceptions of others' knowledge, attitudes or behaviours ('others'). For simplicity, we note where this distinction could apply but define the outcome using language applicable to measures related to participants themselves.

Trust:
- Trusting (self/others): Measure of respondents' trust towards others
- Mistrust (self/others): Measure of respondents' lack of trust towards others
- Trustworthiness (self/others): Measure of respondent acting in a trustworthy way

Sense of belonging:
- Shared: Measure of a shared sense of belonging (positive measures)
- Divisions: Measure of a sense of divisions or tensions between groups
- Separate: Measure of an exclusionary sense of belonging (negative measures)

Willingness to participate:
- Active (self/others): Measure of respondents' actual participation
- Openness (self/others): Measure of respondents' openness to hypothetical participation
- Anxiety: Measure of respondents' anxiety about participation
- Perceived value (self/others): Measure of respondents' perceived value (or lack thereof) of participating
- Refusal (self/others): Measure of respondents' rejection of actual or hypothetical participation

Willingness to help:
- Active (self/others + receive/give): Measure of active giving/receiving help
- Openness (self/others + receive/give): Measure of respondents' openness to hypothetical help
- Perceived value (self/others + receive/give): Measure of respondents' perceived value (or lack thereof) of helping
- Choosing not to help (self/others + receive/give): Measure of respondents' choice not to help in the context of a behavioural game

Acceptance of diversity:
- Tolerance (self/others): Measure of respondents' tolerance of others
- Intolerance (self/others): Measure of respondents' intolerance of others
- Acceptance of multiple perspectives (self): Measure of respondents' recognition of multiple valid perspectives
- Non-recognition of multiple perspectives (self): Measure of respondents' rejection of multiple perspectives (not simply ignorance)

- Inclusive victimhood: Measure of respondents' sense of common suffering or blame (both suffered / both at fault)
- Exclusive victimhood: Measure of respondents' sense of exclusive suffering ('only my group') or blame ('all their fault')
- Bias (self/others): Measure of respondents' bias towards others

Studies that only reported outcomes of intragroup or interpersonal cohesion or vertical cohesion were excluded.

Secondary outcomes

Secondary outcomes of interest are those related to sustainable peace, including measures of peace and human security and resilience. These outcomes are relevant for answering Review Question 4. We list some potential outcomes that could be used to measure effects on resilience and human security, based on the *Berghof Glossary on Conflict Transformation and Peacebuilding* (Berghof Foundation 2019).

Measures of resilience could include, but are not limited to:
- Social norms and behaviours regarding conflict mitigation and violence;
- Strength and quality of community-based social safety net systems; and
- Strength and quality of (re)integration systems or resource management systems.

Measures of human security could include:
- Economic security;
- Educational security;
- Food security and nutrition;
- Health security;
- Personal and community security;
- Political security; and
- Environmental security.

In addition to extracting data on outcomes related to the primary and secondary categories outlined above, we also included any other outcome measures reported by included studies, including any adverse effects.

Timing of measurement

We included any follow-up duration, extracting data on outcomes at multiple follow-ups if applicable. However, we exclude interventions that aimed to assess a participant's immediate response to a short activity in which they just took part, known as 'lab-in-the-field' studies. Although these types of studies may be useful for identifying potential mechanisms and designing interventions, they are not typically designed to lead to long-term changes on social cohesion outcomes. However, we do include studies that employ lab-in-the-field approaches to analyse the effects of a larger intervention.

3.1.6 Types of settings

We included interventions conducted in fragile contexts in L&MICs only. This includes fragile states, where governments are unwilling or unable to provide basic public services in the areas of security, the rule of law and social services (BMZ 2020). Fragile states often, but not always, overlap with countries that we consider 'conflict affected'. However, contexts of fragility may vary within a particular country; conditions of fragility may be

present in communities even within states not considered to be fragile at a national level. We therefore also included studies of fragile communities in non-fragile states, such as places with high levels of gang or intergroup violence.

Specifically, we included studies where the population falls within at least one of the following categories:

(1) The country received a score of over 90 on the Fragile States Index (FSI; Fund for Peace 2019a) in the year in which the intervention was started. The list includes data dating back to 2006.

(2) For interventions that were implemented pre-2006, or in countries scoring below a 90 on the FSI scale, we included studies that met the following two conditions: (a) the community or communities where the intervention was implemented were affected by intergroup violence, including gang violence, or high levels of intergroup tension, or the potential for tension (for example, communities with a large influx of refugees or migrants); and (b) this is defined by the study as one of the motivations for implementing the intervention. (If the study does not discuss the presence of violence or tensions, even though these existed, the study is excluded.)

Other criteria

We included both completed and ongoing studies, in the form of protocols of ongoing studies that appear to meet all other inclusion criteria and studies listed in registries of ongoing impact evaluations. This is a rapidly expanding field and listing ongoing studies will give an overview of where current knowledge gaps might be filled in the future.

We included studies published in any language, although our search terms were in English only. Following the findings of the *Building peaceful societies* EGM (Sonnenfeld et al. 2020), which found no potentially relevant studies published earlier than 2000, we included studies published in 2000 or later.

3.2 Search methods for identification of studies

We developed a systematic and extensive search strategy in consultation with an information specialist. It consisted of a number of different approaches, including electronic searches of academic databases, hand searches of organisational websites, and forwards and backwards citation searches of included studies.

3.2.1 Electronic searches

The comprehensive search strategy we developed for academic database searching aimed to capture key synonyms for interventions targeting intergroup cohesion, alongside search strings focused on impact evaluation terms and L&MIC geographies. The search string was adapted for each of the six academic databases searched: EconLit (Ovid); Scopus; Social Sciences Citation Index (SSCI); PsycINFO; Ebsco Databases (ERIC; International Political Science Abstracts; Communication & Mass Media); and Ebsco Discovery Service (World Bank e-Library REPEC). An example of a search string for an academic database is included in Online appendix A.1 Sample academic search string.

In addition to the academic searches, we further searched over 30 different specialist organisational databases, bilateral and multilateral agencies, and general repositories of impact evaluations in international development in an effort to identify grey literature

(unpublished or published non-commercially). The search strategy for the grey literature aimed to balance rigour with practical constraints due to the low sophistication of most sites' search engines. Keywords focused on impact evaluation terms were searched, with relevant controls offered by the site applied (such as restricting the search to publications instead of blogs and press releases). The full list of sites searched and the detailed grey literature search strategy can be found in Online appendix A.2 Grey literature search strategy.

3.2.2 Searching other resources

In addition to the electronic searches, we further carried out forwards and backwards citation searches of included papers. We sought suggestions for potentially relevant studies through experts on our advisory group.

3.3 Data collection and analysis

In order to synthesise the evidence for what works, for whom, and under what conditions, for building intergroup social cohesion in fragile contexts, we conducted a rigorous and systematic search to identify all papers that met our inclusion criteria. We then extracted substantial descriptive and quantitative data from included studies, and undertook an in-depth critical appraisal to identify potential risks of bias in the designs and analyses. In order to appropriately compare findings across studies drawing on different designs and analysis methods, we first calculated standardised effect sizes for all outcomes reported in included studies. We then used meta-analysis to synthesise the findings for particular intervention-outcome combinations wherever possible. Given the heterogeneity of included studies, however, this was not always possible, and so in these cases we synthesised the effects through narrative analysis. Following the quantitative synthesis, we strove to interpret the findings and explain heterogeneity in effects. To do this, we drew on descriptive data and a framework synthesis of key factors that may influence the ways in which different interventions work for different populations under different circumstances, as well as drew on the findings from the risk of bias analyses. The following sections detail in depth the systematic process we followed to undertake the selection and synthesis of studies in this review.

3.3.1 Selection of studies

We imported all search results into the web-based software programme for review management EPPI-Reviewer 4 and removed duplicates. All studies were double-screened against the review inclusion criteria using information available in the title and abstract by two independent research assistants. Any disagreements were then resolved through conversations with a core review team member. Where a study's title and abstract did not include sufficient information to determine relevance, the study was included for review at full text.

We used the text mining capacity in EPPI-Reviewer to classify studies into groups based on their probability of inclusion in the review. We then used these groups to prioritise allocation of study records to the research assistants for screening. While screening continued at the title and abstract stage, we obtained the full text version of studies initially included and screened them against the complete set of inclusion criteria. We also added results of the grey literature screening to EPPI and screened those all at full text. All of the full text screening was completed by two coders, with one of them being a

core member of the review team and the other being a research assistant. When disagreements on inclusion/exclusion arose, another core team member oversaw the reconciliation discussion.

3.3.2 Data extraction and management

We extracted the following descriptive, methodological qualitative and quantitative data from each included study, using a standardised data extraction form (form provided in Online appendix D.2 Quantitative data extraction):

- Descriptive data including authors, publication date and status, country, type of intervention, population and context;
- Methodological information on study design, analysis method, type of comparison and external validity;
- Quantitative data for outcome measures, including outcome descriptive information, sample size in each intervention group, outcome means and standard deviations, test statistics (e.g. t-test, F-test, p-values, 95% confidence intervals) and cost data; and
- Information on intervention design, including how the intervention incorporates the different pathways of social cohesion (knowledge, attitudes, behaviours and norms), implementation fidelity and adherence to the planned interventions, contextual factors and programme mechanisms.

We used Excel to record the descriptive, methodological and qualitative data. We also used Excel to extract quantitative data for outcomes analysis. Descriptive and qualitative data were double-coded by two reviewers, who then reconciled their answers. A third reviewer undertook spot-checks for quality assurance. The quantitative data were also double-coded by two reviewers who then reconciled their answers.

3.3.3 Assessment of risk of bias in included studies

We assessed the risk of bias in the included studies by drawing on the signalling questions in the 3ie risk of bias tool, which covers both internal validity and statistical conclusion validity of experimental and quasi-experimental designs (Hombrados and Waddington 2012) and the bias domains and extensions to Cochrane's ROBINS-I tool and RoB2.0 (Sterne et al. 2016; Higgins et al. 2016). The risk of bias assessment helps us to determine the extent to which the findings in each study are reliable. The complete risk of bias tool can be found in Online appendix D.3 Risk of Bias Assessments (RoB).

We assessed the risk of bias for each included study according to the domains listed below. For each domain, we assessed whether studies sufficiently addressed sub-questions corresponding to each risk of bias domain in order to be considered free from the source of bias, coding each study as 'Yes', 'Probably Yes', 'Probably No', 'No' or 'No Information'.

Risk of bias domains assessed for all study designs:

- Selection bias: factors relating to differential selection into or out of (attrition) the study;
- Confounding/group equivalence: factors relating to comparability of groups at baseline;
- Deviations from intended interventions: factors relating to spillover of the intervention to the control group, crossovers of participants from control and treatment groups and contamination from other related interventions;

- Performance bias: factors related to the act of monitoring affecting the performance of participants in treatment and control in different ways (including Hawthorne and John Henry effects)[5];
- Outcome measurement bias: factors related to outcome assessor impartiality, participant recall, participant incentives or social desirability bias – they may, for example, have employed list experiments to determine if people are responding truthfully (Blair and Imai 2012) or triangulation to determine whether two or more methodologies that rely on different assumptions produce similar results (Munafò and Smith 2018); and
- Reporting bias: factors related to analyst impartiality, use of appropriate analytical methods and selective reporting of outcomes – the latter can be partially assessed through the use of pre-registration and the publication of pre-analysis plans (PAPs) on platforms such as ClinicalTrials.gov, the Open Science Foundation, the American Economic Association (AEA) RCT Registry or 3ie's Registry for International Development Impact Evaluations (RIDIE).

For experimental methods, we assessed an additional two risks:
- Assignment mechanism: factors related to whether the assignment mechanism was plausibly random; and
- Unit of analysis: factors related to whether adjustments were made to ensure that the unit of analysis takes into account the unit of randomisation by including clustered standard errors as appropriate.

The risk of bias coding was conducted by two independent reviewers who then reconciled their coding. If there were any disagreements, they were resolved through discussion with a third reviewer.

We used the results of the risk of bias assessments to produce an overall rating for each study as 'High risk of bias', 'Some concerns' and 'Low risk of bias', drawing on the decision rules in RoB2.0 (Higgins et al. 2016), and rated studies as follows:
- 'High risk of bias': if any of the bias domains were assessed as 'No' or 'Probably No';
- 'Some concerns': if one or several domains were assessed as 'Unclear' and none were 'No' or 'Probably No'; and
- 'Low risk of bias': if all of the bias domains were assessed as 'Yes' or 'Probably Yes'.

3.3.4 Synthesis of intervention effects (RQ1 and RQ2)
Measures of treatment effect
An effect size expresses the magnitude or strength of the relationship of interest (Borenstein 2009; Valentine and Aloe 2016). In this review, we are interested in the effect of various interventions on measures of social cohesion; the effect size will allow us to determine the magnitude of the difference that the intervention made on social cohesion.

---

[5] Hawthorne effects are related to the act of being observed; John Henry effects are related to the act of being compared.

We extracted data from each individual study to calculate standardised effect sizes for cross-study comparison where possible. For continuous outcomes comparing group means in a treatment and control group, we calculated the standardised mean difference (SMD), or Cohen's *d*, its variance and standard error using formulae provided in Borenstein et al. (2009). An SMD is the difference in means between the treatment and control groups divided by the pooled standard deviation of the outcome measure. Cohen's *d* can be biased in cases where sample sizes are small. Therefore, in all cases we adjusted *d* using Hedges' method, adjusting Cohen's *d* to Hedges' *g* using the following formula (Ellis 2010):

$$g \cong d(1 - \frac{3}{4(n_1 + n_2) - 9})$$

Different formulas for effect size calculations were used depending on the type of data provided in included studies. For example, for studies reporting means (*x*) and pooled standard deviation (*SD*) for treatment (*T*) and control or comparison (*C*) at follow-up only, we used:

$$d = \frac{x_{Tp+1} - x_{Cp+1}}{SD}$$

If the study did not report the pooled standard deviation, we calculated it using the following formula:

$$SD_{p+1} = \sqrt{\frac{(n_{Tp+1} - 1)SD_{Tp+1}^2 + (n_{Cp+1} - 1)SD_{Cp+1}^2}{n_{Tp+1} + n_{Cp+1} - 2}}$$

Where the intervention is expected to change the standard deviation of the outcome variable, we used the standard deviation of the control group only.

For studies reporting means ($\underline{X}$) and standard deviations (*SD*) for treatment and control or comparison groups at baseline (*p*) and follow up (*p*+1), we used:

$$d = \frac{\Delta \underline{X}_{p+1} - \Delta \underline{X}_p}{SD_{p+1}}$$

For studies reporting mean differences ($\Delta \underline{X}$) between treatment and control and standard deviation (*SD*) at follow-up (*p*+1), we used:

$$d = \frac{\Delta \underline{X}_{p+1}}{SD_{p+1}} = \frac{\underline{X}_{Tp+1} - \underline{X}_{Cp+1}}{SD_{p+1}}$$

For studies reporting mean differences between treatment and control, standard error (*SE*) and sample size (*n*), we used:

$$d = \frac{\Delta \underline{X}_{p+1}}{SE\sqrt{n}}$$

For studies reporting regression results, we followed the approach suggested by Keef & Roberts (2004), using the regression coefficient and the pooled standard deviation of the outcome. Where the pooled standard deviation of the outcome was unavailable, we used

regression coefficients and standard errors or t-statistics to do the following, where sample size information was available in each group:

$$d = t \sqrt{\frac{1}{n_T} + \frac{1}{n_C}}$$

where n denotes the sample size of treatment group and control. We used the following where total sample size information (n) was available only (as suggested in Snilstveit et al. 2019):

$$d = \frac{2t}{\sqrt{N}} \quad Var_d = \frac{4}{N} + \frac{d^2}{4N}$$

We calculated the t-statistic (t) by dividing the coefficient by the standard error. If the authors only reported confidence intervals and no standard error, we calculated the standard error from the confidence intervals. If the study did not report the standard error but did report $t,$ we extracted and used t as reported by the authors. In cases in which significance levels were reported rather than t or $SE(b)$, then t was to be imputed as follows:
Prob > 0.1: t = 0.5
0.1 ≥ Prob > 0.05: t = 1.8
0.05 ≥ Prob > 0.01: t = 2.4
0.01 ≥ Prob: t = 2.8.

Where outcomes were reported in proportions of individuals, we calculated the Cox-transformed log odds ratio effect size (Sanchez-Meca et al. 2003):

$$d = \frac{ln(OR)}{1.65}$$

where OR is the odds ratio calculated from the two-by-two frequency table.

Where outcomes were reported based on proportions of events or days, we used the standardised proportion difference effect size:

$$d = \frac{p_T - p_C}{SD(p)}$$

where p_T is the proportion in the treatment group and p_C the proportion in the comparison group, and the denominator is given by:

$$SD(p) = \sqrt{p\,(1-p)}$$

where p is the weighted average of p_C and p_T:

$$p = \frac{n_T\,p_T + n_C\,p_C}{n_T + n_C}$$

For each of the effect sizes we extracted, we reversed the sign where applicable to make sure that the effect sizes were positive if they showed a change in a direction that increased social cohesion. For example, if the original effect size measured 'distrust', an

increase in that effect size would represent a decrease in social cohesion. As such, we reversed the sign so that if that measure of distrust goes down, it is measured as a positive change in our meta-analysis.

Unit of analysis issues

Unit of analysis errors can arise when the unit of allocation of a treatment is different from the unit of analysis of the study, and this is not accounted for in the analysis. We have assessed studies for unit of analysis errors (The Campbell Collaboration 2014). If unit of analysis errors exist, we corrected for this by adjusting the standard errors (Higgins and Green 2011; Waddington et al. 2012):

$$SE(d)' = SE(d) * \sqrt{1 + (m-1)c}$$

where *m* is the average number of observations per cluster and *c* is the intra-cluster correlation coefficient. Where included studies use robust Huber-White standard errors to correct for clustering, we calculated the standard error of *d* by dividing *d* by the t-statistic on the coefficient of interest.

Criteria for determination of independent findings

When extracting the effect sizes, we wanted to avoid double-counting of evidence from different papers that focus on the same study. As such, we linked papers prior to analysis. We extracted data from what we deemed the 'main paper'. Where possible, we designated as the main paper the one that had been peer-reviewed. Otherwise, we chose the most recent version. After extracting the effect sizes from the main paper, we only extracted data on outcome measures and samples from the linked papers that did not appear in the main paper.

Where studies reported data for the same outcome for an intervention over multiple time periods, we extracted data for each of the reported time periods. We then made decisions about which time point to include in each meta-analysis based on comparability with outcomes included from other studies. Where authors reported the same outcome using more than one analytical model, we extracted the data from the authors' preferred model specification. Where the authors did not specify a preference, we extracted data from the model with the most controls.

Where studies reported an index of different outcomes and the effects on the individual factors that comprised the index, we extracted data for the overall index measure. Where studies reported outcomes or evidence according to subgroups of participants, we extracted data on both the full sample (where possible) and on the individual subgroups, in order to answer Review Question 3 regarding differential effects by population type.

If a study reported outcomes related to multiple treatment arms and only one comparison group, we extracted the data and estimated an effect size for each of the treatment arms. We then chose to include the effect size estimate from the treatment arm that tested an intervention that most commonly resembled the other interventions included in the relevant meta-analysis.

Dealing with missing data

In cases of relevant missing or incomplete data, we sought the data within published replication files or contacted study authors to obtain the required information. In cases where we were unable to obtain the necessary data, we have reported the characteristics of the study but do not include it in the meta-analysis or reporting of effect sizes due to missing data.

Data synthesis

Once we had identified all included studies, we mapped out all interventions, social cohesion pathways and outcome measures provided in the included studies.

We synthesised outcomes based on the five different dimensions of intergroup social cohesion: trust, sense of belonging, willingness to participate, willingness to help, and acceptance of diversity. The intervention criteria for the review was broad, which has led us to include studies with a relatively diverse set of interventions. We undertook a detailed analysis of each included intervention's components to identify similarities and differences across the interventions. From this matrix, we identified five distinct groups of sufficiently similar programmes, which we then used to structure the synthesis of results in Section 4.4. Where we had three or more studies assessing the same outcome construct, we conducted meta-analysis using the methodology described above. Where we did not have enough studies presenting comparable outcome constructs, we presented a narrative synthesis of individual standardised effect sizes.

We used the *metafor* package in R software for each of the meta-analyses we conducted (R Development Core Team 2008; Viechtbauer 2010). Due to the relatively small number of studies included in our analyses, we were unable to use the robust variance estimation method of meta-analysis that would have allowed us to use multiple effect sizes from the same paper in the models (Hedges et al. 2010). Instead, we used the restricted maximum likelihood model for the majority of our analysis. This required us to choose a single effect size from each study included in each individual meta-analysis. We followed a clear set of criteria to make those choices without taking the size or direction of effect sizes into account. The decisions made for each analysis can be found in Online appendix C.3.

Assessment of heterogeneity

We assessed heterogeneity in each meta-analysis by calculating the Q-statistic, I^2, and Tau^2 to provide an estimate of the amount of variability in the distribution of the true effect sizes (Borenstein et al. 2009), and complemented this with visual inspections of forest plots.

Assessment of reporting biases

We attempted to reduce publication bias by searching for and including unpublished studies in the review. We also undertook exploratory tests for the presence of publication bias using contour-enhanced funnel graphs (Peters et al. 2008) and statistical tests (Egger et al. 1997).

Sensitivity analysis

We conducted a sensitivity analysis to assess whether the results of the meta-analysis are sensitive to the removal of any single study. We did this by removing studies from the meta-analysis one by one and assessing changes in results.

3.3.5 Subgroup analysis and investigation of heterogeneity (RQ3)

To the extent possible, we collected data on differential effects and experiences for sub-populations and coded information according to the PROGRESS-plus criteria. These criteria include place of residence, race/ethnicity, occupation, gender, religion, education, socio-economic status, social capital, age, disability and sexual orientation (O'Neil et al. 2014). Unfortunately, the data available in the included studies were too heterogeneous to conduct statistical subgroup analysis. However, where subgroup analyses were conducted by individual studies, we report differential effects for different population groups in the syntheses in Section 4.4.

3.3.6 Barriers and facilitators analysis (RQ4)

We collected substantial descriptive information from included studies about the interventions, their design, implementation process, and any identified barriers and facilitators along the causal chains. The data extraction and analysis were based on framework synthesis approaches, wherein a framework of potential characteristics was initially developed based on factors identified as important in the literature, and data were then extracted from studies according to the framework. The list of factors was designed to focus on those that may be relevant across contexts. For example, we did not include factors such as duration of displacement that have been identified as important for interventions targeting relationships between host communities and displaced populations, but we did extract data on the relationship between groups at baseline (De Berry and Rogers, 2018). Not all studies report information that informs every factor covered. Due to the heterogeneity of interventions included in the review, we erred on the side of inclusivity in order to capture all potentially relevant factors. These include:

- Intervention characteristics, including duration, relationship between implementers and target population;
- Target population characteristics, including baseline relationships between groups;
- Context characteristics, including characteristics of conflict and violence;
- The use and components of theories of change;
- Intervention objectives, including changing knowledge, beliefs, attitudes, behaviours or some combination thereof;
- The extent to which members of both targeted groups are engaged (equally or primarily one group);
- Alignment between intervention activities, identified drivers of tensions, and targeted levels of change; and
- Specific dimensions of social cohesion to which the intervention aims to contribute.

The full descriptive coding tool is included in Online appendix D.1. In order to identify the barriers and facilitators that influenced intervention effects from within all of the challenges faced during implementation or best practices identified in included studies, we undertook a three-part process of qualitative synthesis that aimed to explain the quantitative findings using qualitative data on implementation processes, barriers and facilitators alongside descriptive data on intervention and implementation characteristics. First, we reviewed the included impact evaluation papers and additional identified documents related to included studies, such as implementation reports, case studies and baseline documents, to identify any potential barrier or facilitator during implementation

identified by authors or implementers. We kept note of how these factors were identified, including from qualitative research, implementer reports or reflections, or evaluator interpretations. We then placed these alongside the findings from the quantitative synthesis and the descriptive information of intervention and implementation characteristics, and undertook both horizontal and vertical case synthesis to identify trends in the data. From this, we identified a series of factors that may be associated with better or worse intervention effects.

3.3.7 Cost-effectiveness analysis (RQ5)

Analysis under Review Question 5 aimed to address unit cost, cost-efficiency, cost-effectiveness or benefit–cost evidence on interventions in particular contexts, where available. We planned to incorporate cost data from impact evaluations and additional programmatic documents related to included studies, such as economic evaluations of included programmes, and draw on standard approaches to synthesis of economic appraisal evidence (Shemilt et al. 2011). We expected that included studies would incorporate unit and total costs to implementers and participants (and non-participants, as relevant). Unfortunately, given minimal cost reporting in included studies, analysis of RQ5 has been limited to narrative synthesis of the different types of cost data reported in included studies and methodologies used where reported.

4. Results

4.1 Results of the search

The results of the search and screening process can be seen in the PRISMA flow diagram in Figure 2, following guidance for reporting in systematic reviews (Moher et al. 2009). The initial academic search identified 76,148 records. Hand searches of relevant grey literature sites identified a further 409 studies, while forwards and backwards citation searches of included studies identified 225 additional potentially relevant studies. Following removal of duplicates, a total of 71,077 studies were left for screening at title and abstract. Title and abstract screening identified 875 potentially relevant records for inclusion in the review, which were then screened at full text. The full-text screening ultimately identified 37 impact evaluation papers, corresponding to 24 studies of 31 unique intervention arms for inclusion in this review. In addition, five ongoing studies were also identified that appear likely to meet inclusion criteria. Although not included in the review, these are listed in Section 9.3.

We include full explanations and a table summarising the reasons for which studies were excluded in Section 9.2. There were often multiple reasons for which a study could be excluded, but once we had identified one failed criterion, we did not search for all possible reasons. We most commonly excluded studies during full-text screening because they did not focus on an intervention aiming to build intergroup social cohesion (n = 277). We also frequently excluded studies that failed to meet our study design inclusion criteria, particularly such as not measuring intervention effects on outcomes (n = 103) or by not having a comparison group (n = 119). During title and abstract screening, the most common reasons for exclusion included studies that did not relate to an intervention, including many primary studies focused on understanding – but not trying to change – intergroup cohesion (n = 48,176), and studies where the interventions did not relate to social cohesion (n = 10,635).

Figure 2: PRISMA flow diagram of systematic search and screening results

76,148 records identified through academic database searching	636 records identified through grey literature search, backwards and forwards citation tracking

71,079 records screened at title and abstract (after 5,705 duplicates removed) → 70,202 records excluded

877 articles screened at full text

REASONS FOR EXCLUSION
- Published pre-2000: 2
- High-income country: 28
- Intervention not relevant: 19
- No intervention: 23
- Not intergroup social cohesion: 277
- Not a primary study: 45
- Does not address impact: 103
- Not fragile context: 26
- Qualitative theory-based evaluation: 68
- Comparison conditions: 119
- Does not meet study design criteria: 41
- No core social cohesion outcomes: 6
- Lab in the field only: 12
- Unclear – no access to full text: 13
- Duplicate paper of included study: 53

37 included impact evaluation papers, corresponding to 24 studies of 31 unique interventions / arms

5 ongoing studies

4.2 Descriptive analysis of included studies

This section gives an overview of the characteristics of included studies, including geographic coverage, target populations, intervention groups and study design characteristics. A comprehensive overview of included studies is presented in Section 9.1.

4.2.1 Intervention groups

In this section, we describe the interventions, including their components, duration and target populations, addressed in the included studies (Table 1). We also discuss the ways in which the interventions aimed to build intergroup social cohesion, including both the primary mechanisms they sought to trigger and the core dimensions of social cohesion they aimed to influence. These core dimensions include trust, sense of belonging, willingness to participate, willingness to help, and acceptance of diversity.

Table 1: Intervention groups

Intervention group	Core components	Included studies
School-based peace education	Peace education	• Aladysheva et al. (2017) – Kyrgyzstan • Alan et al. (2020) – Turkey • Biton and Salomon (2006) – Palestine • Cleven (2020) – Bosnia and Herzegovina
Collaborative contact	Collaborative contact	• Alaref et al. (2019) – Lebanon • Mousa (2019) – Iraq • Okunogbe (2018) – Nigeria • Scacco and Warren (2018) – Nigeria
Intergroup dialogues	Intergroup contact + facilitated dialogue sessions	• Cilliers et al. (2018) – Sierra Leone • Hartman et al. (2018) – Liberia • Lonergan (2016) – Sri Lanka • Rime et al. (2011) – Rwanda • Schiller (2012) – Indonesia • Svensson and Brouneus (2013) – Ethiopia
Workshop-based peace education with intergroup contact and economic support (workshops-contact-econ)	Peace education + intergroup contact + economic support	• Causal Design (2016) – Nigeria • Ferguson (2019) – Jordan • Finkel et al. (2018) – Burkina Faso, Chad and Niger • IMPAQ International (2017) – Bosnia and Herzegovina • Dawop et al. (2019) – Nigeria
Media for peace	Radio programmes or media campaigns	• Bilali et al. (2016) – Burundi • Bilali and Vollhardt (2015) – DRC • Bilali and Vollhardt (2013) – Rwanda • Paluck (2009) – Rwanda • Vicente and Vilela (2020) – Mozambique

School-based peace education

We identified four studies of interventions providing school-based peace education (Alan et al. 2020; Aladysheva et al. 2017; Biton and Salomon 2006; Cleven 2020). This group of interventions includes studies based in primary and secondary schools, typically held over one school year. Students were taught sociocognitive and psychosocial skills seen as critical to building bridging ties across differences and transforming conflicts in a peaceful way.

All four of the studies in this group targeted children and youth in schools. The interventions were implemented by the students' teachers. In all four cases, the facilitators received specific training and support to build specific skills designed to help 'teach' social cohesion.

Classroom-based peace education interventions tend to adopt a 'head-first' approach to strengthening social cohesion (McCauley 2002). Interventions in this group focus primarily on strengthening key social and emotional skills among participants, including

empathy, respect, conflict resolution and tolerance for diversity. All four interventions were designed to create a safe space in which participants from both groups could build empathy with each other and reduce anxiety around interactions with individuals from the other groups. However, in one case (Biton and Salomon 2006), outbreaks of violence prevented the contact portion of the intervention from being implemented.

The primary mechanisms targeted by these interventions are those focused on 'seeing the other' as a means of building trust and a shared sense of belonging (Blewitt n.d.). By building awareness of and comfort with multiple perspectives, these interventions aimed to strengthen acceptance of diversity. In one case, the intervention studied in Aladysheva et al. (2017) further aimed to build participants' willingness to help by supporting community projects that students from both ethnic groups worked together to design and implement.

Collaborative contact

We identified four studies, reported in six papers, that evaluated interventions that introduced collaborative forms of contact between different groups, yet did not explicitly address the nature of the conflict, tensions or even the purpose of the activities (Alaref et al. 2019; Okunogbe 2018; Mousa 2019; Scacco and Warren 2018). In this sense, participants were 'blind' to the aim to build social cohesion. Participants in these interventions attended computer literacy training (Scacco and Warren 2018), completed a mandatory year of civil service (Okunogbe 2018), participated in volunteering activities to acquire soft skills (Alaref et al. 2019) or signed up to play soccer (Mousa 2019). Therefore, participants accepted the intergroup element (i.e. the presence of participants from the other group) without being told that it was the primary aim of the activity.

Three of the interventions (Alaref et al. 2019; Mousa 2019; Scacco and Warren 2018) were relatively short term in nature, lasting from two and a half to four months. Two only targeted young men, most of whom were between 18 and 25 years of age (Mousa 2019; Scacco and Warren 2018).

Unlike many of the other interventions in this study, these interventions implemented a 'feet-first' approach to change (McCauley 2002), by bringing participants together and getting them to work with each other towards common goals. In this way, they first aimed to build social cohesion by triggering mechanisms of 'collaborating with the other', which were expected to, in turn, strengthen willingness to participate and help members of the other group. From there, the programmes hoped to build trust, sense of belonging, and acceptance of diversity.

Intergroup dialogues

We identified six studies, reported in 13 papers, that aimed to facilitate intergroup dialogues (Hartman et al. 2018; Cilliers et al. 2018; Lonergan 2016; Rime et al. 2011; Schiller 2012; Svensson and Brouneus 2013). The interventions of these studies aimed to build intergroup cohesion and sustainable peace by bringing together people from groups in conflict and facilitating constructive dialogues on their differences and causes of the conflict. Each intervention applied a 'head-first' approach (McCauley 2002) focused on changing intergroup knowledge and understanding through approaches grounded in 'talking with the other' (Blewitt n.d.). These interventions ranged in duration but were relatively short. One consisted of a few days of intensive interactions (Schiller 2012).

Several operated for one to three months (Cilliers et al. 2018; Hartman et al. 2018; Lonergan 2016; Rime et al. 2011). One ran for six months (Svensson and Brouneus 2013).

In addition to strengthening intergroup understanding, these interventions further aimed to resolve conflicts or heal emotional wounds from violent conflicts between participants. Two studies focused on resolving conflicts between participants and institutionalising non-violent dispute resolution mechanisms (Hartman et al. 2018; Svensson and Brouneus 2013). In Liberia, the intervention aimed to build skills and norms for community-based alternative dispute resolution. The goal was to strengthen communities' abilities to resolve longstanding conflicts, which were often over land and related to the aftermath of the civil war and displacement (Hartman et al. 2018). In Ethiopia, the intervention engaged university students on a campus wracked by violent conflict across social cleavages in sustained dialogue sessions that aimed to address grievances (Svensson and Brouneus 2013). These interventions primarily aimed to build trust and willingness to participate.

Two further studies focused on reconciling members of different groups following conflict (Lonergan 2016; Schiller 2012). In Sri Lanka, the intervention brought together student representatives from different ethnic backgrounds to try to build intergroup trust following the civil war (Lonergan 2016). In Indonesia, the intervention brought together young men affiliated with two parties to the conflict to learn conflict resolution skills, build intergroup understanding and serve as ambassadors for peace in their respective communities (Schiller 2012). These interventions primarily aimed to build trust and acceptance of diversity.

Finally, two studies focused on interventions that facilitated communal rituals of recognition and forgiveness between victims and perpetrators (Cilliers et al. 2018; Rime et al. 2011). In Sri Lanka, the intervention incorporated substantial community mobilisation culminating in a communal bonfire ceremony in which victims shared their stories, perpetrators gave further details and asked for forgiveness, followed by a cleansing ceremony to heal the hurt and move forward (Cilliers et al. 2018). The intervention aimed to solidify the reconciliation through the symbolic installation of peace trees and facilitate collaborative intergroup contact through communal farms. In Rwanda, the intervention served as an extension of the formal justice system that focused equally on community reconciliation (Rime et al. 2011). The community-based gacaca courts led by local leaders gave victims an opportunity to share their experiences and ask further detail of imprisoned perpetrators, who in turn could ask for forgiveness and confess in return for lighter sentencing. This intervention thus straddled both vertical and horizontal cohesion, including government-administered access to justice as well as aiming to heal intergroup relationships. These interventions aimed to build trust, sense of belonging, and acceptance of diversity.

Workshop-based peace education with intergroup contact and economic support
We identified five studies, reported in six papers, of interventions that combined workshop-based peace education with intergroup contact and economic support (IMPAQ International 2017; Causal Design 2016; Ferguson 2019; Finkel et al. 2018; Dawop et al. 2019). The discrepancy in numbers is due to one paper, by Finkel and colleagues (2018), reporting on an intervention implemented and evaluated separately in three countries.

The interventions in this group (workshops-contact-econ) aimed to build intergroup social cohesion, both through activities aimed directly at addressing the relationships between the groups and through components aimed at mitigating some of the underlying drivers of tensions between the groups. These interventions include three core components: workshops where participants were taught conflict resolution skills; opportunities for intergroup contact; and funding for small to medium-scale projects. The duration of interventions in this grouping was typically longer than those described in other groups in this section, lasting between three and a half to five years.

The interventions were targeted primarily at adults, with two exceptions. Finkel and colleagues (2018) included a substantial component of the intervention aimed directly at youth, and IMPAQ International (2017) also included youths. In all five studies, the interventions included activities directed towards community leaders, including political, traditional and religious authority figures.

In four cases, these projects were designed not only to contribute to the well-being of communities, but also as opportunities for collaborative intergroup contact, as participants were tasked with identifying and implementing the projects together. In the fifth case (Finkel et al. 2018), contact was introduced through interfaith dialogues and peace conferences, while the economic support was provided in the form of vocational training and life-skills training for youth.

While the interventions shared the same three core components of conflict resolution training, intergroup contact and economic support, in three cases they also included additional components. Civic education was also included in one study (Finkel et al. 2018); media for peace was included in two studies (Causal Design 2016; Finkel et al. 2018); and conflict prevention forums was included in two studies (Causal Design 2016; Dawop et al. 2019).

There was greater variation among the interventions in this group in terms of the mechanisms they aimed to trigger to build social cohesion as compared with the previous three intervention categories. Since the interventions targeted community leaders, they incorporated diverse strategies for different target populations. The intergroup contact was typically focused on bringing diverse leaders together, and efforts aimed at triggering mechanisms of 'seeing the other', 'talking with the other' and 'collaborating with the other' were typically targeted towards these individuals (Blewitt n.d.). Three interventions also included opportunities for the general adult population to come together, targeting a mechanism of 'meeting the other' through fun community activities to which individuals from all targeted groups were encouraged to attend (Ferguson 2019; Finkel et al. 2018; IMPAQ International 2017). The dimensions of social cohesion that these interventions most directly aimed to influence were acceptance of diversity and willingness to participate, although three interventions also explicitly worked to build trust (Causal Design 2016; IMPAQ International 2017; Dawop et al. 2019).

Media for peace
We identified five studies, reported in eight papers, evaluating four different radio programmes designed to build social cohesion between groups in fragile contexts (Bilali et al. 2016; Bilali and Vollhardt 2015; Bilali and Vollhardt 2013; Paluck 2009; Vicente and

Vilela 2020).[6] One study evaluated a media campaign that aimed to spread messages of tolerance in order to reduce violence caused by religious extremism (Vicente and Vilela 2020). The other three interventions in this group aimed to teach peace-positive values through locally relatable characters and storylines, which were broadcast over multiple years. Paluck (2009) evaluates the effects of listening to one year of the drama, while the other three evaluate a mixture of short- and long-term effects depending on when participants happened to start listening to the drama. Two studies included only individuals who had listened to the drama for at least six months in the treatment group (Bilali et al. 2016; Bilali and Vollhardt 2013), while the third did not report the length for which participants needed to have reported listening to the drama to be classed as 'listeners' (Bilali and Vollhardt 2015). The maximum duration anyone could have been listening to these dramas is roughly six years.[7]

The dramas are designed to include characters who reflect members of different groups in each country, though ethnicities are fictionalised, and are targeted at an adult audience. These interventions aim to present positive role models in the form of sympathetic characters from both sides of the conflict, encouraging listeners to learn to relate to people from diverse backgrounds and with diverse experiences.

In this way, the primary mechanisms through which the interventions aim to build intergroup social cohesion is through building participants' familiarity and comfort with 'seeing the other' (Blewitt n.d.), thereby changing attitudes to build tolerance for diversity.

4.2.2 Social cohesion outcomes reported in included studies
This section presents an overview of the different outcome measures reported in included studies. We only include outcomes for which we were ultimately able to extract effect size data (reported in Section 4.4); as a result, not all outcomes are included in this analysis.[8]

As we define in Section 3.1 above, we included studies that measured one or more core dimension of social cohesion. Table 2 shows the high heterogeneity of outcome measures across studies across and within intervention groups. Overall, each dimension of social cohesion was measured by 9 to 17 studies, suggesting that broadly the five dimensions we identified are indeed important to study authors and accurately capture the measures of social cohesion. The most frequently measured dimension of social cohesion was acceptance of diversity, measured in almost three quarters (n = 17) of included studies, which is understandable given the focus of this review on intergroup cohesion. The least frequently measured dimension was willingness to help, which was measured only in nine studies overall, although including in three of four collaborative contact studies. While about half of included studies reported at least one measure of sense of belonging, measures of this dimension were the least frequently reported of all

---

[6] Bilali and Vollhardt (2013) and Paluck (2009) both evaluate the impact of the Musekeweya or 'New Dawn' radio programme in Rwanda, at different follow-up times and in separate studies.
[7] None of the three longer-term studies report when data were collected, which makes it difficult to estimate exposure times.
[8] Unfortunately, despite efforts to source missing data from authors, we were not able to calculate effect sizes for all outcomes reported in all studies. For details of the information needed to calculate effect sizes, see Section 3.3.4. The only study for which no effect sizes could be calculated was Rime et al. (2011), which did not report standard errors.

social cohesion outcomes. Across all social cohesion outcomes reported across all studies, only 8 per cent (n = 15) measured sense of belonging. By contrast, all other dimensions were measured at relatively equal proportions, ranging from 20 to 27 per cent (n = 36–50). This data show that studies typically only included one or two measures of sense of belonging, but multiple measures of other social cohesion dimensions. It is unclear whether this is due to measurement challenges or perceptions among authors that sense of belonging is less relevant to intergroup social cohesion.

There is seemingly the most overlap among collaborative contact interventions, wherein all four studies reported measures of acceptance of diversity and three studies measured outcomes of willingness to participate and help as well. However, as described in more detail in Section 4.4.3, many of the underlying measures were substantially different from each other, which prevented useful meta-analysis. In terms of similarity of measures, the interventions in the workshops-contact-econ group were most similar, likely because a single organisation, Mercy Corps, implemented three of the included interventions in this group. Among media for peace interventions, as noted earlier, the radio drama programmes on which four studies were based were implemented by a single organisation as well, Radio La Benevolencija. While the table below suggests less overlap within these studies, the majority of outcomes reported related to measures of acceptance of diversity, and we were ultimately able to conduct four meta-analyses of outcomes related to this dimension.

Table 2: Frequency of studies reporting at least one outcome per social cohesion dimension, by intervention group

	Total n studies	Trust	Sense of belonging	Willingness to participate	Willingness to help	Acceptance of diversity
All studies	23*	61%	52%	52%	39%	74%
School-based	4	3	2	2	2	3
Collaborative contact	4	1	2	3	3	4
Intergroup dialogues	5	3	3	1	1	4
Workshops-contact-econ	5	4	3	4	1	2
Media for peace	5	3	2	2	2	4

Notes: Shading corresponds to the percentage, with 0–40% the lightest shade, 41–60% the medium shade, and 61–100% the darkest shade.
* The total number of studies is smaller here to account for the fact that we were unable to extract effect sizes for one study (Rime et al. 2011).

How outcomes were measured

In terms of data collection, three quarters of social cohesion outcomes were based on self-reported perceptions (n = 136).[9] A further 18 per cent were based on behaviours within the contexts of 'explicit' games (games in which the participants were aware they were playing a game) (n = 33), and seven per cent were measured by behaviours in 'natural' games (games in which the participants were not expected to be aware that they were participating in an exercise that was part of the study) (n = 13).

Drawing on our framework for how intergroup cohesion interventions may work (see Section 1.5), we further analysed the extent to which studies reported outcomes along the causal pathways we identified, which include knowledge, attitudes, behaviours and norms. Of all social cohesion outcomes reported by all studies, 59 per cent (n = 107) were measures of attitude changes. A further 38 per cent (n = 70) were measures of behaviour change. Only 3 per cent of social cohesion outcomes were measures of knowledge changes (n = 5), which is striking given how many of the interventions aimed to build participants' knowledge. We identified only eight studies measuring knowledge outcomes of intermediate steps that may contribute to social cohesion, including measures of intergroup sociocultural awareness and understanding of concepts related to peace and conflict. We did not identify any measures of changes in social norms. Although some studies made efforts to understand norm changes, they measured outcomes related to individual-level changes, such as by asking individuals how they would respond to a situation and interpreting aggregate measures as indicative of shifts in norms.

Trust: There are many different ways to measure trust. Most trust outcome measures (75 per cent, n = 27) estimated respondents' feelings of trust towards other people, while one measured people's perceptions of others' levels of trust. We also identified four outcomes that measured whether respondents acted in a trustworthy way, typically via observation of behaviours in lab-in-the-field games. A further four outcome measures were phrased in a negative way to capture respondents' sense of mistrust towards others.

Sense of belonging: Among the 15 outcomes measuring sense of belonging, seven were based on measures of a shared sense of belonging, typically by measuring respondents' agreement with statements phrased in an inclusive way that prioritised a national identity over a smaller group-based identity. Only three outcomes measuring sense of belonging were phrased in an exclusive way, measuring alignment with entrenched senses of separate identities. Finally, five outcomes measured respondents' perceptions of division and tension between groups in their communities.

Willingness to participate: Just over half of outcomes measuring respondents' willingness to participate (n = 19) reported active measures of self-reported or observed intergroup participation (n = 15) or general community participation (n = 4). A further third of willingness to participate measures (n = 15) were related to respondents' hypothetical openness to participation, all but three of which were explicitly presented as intergroup interactions. Finally, we identified a single outcome of respondents' anxiety around intergroup participation, and a single outcome of the perceived value of intergroup participation.

---

[9] This measure is based on outcomes reported for whole samples only to avoid double-counting outcomes of disaggregated effects.

Willingness to help: Roughly half of reported outcomes measuring willingness to help were measures of a hypothetical openness to providing help (n = 24). Note, however, that we include explicit game-based measures of altruism and in-game donations within this count, to differentiate from outcomes measuring actions of intergroup help in participants' regular lives. There were only six outcomes reported among included studies that related to active measures of giving or receiving help, such as a respondent's experience receiving help from classmates. Ten outcomes measured behaviours of refusing to help, all of which were measured through behaviours in explicit games. Finally, five outcomes related to perceptions of the value of providing help through volunteering.

Acceptance of diversity: As noted, authors reported outcomes measuring acceptance of diversity more frequently than measures of other social cohesion dimensions. Over a third of measures of acceptance of diversity (n = 19) assessed changes in respondents' tolerance or intolerance for diversity within their lives, such as measures of acceptance for intergroup marriage or measures of intergroup friendships or feelings about these. This was followed by measures of respondents' bias towards others (n = 15), such as positive or negative traits associated with the other group. Ten outcomes measured perceptions of inclusive or exclusive victimhood – the extent to which respondents felt that the other group had not suffered the way the the respondent's group had suffered, as compared with the feeling that all groups had suffered from conflict. Finally, six outcomes measured awareness, acknowledgement, rejection or ignorance of differing perspectives, particularly in a historical sense.

4.2.3 Geographic coverage
Figure 3: Geographic distribution of included studies
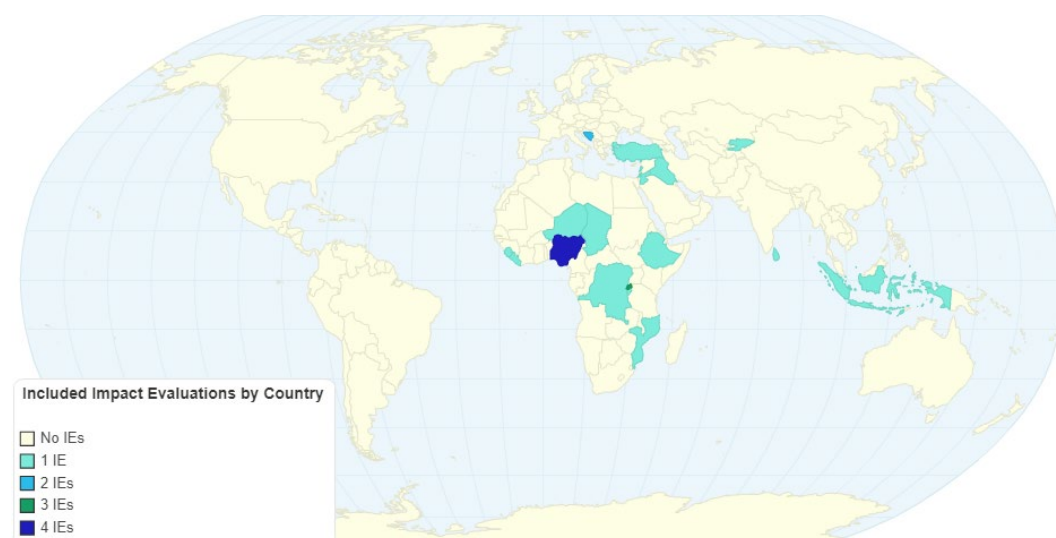
Created with chartsbin.com

The map in Figure 3 shows the geographical distribution of the studies included in the review. The overwhelming majority – 62 per cent – of studies took place in Sub-Saharan Africa (n = 16). Within the region, most studies took place in West Africa, but there was a clear spread across the continent. Four studies took place in Nigeria (Dawop et al. 2019; Okunogbe 2018; Scacco and Warren 2018; Causal Design 2016), and three in Rwanda

(Paluck 2009; Bilali and Vollhardt 2013; Rime et al. 2011).[10] The other countries with studies each had one: Burkina Faso, Chad and Niger (Finkel et al. 2018); Burundi (Bilali et al. 2016); the Democratic Republic of the Congo (DRC) (Bilali and Vollhardt 2013); Ethiopia (Svensson and Brouneus 2013); Mozambique (Vicente and Vilela 2020); Liberia (Hartman et al. 2018); and Sierra Leone (Cilliers et al. 2018).

We identified four studies from the Middle East and North Africa: Iraq (Mousa 2019); Jordan (Ferguson 2019); Lebanon (Alaref et al. 2019); and Palestine (Biton and Saloman 2006). From Europe and Central Asia, we identified four further studies: two studies in Bosnia and Herzegovina (Cleven 2020; IMPAQ International 2017); one in Kyrgyzstan (Aladysheva et al. 2017); and one in Turkey (Alan et al. 2020). Finally, we identified one study from East Asia and Pacific, in Indonesia (Schiller 2012).

Table 3: Regions

Region	Frequency (per cent)
East Asia and Pacific	1 (4%)
Europe and Central Asia	4 (15%)
Middle East and North Africa	4 (15%)
South Asia	1 (4%)
Sub-Saharan Africa	16 (62%)

Source: Regions based on World Bank (n.d.).

Table 4: Income categories

Income category	Frequency (per cent)
Low-income countries	12 (46%)
Lower-middle-income countries	8 (31%)
Upper-middle-income countries	6 (23%)

Source: Income categories based on historical classifications for the first year of implementation (World Bank, n.d.).

Table 5: Fragility levels

Country fragility	Frequency (per cent)
≥ 100	3 (12%)
90–99	10 (40%)
80–89	7 (28%)
70–79	5 (20%)

Source: Country fragility data from the FSI (Fund for Peace 2019a), based on the first year for which interventions were implemented. Scores are out of 120 maximum. The FSI dates back to 2006, so for studies where implementation began before 2006, we used the 2006 data where implementation was ongoing in 2006. Note that this table lacks data for one study that was implemented in 2001–2002 in Palestine.

---

[10] Both Paluck (2009) and Bilali and Vollhardt (2013) evaluated the effects of the same radio drama, albeit at different follow-up periods and in different ways.

Tables 3–5 present overview information of the geographic regions, country income categories and country fragility levels of included studies. The studies that emerged in our search suggest that intergroup social cohesion interventions are being studied in especially challenging contexts: more than half of the studies were conducted in countries with scores of more than 90 on the FSI during the first year of implementation (Fund for Peace 2019a). The largest proportion of studies (40 per cent) were undertaken in countries that scored between 90 and 99 on the FSI scale, while three studies had scores equal to or greater than 100. Of the combined 13 studies featuring countries with scores above 90, all but three (Alaref et al. (2019) in Lebanon; Ferguson (2019) in Jordan; Lonergan (2016) in Sri Lanka) were in Sub-Saharan Africa. All other countries outside of the region had scores below 90. This finding is consistent across two different measures of state-level fragility as well – the OECD States of Fragility platform and the INFORM Global Risk Index. Comparative analysis across the three indices is reported in Online appendix B.1.

The majority of Sub-Saharan African countries where interventions took place were low-income economies; only Nigeria was classified as a lower-middle-income economy during the first year in which included studies were implemented. The interventions that took place in the Middle East and North Africa, Europe and Central Asia, and East Asia and Pacific ranged from lower-middle to upper-middle-income countries.

Types and drivers of conflict
The conflict and fragility contexts for the interventions included in this review vary widely. The levels of violence range from recurrent outbursts of violence, such as the Middle Belt of Nigeria (Causal Design 2016; Dawop et al. 2019), to active fighting by rebel groups in the DRC (Bilali and Vollhardt 2015). In some more post-conflict contexts, such as Sri Lanka, ongoing simmering tensions between different groups are seen as being at risk of escalating into violence (Lonergan 2016). In Rwanda, the wounds of genocide have yet to heal (Bilali and Vollhardt 2013; Paluck 2009; Rime et al. 2011). We further include studies evaluating interventions in places where violent conflict has not yet broken out, but is seen as a risk due to heightening levels of intergroup tensions such as between refugee and host communities in south-eastern Turkey (Alan et al. 2020). Particularly in this last case, the instances of violence identified in the study extend beyond physical violence to psychological violence in the form of bullying and social isolation.

Included interventions aimed to build social cohesion across a range of social cleavages. Ethnic differences were the most commonly identified social cleavage across which tensions flared, as 16 studies targeted tensions between ethnic groups.[11] Many studies also looked at tensions across religious identities (n = 10). We identified three studies that looked at tensions between host communities and displaced populations, two where this was the primary source of tensions (Alan et al. 2020; Ferguson 2019), and one where displacement was one of multiple social cleavages across which tensions were identified

---

[11] Note that some studies identified multiple cleavages across which tensions were persistent. Sometimes this was because both ethnic and religious differences separated two groups, while in other cases, there were more than two groups targeted with overlapping identities. We count here all instances in which authors noted a social cleavage targeted by the intervention, and thus the number is greater than the number of included studies.

(Hartman et al. 2018). Two studies looked at interventions bringing together victims and perpetrators of atrocities (Cilliers et al. 2018; Rime et al. 2011), while one study brought together individuals affiliated with different parties to an armed conflict (Schiller 2012). Finally, two studies looked at tensions between settled farmer and pastoralist communities (Dawop et al. 2019; Causal Design 2016).

In many cases, studies identified additional drivers of tensions that went beyond social cleavages. These underlying drivers of tension tended to relate to land access, shared resources and public services. The government was often identified as playing a role in perpetuating low levels of intergroup cohesion, either through policies that exacerbated tensions or through poor governance and service delivery. In three cases, weak institutions or inadequate service delivery were identified as underlying drivers of tensions (Alaref et al. 2019; Ferguson 2019; Hartman et al. 2018). For example, the Jordanian government's decision to implement an IMF-backed austerity measure was seen as exacerbating the Jordanian host community's frustration with the availability and quality of public services, already thinly stretched due to the influx of refugees (Ferguson 2019). In other cases, tensions were seen as stemming from inadequate government policies regarding reconciliation following violent conflict. For example, in Rwanda, the government's efforts to propagate an official historical narrative of the genocide were seen as stifling and preventing effective dialogue between the affected groups (Bilali and Vollhardt 2013).

Populations targeted in interventions
Interventions aiming to improve social cohesion in conflict-affected and fragile contexts may focus on the general population across group lines or on different target populations.

People under the age of 30 were the most frequently targeted population in studies included in this review: out of 24 total studies, 15 (63 per cent) targeted populations under the age of 30, although 5 of these studies were not exclusively targeted to youth. Nine studies included activities specifically targeted to youth (ages 15–30) outside of the education system (Alaref et al. 2019; Finkel et al. 2018; IMPAQ International 2017; Mousa 2019; Okunogbe 2018; Scacco and Warren 2018; Vicente and Vilela 2020). This included all four collaborative contact interventions and three of five workshops-contact-econ interventions. The four school-based interventions all targeted schoolchildren (Alan et al. 2019; Aladysheva et al. 2017; Biton and Salomon 2006; Cleven 2020), and a further two studies targeted university students (Lonergan 2016; Svensson and Brouneus 2013). This focus on youth is not surprising since studies show that large youth populations are associated with various forms of violence (Urdal 2006) and young people are consequently a key target of many post-conflict interventions.

Community leaders were the next most frequently targeted population group: eight studies incorporated activities designed for local civic, religious and traditional leaders, including all five workshops-contact-econ studies. Two studies targeted both religious and traditional leaders (Causal Design 2016; Cilliers et al. 2018); two targeted religious and civic leaders (Finkel et al. 2018; IMPAQ International 2017); and one targeted traditional and civic leaders (Ferguson 2019). Finally, two studies only targeted traditional leaders (Hartman et al. 2018; Dawop et al. 2019), and one only targeted civic leaders (Cleven 2020). The logic of targeting leaders is different from that of targeting youth: interventions encourage leaders to work well with other leaders and also behave in ways that motivate their followers to promote social cohesion towards other groups.

Two studies targeted both victims and perpetrators of violence in reconciliation processes (Cilliers et al. 2018; Rime et al. 2011), and one study targeted ex-combatants from both sides of the conflict (Schiller 2012). Two studies targeted refugee and host populations (Alan et al. 2020; Ferguson 2019), and one study targeted an internally displaced population (Mousa 2019).

In terms of gender, five studies evaluated interventions that only targeted men (Biton and Salomon 2006; Mousa 2019; Scacco and Warren 2018; Schiller 2012; Vicente and Vilela 2020). One additional study noted that almost all of the participants were men, and it identified issues in the participant targeting that may have led to the imbalance (Svensson and Brouneus 2013). Four studies evaluated interventions that included activities specifically designed for women and girls (Causal Design 2016; Cilliers et al. 2018; Ferguson 2019; IMPAQ International 2017). However, for one of those studies, the women-specific activities were only in an intervention arm that was not covered by the impact evaluation (Causal Design 2016). Five studies identified explicit efforts undertaken to ensure women's participation in the activities (Aladysheva et al. 2017; Finkel et al. 2018; Hartman et al. 2018; Lonergan 2016; Dawop et al. 2019), although one study presented qualitative evidence suggesting that actual engagement of women in activities varied widely by study location (Finkel et al. 2018). The meaningful participation of not only men, but also women, in social cohesion and peacebuilding activities is increasingly noted in the literature (Dayal and Christien 2020; O'Reilly et al. 2017).

Finally, four of the five included studies of media for peace interventions did not undertake any sort of specific population targeting in assignment to treatment (Bilali et al. 2016; Bilali and Vollhardt 2015; Bilali and Vollhardt 2013; Paluck 2009). These researchers did make efforts to ensure study samples included diverse population groups that would have been on opposite sides of conflict.

4.3 Risk of bias in included studies

4.3.1 Summary risk of bias assessment
Twenty-four studies met the study design inclusion criteria for impact evaluations using rigorous quantitative methods. As such, they already meet a relatively high methodological bar. The reliability of the results they present, however, depends on how effectively those rigorous methods have been applied and whether risks of bias have been avoided. Figures 4 and 5 present a summary of the risk of bias assessments across the included impact evaluations for experimental and quasi-experimental designs respectively, and the full risk of bias assessments for each study can be found in Online appendix D.3.

Figure 4: Risk of bias for included studies based on experimental (randomised) designs
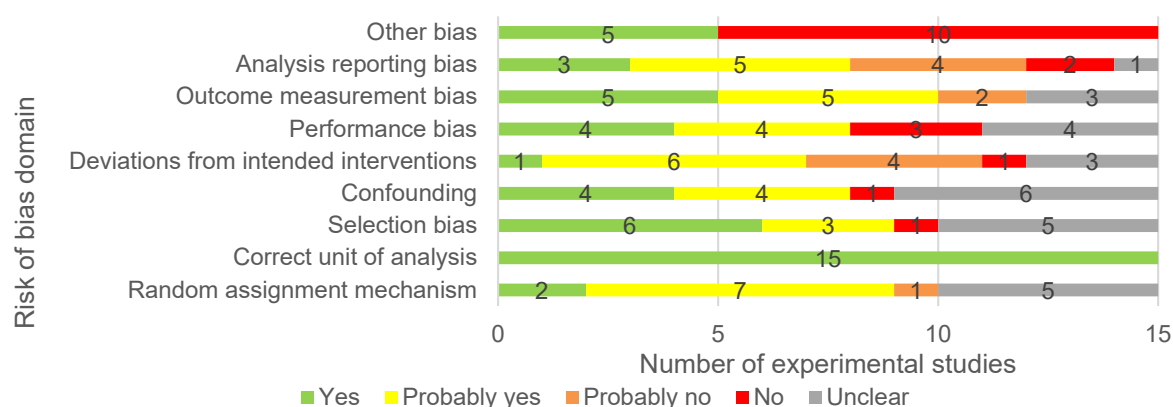
Figure 5: Risk of bias for included studies based on quasi-experimental designs
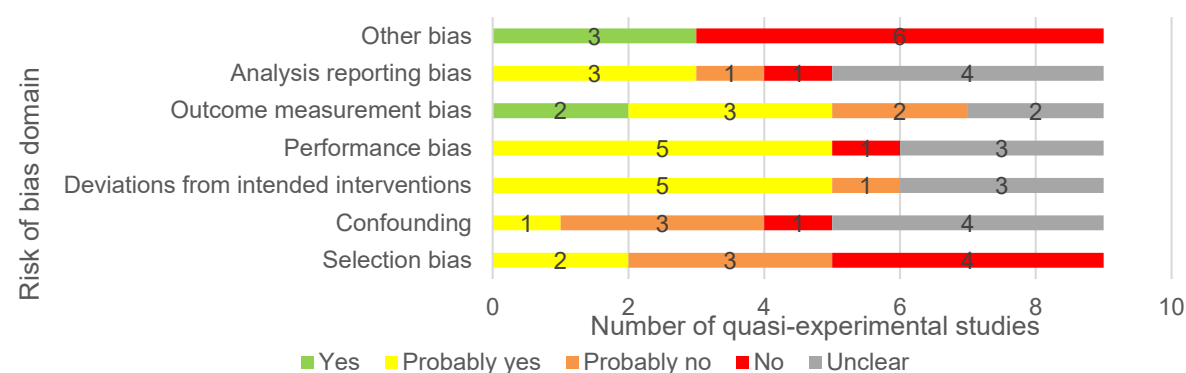
The majority of experimental designs were able to sufficiently address selection bias, with 9 of the 15 studies likely to be free from it. This is in contrast to the quasi-experimental studies, where two out of nine studies were assessed as likely to be free from selection bias. This is in part because, for some social cohesion interventions, participation was based on self-selection and it was difficult to effectively control for unobserved variables that might predict people's participation. For example, people made choices about whether or not to listen to radio programmes within some of the media for peace interventions.

In terms of confounding, 8 of the 15 experimental studies were assessed to be likely free of that bias. However, only one of the nine quasi-experimental studies were assessed as being free from confounding bias. The rest were equally divided between likely being affected by it and not being clear as to whether their methodology was successful in ensuring comparability of groups throughout the study. The experimental studies were generally better able to address confounding issues due to their use of randomisation techniques, which were assessed as plausibly random in nine of the studies.

We assessed half of the studies as likely being free from deviations from the intended intervention, which may include spillovers of intervention effects between groups, crossovers of participants between the treatment and control groups, or contamination by other relevant interventions. This includes seven of the experimental studies and five of the quasi-experimental studies. In some cases, such as when study participants in treatment and control groups were based within the same school or university over the duration of the intervention, preventing interaction between treatment and control groups

may have been challenging. Similarly, in some of the quasi-experimental studies that examined the effect of radio programmes on listeners, it would not be easy to guard against interactions between listeners and non-listeners, or influence from other relevant interventions within participants' daily lives.

We assessed a little over half of the studies as likely being free from performance bias by adopting measures to avoid the risk of participants being affected by being monitored or being compared with others. This included eight of the experimental and five of the quasi-experimental studies. There were, however, a substantial number of studies for which the extent of performance bias was unclear: four and three of the experimental and quasi-experimental studies, respectively. There were also some clear cases where performance bias was not effectively addressed within both types of study design (Alaref et al. 2019; Blattman et al. 2014; Lonergan 2016; Dawop et al. 2019).

We also rated the majority of studies as likely not being affected by outcome measurement bias, which includes issues of social desirability bias that might be expected to have an influence on self-reported outcomes related to social cohesion. Ten of the experimental studies were assessed as likely being free of outcome measurement bias, as were five of the quasi-experimental studies. In some cases, studies included methods that might reduce the likelihood for social desirability bias by using a separate implementation and research team or including games where the factor that was being tested may not have been obvious. For example, through the use of implicit cues about the opponent's ethnic or religious identity in lab games.

The majority of experimental studies likely did not suffer from analysis reporting bias: eight of those study designs were assessed as free or probably free from that bias. This is in contrast to the quasi-experimental studies where three out of nine were likely unaffected by that form of bias. The publication of a PAP can be useful for checking reporting bias to see if there is some selective reporting of results. Among the included studies, only six reported registering a PAP. These were all among the experimental studies. While this is a relatively low proportion of the total number, it is useful to note that the more recently published studies were more likely to have registered a plan, suggesting that the increased attention to research transparency in the social sciences is improving research practices.

Other types of risk of bias were identified in most of the studies. This amounts to ten of the experimental studies and five of the quasi-experimental studies. Some of the other forms of bias we identified included the lack of reporting of full survey instruments, a very limited time between the intervention implementation and the assessment occurring, involvement of the research team in designing the intervention, lack of reporting of geographic distance between treatment and control, displacement of communities between baseline and endline assessment or lack of robustness tests conducted.

4.3.2 Attention to equity and ethics in research and analysis
We also captured information on whether studies explicitly stated that the authors had ethical clearance to undertake the study. Despite the heavy reliance on collecting primary data for their analysis, only seven of the studies reported that they had received ethical clearance for their research (Aladysheva et al. 2017; Alan et al. 2020; Cilliers et al. 2018; Lonergan 2016; Paluck 2009; Scacco and Warren 2018; Vicente and Vilela 2020).

In some cases, the ethics were considered but official ethical review board approval was not received. One study noted that institutional review board approval was not available to them in Jordan, and so they liaised with the Jordanian government to ensure that they had government permission to carry out the work (Ferguson 2019). Three other studies did not reference seeking and gaining explicit ethical approval but described that their research teams were trained on ethical issues prior to commencing the work (Bilali and Vollhardt 2015; Bilali et al. 2016; Paluck 2009).

The rest of the studies did not report whether ethical clearance to undertake the research was sought or granted. It is possible, however, that they did seek or gain that clearance but the evidence is not provided in the reports.

We reviewed all included studies to assess whether and in what ways they made efforts to ensure equity in the research process.[12] The first part of the analysis assessed the ways in which equity was considered, such as through taking into account the specific needs of subpopulation groups in data collection or through undertaking heterogeneity analysis to explore differential impacts of the intervention. The second part looked at the dimensions of equity addressed, meaning the different characteristics that may account for differences in participants' experiences, such as gender, religion or socio-economic status.

Overall, 63 per cent (n = 15) of studies in this review incorporated at least one method of addressing at least one dimension of equity. Among most intervention groups, this corresponded to all but one study including some efforts to address equity. Media for peace interventions were the exception, where only one study out of five incorporated equity considerations into the intervention or research design (Paluck 2009).

Figure 6 presents the frequency with which studies used different methods for addressing or incorporating equity considerations. The most frequent way in which equity was addressed was through subgroup analysis, yet less than half of included studies (n = 11) presented disaggregated effects for key population groups. Of these, the majority looked at differential effects for the different population groups targeted in the intergroup cohesion interventions, and split the analysis along the salient social cleavages. Five studies disaggregated findings by ethnic group (Aladysheva et al. 2017; Hartman et al. 2018; IMPAQ International 2017; Lonergan 2016; Paluck 2009); three studies by religious group (Biton and Salomon 2006; Hartman et al. 2018; Scacco and Warren 2018);[13] and one by displacement status (Alan et al. 2020).

We further identified five studies that looked at heterogeneous impacts for individuals exposed to violence, including victims of violent conflict (Cilliers et al. 2018; Lonergan 2016; Paluck 2009; Rime et al. 2011; Scacco and Warren 2018). Two of those studies that analyse the impacts of truth and reconciliation processes further include effects for ex-combatants (Cilliers et al. 2018; Rime et al. 2011). We also identified four studies that undertook subgroup analysis by sex (Aladysheva et al. 2017; Alan et al. 2020; Cilliers et

---

[12] We include efforts to ensure that research design was culturally appropriate in consideration of risk of measurement bias.

[13] Mousa (2019) reported all main results for a single religious group, Christians, and exploratory findings for effects on Muslim participants. The limited number of Muslims participating in the intervention prevented stronger analysis.

al. 2018; Hartman et al. 2018): two that conducted analysis of heterogeneous effects by age group, although only for some outcomes (Hartman et al. 2018; Okunogbe 2018), and two that disaggregated findings by socio-economic group (Hartman et al. 2018; Scacco and Warren 2018). Finally, two studies used qualitative research to investigate the particular experiences and impacts of the interventions for women (Finkel et al. 2018; IMPAQ International 2017).

Figure 6: Methods for addressing equity in included studies
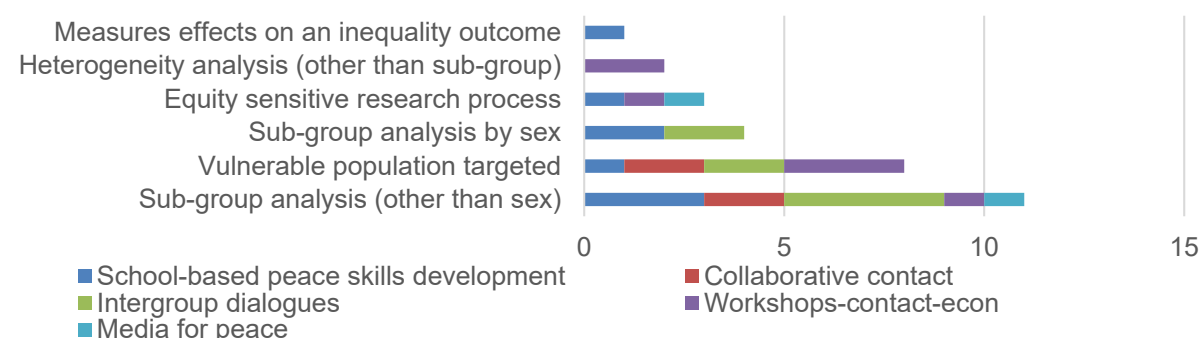
Beyond analysis of heterogenous effects, the most frequent way in which studies addressed equity was by targeting a vulnerable, underserved population (Alan et al. 2020; Cilliers et al. 2018; Ferguson 2019; Finkel et al. 2018; Dawop et al. 2019; Mousa 2019; Rime et al. 2011; Scacco and Warren 2018).[14] Three studies noted equity-sensitive research processes, such as ensuring disaggregated focus groups so more vulnerable populations would be comfortable participating (Alan et al. 2020; Finkel et al. 2018; Paluck 2009). Finally, one study presented exploratory analysis of effects on an inequality outcome (Alan et al. 2020), where a measure was developed to assess levels of classroom segregation between host and refugee children.

4.3.3 External validity

We considered several factors when assessing the external validity of studies. These include the approach used by researchers to select the study population, whether the programme implemented was a small-scale pilot or a large-scale established programme, and the characteristics of the population and setting of the study. In addition to collecting information on those factors, we extracted information on the authors' own discussion of the generalisability of their findings.

Some of the ways that authors attempted to support external validity were by ensuring that their random sample reflected the ethnic make-up of the population (Paluck 2009) or by targeting members of society that were typically in greatest need of the intervention in the broader population (Vicente and Vilela 2020). Some challenges to external validity in the studies included the use of convenience sampling or country-specific contextual factors. For example, a set of media for peace studies selected participants by approaching people in public settings to ask them to join the study, which rendered them

---

[14] Given the focus of this review, we did not code studies as targeting vulnerable populations where the entire local population could be considered vulnerable due to the conflict contexts in which many studies were implemented. Rather, we limited the code to interventions where exposure to violence or conflict was a condition of being eligible for treatment, or a vulnerable population subgroup was a key population of focus in the analysis.

unlikely to be representative of the general population (Bilali and Vollhardt 2013; Bilali and Vollhardt 2015; Bilali et al. 2016). Another study that focused on promoting dispute resolution reflected on the fact that Liberia's specific history of informal institutions likely provided a basis on which the intervention could build, which might be absent in other contexts (Blattman et al. 2012).

4.3.4 Alignment between final reports and pre-analysis plans

To assess research transparency and risk of bias due to changes in outcome reporting, we assessed the extent to which final studies report on all hypotheses and outcomes proposed in pre-analysis plans (PAPs), where available. This analysis provides a summary of our findings; details for each study are included in Online appendix B.2. We analysed whether discrepancies were reported or not, and if they were, what rationale was provided, aiming to undertake a simple analysis to check for this key risk of bias (Goldacre et al. 2016). Pre-registration for social science studies became available in 2011,[15] so we did not include the seven studies based on data collected prior to 2011 in this analysis.[16]

Of the 17 studies based on data collected in 2012 or later, 41 per cent (n = 7) were based on PAPs (Alan et al. 2020; Cilliers et al. 2018; Ferguson 2019; Dawop et al. 2019; Mousa 2019; Scacco and Warren 2018; Vicente and Vilela 2020).[17] This represents over half of the studies published in the past three years, suggesting that the use of PAPs is becoming increasingly common. Of those seven, all were pre-registered except for one, which was based on a PAP that was not published but was shared upon request (Ferguson 2019). Four of the pre-registered studies were registered before final data collection was completed (Alan et al. 2020; Cilliers et al. 2018; Dawop et al. 2019; Mousa 2019). Two of the pre-registered studies were only registered after data collection was completed (Scacco and Warren 2018; Vicente and Vilela 2020), leaving open the risk that findings could have informed the PAPs. As no registration details are available online for Ferguson (2019), the date of the PAP cannot be confirmed. All six studies for which we could access the PAP included discrepancies between the PAPs and final reports.[18] Three studies reported all changes to the PAP in the final report (Cilliers et al. 2018; Ferguson 2019; Mousa 2019). One study reported some changes to secondary outcomes, but did not

---

[15] The Evidence in Government and Politics (EGAP) design registry was launched in 2011. Options increased in 2013, with the launch of the AEA RCT Registry and 3ie's RIDIE, for impact evaluations based on experimental and quasi-experimental designs.

[16] The following studies were based on data collected prior to 2011: Bilali et al. 2016; Bilali and Vollhardt 2015; Bilali and Vollhardt 2013; Biton and Salomon 2006; Paluck 2009; Rime et al. 2011; Svensson and Brouneus 2013. One study (Hartman et al. 2018, p.11) reported that both of the endline surveys, which took place in 2011 and 2013, took place before the development of a social science registry. However, EGAP's design registry includes trial registries dating as far back as March 2011. One of the included studies that was pre-registered, Cilliers et al. (2018), registered its PAP on EGAP in December 2012 – calling into question this claim.

[17] Of studies based on data collected after 2011, 59 per cent (n = 10) did not pre-register or report being based on PAPs, including: Aladysheva et al. 2017; Alaref et al. 2019; Causal Design 2016; Cleven 2020; Finkel et al. 2018; Hartman et al. 2018; IMPAQ International 2017; Lonergan 2016; Okunogbe 2018; and Schiller 2012.

[18] The PAP for the seventh study, Vicente and Vilela 2020, has been requested. Insufficient information is provided on the registration page to undertake an adequate assessment.

report changes to hypotheses or primary outcomes (Alan et al. 2020). Two studies were not transparent about any of the changes made to the final analyses (Dawop et al. 2019; Scacco and Warren 2018). In one case, the authors had merged hypotheses into new index measures (Dawop et al. 2019), while in the other, authors had organised the five pre-specified outcome groups into three hypotheses, and dropped a number of pre-specified outcomes from the final report (Scacco and Warren 2018).

4.4 Quantitative analysis of social cohesion outcomes

Below, we report the results for Review Question 1, looking at the effects of interventions that aim to build intergroup social cohesion on social cohesion outcomes. As described in detail above, where three or more different studies contributed effect sizes for similar constructs for a particular intervention-outcome combination, we conducted a meta-analysis. We further conducted meta-analyses of effects on each dimension of social cohesion across all intervention types. In addition, we conducted an overall meta-analysis using all available data, estimated with robust variance estimation to account for the dependent effect sizes. In the context of the overall analysis, we also looked at three potential moderators of the relationship between interventions and social cohesion outcomes: (1) risk of bias; (2) study design; and (3) length of intervention exposure. Where meta-analysis was not possible, we provide a short narrative description of the effects found in the individual studies.

We first present a narrative summary of findings across all interventions, before we present a detailed analysis of effects on different social cohesion outcomes for each intervention group, in Sections 4.4.2 to 4.4.6.

Throughout, where a study reports outcomes for which a negative movement indicates an improvement in social cohesion, such as a reduction in levels of bias, the sign has been reversed. For all the meta-analyses reported below, we conducted sensitivity analyses and created funnel plots to investigate publication bias. In the following, we note only instances where we detected such bias. The full results of these analyses can be found in Online appendix C.3.

4.4.1 Summary of effects on social cohesion outcomes
Table 6 presents an overview of the findings for each intervention group and each dimension of social cohesion, followed by a detailed summary of the key findings by each dimension of social cohesion. Each cell in the matrix represents the outcomes reported by studies within an intervention group.

In each of the blue boxes (studies), each bubble represents a relevant reported effect size. The size of the bubbles relates to the absolute value of the effect size: we calculated Hedges' *g* for each outcome, to enable comparison across studies (see Section 5.3 for a detailed description of the approach applied). In short, the smaller the bubble, the smaller the effect size. We further differentiate between outcomes where the confidence interval for the effect size does or does not cross the line of no effect. This means that green-shaded bubbles correspond to positive effects, wherein the 95% confidence interval does not cross the line of no effect. On the other hand, red-shaded bubbles correspond to negative outcomes where the confidence interval does not cross the line of no effect, wherein the intervention actually had a negative impact on that particular measure of

social cohesion in that particular study. Grey-shaded bubbles correspond to outcomes where the confidence interval crosses the line of no effect; this means that, while the (+)/(-) sign denotes the direction of the effect, the 95% confidence interval includes the possibility that the true effect is null. Finally, the darkness of the shading corresponds to the precision of the estimate, based on the absolute value of the 95% confidence interval. The darkest shade is for effects wherein the 95% confidence interval is relatively short (< 0.2 SD). By contrast, findings with the lightest shade are those where the absolute value of the 95% confidence interval is comparatively longer (> 0.6 SD). Where the confidence intervals are shorter, we can have more confidence that the true effect is close to the estimated effect. We undertake a detailed analysis of each cell in the sections below, and here present an overview analysis of the findings.

Table 6: Summary of effects of included studies on intergroup social cohesion

The table presents a grid of effect-size bubbles. Columns: Intervention group, Trust, Sense of belonging, Willingness to participate, Willingness to help, Acceptance of diversity, Social cohesion index. Rows: School-based peace education, Collaborative contact, Intergroup dialogues, Workshops-contact-econ, Media for peace.

Legend:

Bubble size key: Each bubble represents a single effect, based on the size of the Hedges' g standardised effect, the direction of the effect, and if it crosses line of no effect (grey if yes)

Shading key: Absolute value of 95% confidence interval

	< 0.2		< 0.2		< 0.2
	0.2 - 0.4		0.2	-	0.4		0.2 - 0.4
	0.4 - 0.6		0.4	-	0.6		0.4 - 0.6
	> 0.6		> 0.6		> 0.6
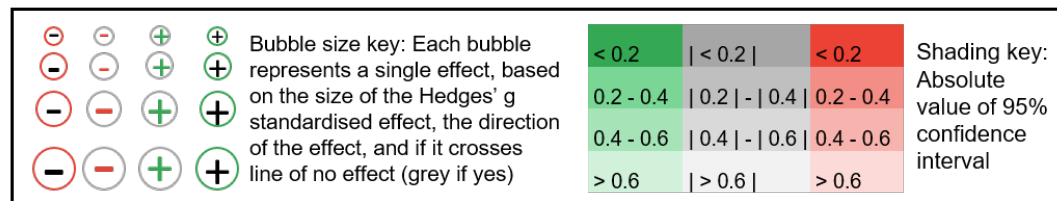
In analysing the effects of interventions that aim to build intergroup social cohesion on trust, we typically find small effects, many of which are in a positive direction but cross the line of no effect. Our meta-analysis identifies small, but statistically significant positive effects on outcomes of trust among workshops-contact-econ studies and media for peace interventions, but no effect of intergroup dialogue interventions. We find that the average effect of workshops-contact-econ interventions on trust is positive and statistically significant (g = 0.08, [0.00, 0.16]). Among the media for peace interventions, the meta-analysis identifies a positive effect on trust of three radio drama programmes (g = 0.10, [0.02, 0.18]).

For both workshops-contact-econ and radio dramas, the individual effects on trust were often imprecise or crossed the line of no effect. The fact that the meta-analyses of these two intervention groups identify positive effects with a smaller 95 per cent confidence interval that does not cross the line of no effect demonstrates the value of meta-analyses for comparable programmes measuring similar outcomes.

We identify larger positive effects for some measures of trust among school-based peace education interventions, yet negative effects were identified in a third study. There are important differences between the programmes, however. The two studies identifying positive effects reported effects on outcomes for the schoolchildren who were the primary recipients of the treatment, whereas the study identifying negative effects surveyed parents of children who received the treatment. While some of the parents also participated in dialogue interventions, not all of them did, and this difference may be driving the divergent findings.

Only one study reported outcomes of trust among the collaborative contact interventions, with mixed findings, both of which cross the line of no effect. We find that intergroup dialogue interventions have mixed effects. Two studies reported small effects that, although in a positive direction, typically cross the line of no effect, while the third found a substantial negative impact.

We find mixed results of intergroup social cohesion interventions on measures of a shared sense of belonging. Studies of collaborative contact, intergroup dialogues, and workshops-contact-econ interventions find generally small effects that are in a positive direction but cross the line of no effect. However, we identify significant negative effects for some school-based peace education and media for peace interventions. The only meta-analysis we were able to run for this dimension of social cohesion identifies a positive impact of workshops-contact-econ interventions, but the 95 per cent confidence interval crosses the line of no effect (g = 0.10, [-0.01, 0.21]).

Of the negative effects identified on sense of belonging, one stems from a school-based peace education intervention that found a negative effect on participants feeling at home in their country. Other school-based peace education interventions identified small effects in a positive direction, but these cross the line of no effect. The other negative effect identified is from a media for peace study of a radio drama, which found an increase in perceptions of the need for politicians who would promote their ethnic group's interest (i.e. a negative impact on intergroup social cohesion), although this outcome was in the context of ongoing violent conflict in the DRC. No other media for peace intervention measured effects on a shared sense of belonging.

We find that collaborative contact interventions have a positive effect on willingness to participate. The majority of other interventions typically yield small positive effects, many of which cross the line of no effect, on this dimension of social cohesion. However, we identify significant negative effects for some school-based peace education and media for peace interventions. We found mixed effects on willingness to participate from two school-based peace education interventions: one study that measured effects on the students who participated had positive effects, while the other study that measured effects on parents found negative effects. The meta-analysis of collaborative contact interventions finds a small, but significant, positive effect that was quite precisely estimated (g = 0.06, [0.01, 0.10]). Only one intergroup dialogues study reported a measure of willingness to participate. It was also the only intervention from this group that incorporated an element of collaborative contact in the intervention, and it found a positive impact on willingness to participate (Cilliers et al. 2018).

Our meta-analysis of workshops-contact-econ interventions finds a small effect in a positive direction, but the estimate is very small (g = 0.06, [-0.05, 0.16]). Finally, two studies of media for peace interventions identified effects on willingness to participate in a negative direction, one of which did not cross the line of no effect.

Fewer effects on willingness to help were reported by studies in this review, and they measured this dimension of social cohesion in a variety of ways. However, we identify a pattern of typically small effects in a positive direction, many of which do not cross the line of no effect. Across all intervention groups, no more than two studies reported on the same outcome construct for willingness to help, which made effective synthesis challenging, particularly within each intervention group. We conducted two review-level meta-analyses for this dimension of social cohesion. We find a small positive effect on self-reported measures of willingness to help (g = 0.08, [0.03, 0.14]) across all intervention types. On willingness to help measured through behaviours within lab games, we also find a small positive effect, but the estimate is imprecise (g = 0.06, [-0.01, 0.12]). Only two studies reported both lab game and self-reported measures of willingness to help, and both identified larger effects on the self-reported measures, which suggests a risk of social desirability bias. Given this point, and the heterogeneity of the underlying measures feeding into these analyses, some caution is required in interpreting these findings, but they are broadly indicative of small positive effects on willingness to help.

Overall, we find limited to no effects of intergroup social cohesion interventions on acceptance of diversity. Within each meta-analysis by intervention group, there were studies reporting effects that pulled in opposite directions, leading to high heterogeneity within the models. Among school-based peace education interventions, effects were typically in a positive direction yet these cross the line of no effect. As the underlying measures were of very different constructs, we were unable to run a useful meta-analysis. Our meta-analysis of the effect of collaborative contact interventions on acceptance of diversity finds a small effect in a positive direction, but it crosses the line of no effect (g = 0.03, [-0.2, 0.46]). This null effect is driven by a single study (Okunogbe 2018), which reported a small negative effect on inter-ethnic friendship. However, it was one of the studies with the lowest risks of bias in its design and analysis included in this review, and as such, we have more confidence in the finding. The meta-analysis of

intergroup dialogue interventions identifies a large effect in a negative direction, although it crosses the line of no effect (g = -0.42, [-1.52, 0.67]). This is primarily driven by a single study, noted above, which reported outlier effects across the review and has a high risk of selection bias and confounding (Lonergan 2016). As such, this finding should be interpreted with caution. Only two workshops-contact-econ studies reported effects on acceptance of diversity, and both were positive, but these cross the line of no effect. Finally, we were able to run four meta-analyses of different outcome constructs measuring the effects of media for peace interventions through radio dramas on acceptance of diversity. Each identifies an effect in a positive direction, yet these cross the line of no effect (g = 0.26, [-0.13, 0.66]; g = 0.07, [-0.02, 0.17]; g = 0.07, [-0.06, 0.21]; g = 0.07, [-0.02, 0.17]). In each case, the imprecision and heterogeneity in the models were driven by a single study, which evaluated the effects of a media for peace intervention during a period of violent conflict in the DRC and had a higher risk of bias than the other studies in these meta-analyses (Bilali and Vollhardt 2015). Our sensitivity analyses find that, when this study is dropped, the effects become larger and no longer cross the line of no effect. This suggests that, in some contexts, media for peace interventions may have a positive effect, but not in contexts of ongoing violent conflict.

While there are some exceptions, the overall pattern across intervention types and different measures of social cohesion suggests small positive effects. We used robust variance estimation to examine the effect of all interventions on any measure of social cohesion using all available coded data (a total of 796 effect sizes). The result of this analysis indicates that overall interventions produce a significant small positive effect on relevant outcomes (g = 0.06, [0.03, 0.10]). The effect remains small and positive when outliers are removed.

We were also able to use that overall analysis to examine three moderators: risk of bias, study design, and length of intervention exposure. We find no effect of studies of higher or lower risk of bias on the results and there is also no effect for whether study designs were experimental or quasi-experimental. This result lends some confidence to the findings, as it suggests that the effects are not driven by studies with a higher risk of bias or a particular study design. In addition, we find no effect of the length of intervention exposure. These results are described in detail in Online appendix B.4. We now present the detailed findings by intervention group.

4.4.2 School-based peace education interventions
Four studies on school-based peace education interventions reported effect sizes on social cohesion outcomes. However, one reported only a single effect, and of the remaining three, two measured effects on students who received the treatment while the third measured effects on parents who did not receive any direct treatment.[19] As such, we were unable to conduct meta-analysis of the results for this group of interventions. Caution should also be applied because we assessed three of the studies as being of high risk of bias. In two cases, the study designs pose risks of selection bias and confounding (Biton and Salomon 2006; Cleven 2020). In the third, there were likely deviations from the intended intervention that may underestimate programme effect

---

[19] Of a sample of 300 respondents from the treated school, only 6 parents reported having participated in one of the dialogue sessions for parents run by the intervention (Cleven 2020).

(Aladysheva et al. 2017). We classed the other study as having some concerns related to risk of bias (Alan et al. 2020).

Overall, we find that studies reporting effects of school-based peace education on outcomes for students themselves are typically in a positive direction, although many cross the line of no effect. The exception is sense of belonging, for which we find a significant negative effect for one study, and two small effects that, although in a positive direction, cross the line of no effect.

Trust

Three studies reported effects of school-based interventions on trust outcomes. These three studies took place in Bosnia and Herzegovina, Kyrgyzstan and Turkey, and the outcome measures included in the analysis focused on measures of intergroup trust.

Among the two studies measuring effects for students who participated in the peace education interventions, the intervention in Turkey has a significant positive effect ($g = 0.11$, [0.06, 0.16]) (Alan et al. 2020). However, subgroup analysis found that this effect was driven by impacts on host children, whose play during the lab game suggested greater trust towards children they did not know ($g = 0.13$, [0.08, 0.19]) compared with non-treated host children. By comparison, the effect on refugee children, although in a positive direction, is very small and crosses the line of no effect ($g = 0.04$, [-0.08, 0.16]). The study in Kyrgyzstan similarly found an effect in a positive direction, but it is small and crosses the line of no effect ($g = 0.04$, [-0.15, 0.24]) (Aladysheva et al. 2017).

In contrast, the study of the intervention in Bosnia and Herzegovina had a small negative effect on parents' self-reported trust of out-groups, but it is not statistically significant ($g = -0.09$, [-0.25, 0.07]) (Cleven 2020).

Sense of belonging

Three studies reported effects of school-based interventions on sense of belonging outcomes, assessing interventions that took place in Bosnia and Herzegovina, Palestine and Kyrgyzstan. The study of the intervention in Bosnia and Herzegovina reported an intergroup measure of how close the respondent felt to the out-group (Cleven 2020). There was no significant effect of the intervention on the outcome for parents with children at the targeted school ($g = 0.15$, [-0.01, 0.31]). The study in Palestine reported effects on the extent to which participants associated the concept of peace with values in line with positive peace ideals of cooperation and harmony across groups (Biton and Salomon 2006). There was no significant effect of the intervention ($g = 0.03$, [-0.20, 0.25]). The study of the intervention in Kyrgyzstan reported a generalised measure of how at home the respondent felt in the country and found a negative and statistically significant effect ($g = -0.27$, [-0.42, -0.13]) (Aladysheva et al. 2017).

Willingness to participate

One study reported effects on children's willingness to participate through cooperation in games, while a second reported effects on parents' participation in school associations.

The former study (Alan et al. 2020) found that treated refugee and host community children in Turkey did not behave in significantly different ways from children in untreated schools when playing with children they did not know ($g = -0.00$, [-0.05, 0.04]). When they played against children they did know, however, there was a small positive effect (g

= 0.07, [0.02, 0.12]). The latter study (Cleven 2020) reported a generalised measure of parent participation in school associations in Bosnia and Herzegovina, finding a large negative effect (g = -0.43, [-0.59, -0.27]). This effect is larger than many of the others reported in this discussion but comes from a study with high risk of bias due to selection bias, which evaluated effects on parents whose children went to the targeted school but who did not participate in the intervention themselves, and should therefore be treated with caution.[20]

Willingness to help
Two studies that focused on interventions in Kyrgyzstan and Turkey reported on willingness to help. Both studies reported behavioural measures of altruism based on the level of donations students gave within a game played in class. The interventions do not appear to have increased willingness to help in either context (g = 0.07, [-0.09, 0.23]; g = 0.01, [-0.03, 0.06]). In Turkey, however, there appears to have been a small positive effect when children played against children they did know (including a mix of refugee and host children) (g = 0.09, [0.04, 0.14]) (Alan et al. 2020). Further, in Kyrgyzstan, the impacts varied by population group (Aladysheva et al. 2017). Sex-disaggregated subgroup analysis found that the intervention had a positive effect on girls (g = 0.28, [0.08, 0.48]), but not on boys (g = -0.01, [-0.27, 0.25]). The effects also varied by ethnicity: the programme had a positive impact on donations in public goods games on students from the ethnic majority (g = 0.23, [0.04, 0.43]), but not on minority students (g = -0.07, [-0.28, 0.13]).

Acceptance of diversity
Two studies that focused on interventions in Bosnia and Herzegovina and in Turkey reported measures of intergroup friendship. The study of the intervention in Turkey reported a positive and statistically significant effect (g = 0.29, [0.03, 0.56]) (Alan et al. 2020). In Bosnia and Herzegovina, the effect was close to zero (g = 0.03, [-0.13, 0.19]) (Cleven 2020).

A third study focused on an intervention in Kyrgyzstan and reported a measure of the need to protect one's culture, religion and language from others (Aladysheva et al. 2017). This study reported no significant effect of the intervention on the outcome (g = 0.07, [-0.07, 0.21]).

4.4.3 Collaborative contact interventions
Four studies on collaborative contact interventions reported effect sizes on social cohesion outcomes. Due to the diversity in outcomes reported and the limited overlap between them, we were only able to conduct meta-analyses for two of the social cohesion outcome groups (willingness to participate and acceptance of diversity). In addition to the relatively low number of studies, caution should be used in interpreting the results because three of the studies are classified as having high risk of bias (Alaref et al. 2019; Mousa 2019; Okunogbe 2018). The other study was classed as low risk of bias (Scacco and Warren 2018). The meta-analysis shows a small positive effect of interventions on willingness to participate. Two studies showed a consistent small and

---

[20] As noted, 6 parents from the sample of 300 did report participating in a dialogue session, but the author states that running all analyses without these 6 participants had no effect on the findings (Cleven 2020).

positive effect on sense of belonging outcomes. These interventions showed no effect on trust, willingness to help or acceptance of diversity outcomes.

Trust

Only one study of collaborative contact reported a measure of trust. This study focused on an intervention in Iraq and reported a measure of Christian trust for Muslims (Mousa 2019). They found no significant effect of the intervention on that measure (g = 0.12, [-0.19, 0.42]).

Sense of belonging

Only two studies of collaborative contact interventions reported effect sizes on a shared sense of belonging, so a useful meta-analysis could not be conducted. These two studies focused on interventions that took place in Iraq and in Lebanon. The study focused in Iraq reported an intergroup measure of beliefs about coexistence prioritising a national identity over an ethnic or religious identity (Mousa 2019). They found a positive and statistically significant effect of the intervention on that measure (g = 0.29, [0.05, 0.54]). The study focused in Lebanon reported a generalised measure of sense of belonging to the Lebanese community (Alaref et al. 2019). They also found a positive and statistically significant effect of the intervention on that measure (g = 0.14, [0.00, 0.28]).

Willingness to participate

Three studies contributed effect sizes to the meta-analysis of the average effect of collaborative contact interventions on willingness to participate outcomes. One of these interventions took place in Iraq and the other two were implemented in Nigeria. The outcomes included in the analysis are measures of intergroup participation, with two studies measuring self-reported openness to participation and one study reporting a behavioural measure.

As can be seen from Figure 7, the meta-analysis shows a small, positive and statistically significant average effect (g = 0.06, [0.01, 0.10]). The positive effect was driven by two of the studies. The intervention in Iraq produced the largest effect size but the confidence interval passes the line of no effect (Mousa 2019). One of the studies in Nigeria produced a smaller, positive but statistically significant effect (Okunogbe 2018). The other Nigerian study reported a negative effect, with the smallest effect size in the analysis and a confidence interval crossing the line of no effect (Scacco and Warren 2018). As such, there is a minimal amount of heterogeneity in the model (I^2 = 0.05%).

Our sensitivity analysis indicates that the main average effect is not affected by the removal of any of the studies. The removal of Okunogbe (2018), however, substantially increases the heterogeneity in the model (I^2 = 64.00%) because it leaves only two effects that pull in opposite directions.

Figure 7: Forest plot of effects on willingness to participate outcomes among collaborative contact interventions
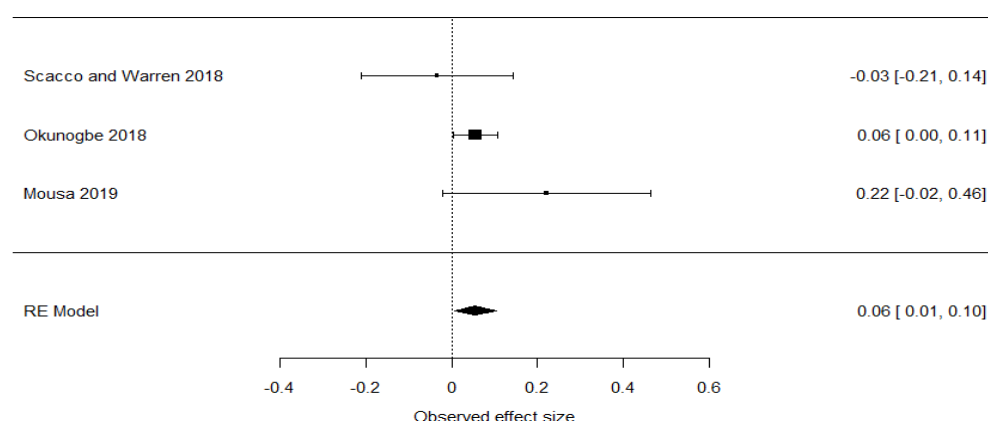
Willingness to help

While three studies reported effect sizes of collaborative contact interventions on willingness to help outcomes, only two contained measures of comparable constructs. We deemed the other measure too dissimilar to use in a meaningful meta-analysis. Two studies, focused on interventions that took place in Iraq and in Nigeria, reported behavioural measures of intergroup help in terms of donations (Mousa 2019; Scacco and Warren 2018). Neither of the studies found a significant effect of the interventions (Mousa: ($g = 0.06$, [-0.18, 0.30]); Scacco and Warren: ($g = -0.07$, [-0.25, 0.10])).

The third study focused on an intervention in Lebanon and reported a measure of perceived contributions of volunteering to social cohesion (Alaref et al. 2019). They also reported no significant effect of the intervention on the outcome ($g = -0.04$, [-0.18, 0.10]).

Acceptance of diversity

Four studies contributed effect sizes to the meta-analysis of the average effect of collaborative contact interventions on acceptance of diversity outcomes. One of these interventions took place in Iraq, another in Lebanon and two took place in Nigeria. The outcomes included in the analysis are all measures of intergroup tolerance of the respondent towards the out-group.

As can be seen from Figure 8, the meta-analysis suggests a small and positive average effect but the confidence interval crosses the line of no effect ($g = 0.03$, [-0.02, 0.46]). The positive effect is driven by two of the studies. Three of the studies reported positive effect sizes; however, all of their confidence intervals cross the line of no effect. Only one effect included in the meta-analysis is statistically significant, the negative effect from one of the interventions in Nigeria (Okunogbe 2018). The relative size and precision of that negative effect helps to create the substantial amount of heterogeneity in the model ($I^2 = 80.14\%$).

Our sensitivity analysis indicates that the main average effect is affected by the removal of one of the studies. When Okunogbe (2018) is removed from the analysis, the main average effect becomes positive and statistically significant ($g = 0.12$, [0.01, 0.22]). This removal also substantially reduces the heterogeneity in the model ($I^2 = 0.01\%$). This study, however, is of a lower risk of bias than two of the other studies that would remain in the model, which raises questions about the reliability of the results once it is removed.

Figure 8: Forest plot of effects on acceptance of diversity among collaborative contact interventions
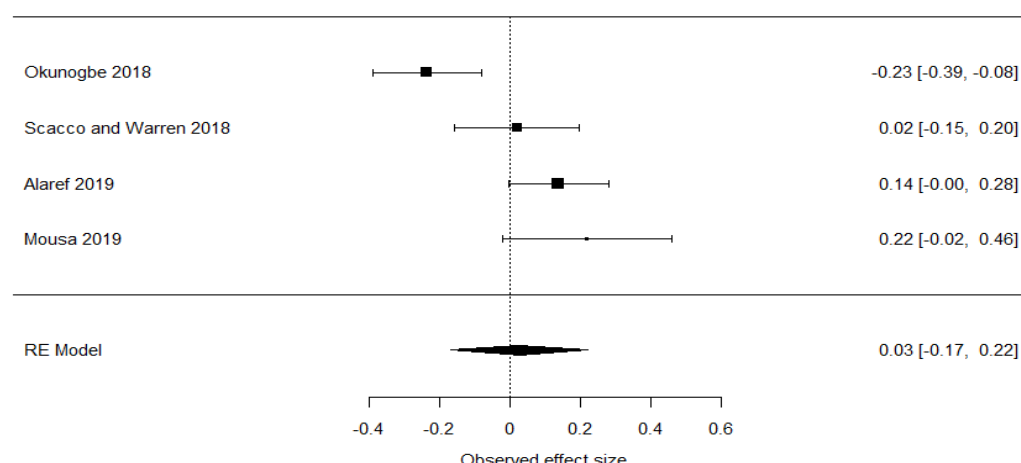
Combined social cohesion outcomes measured as an index

One study of a collaborative contact intervention in Nigeria also reported an effect size for an index that combined several of the social cohesion outcome categories (Okunogbe 2018). The measure was an index of attitudes towards the non-Yoruba out-group in terms of feelings of closeness, trust and openness to inter-ethnic marriage. There was no effect of the intervention on that index measure (g = 0.03, [-0.12, 0.19]).

4.4.4 Intergroup dialogue interventions

Five studies on intergroup dialogue interventions reported effect sizes on social cohesion outcomes.[21] Due to the diversity in outcomes reported and the limited overlap between them, we were only able to conduct meta-analyses for two of the social cohesion outcome groups (trust and acceptance of diversity). In addition to the relatively low number of studies, caution should be used in interpreting the results because all but one (Cilliers et al. 2018) of the studies are classified as having high risk of bias (Hartman et al. 2018; Lonergan 2016; Schiller 2012; Svensson and Brouneus 2013). Overall, there appear to be no or inconclusive effects of these interventions on social cohesion outcomes. The meta-analyses suggest no effect of the interventions on either trust or acceptance of diversity outcomes. Only one study reported outcomes related to willingness to participate and willingness to help; both of these effect sizes were small, positive and statistically significant (Cilliers et al. 2018). However, the results are not conclusive because they come from only a single study. The individual effect sizes examined for the sense of belonging outcomes reported inconsistent results. Only one of those effects was statistically significant but it was from a study with high risk of bias and for which myriad quality concerns were identified. As such, that positive effect should be interpreted with strong caution.

Trust

Three studies contributed effect sizes to the meta-analysis of the average effect of intergroup dialogue interventions on trust outcomes. These three studies took place in
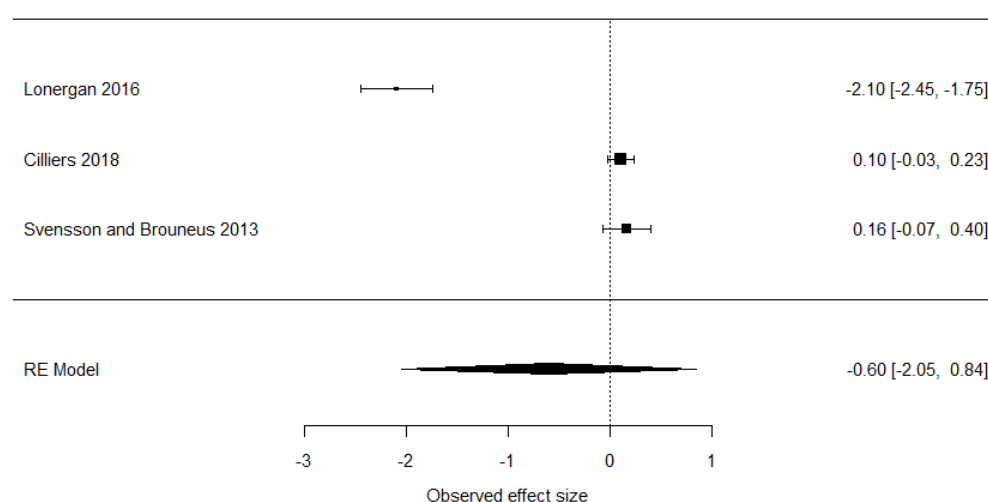
---

[21] We were unable to compute standardised effect sizes based on the information included in the Rime et al. (2011) paper.

Ethiopia, Sierra Leone and Sri Lanka, and the outcome measures included in the analysis focused on measures of intergroup trust that respondents had for the out-group. The studies in Ethiopia and Sri Lanka involved university students (Svensson and Brouneus 2013; Lonergan 2016), while the study in Sierra Leone evaluated a truth and reconciliation process involving victims and perpetrators of violence during the civil war (Cilliers et al. 2018).

As can be seen from Figure 9, the meta-analysis suggests a negative average effect, but the confidence interval crosses the line of no effect (g = -0.60, [-2.05, 0.84]). This is not surprising given the considerable amount of heterogeneity evident from both visual inspection of the forest plots, as well as the results of the statistical test of heterogeneity (I^2 = 99.19%). This heterogeneity appears to be largely driven by the dialogue intervention in Sri Lanka, which reported a strong negative and statistically significant effect (Lonergan 2016). This is in contrast to the two small positive effects reported by the other studies. The confidence intervals of those two positive effects, however, cross the line of no effect.

Our sensitivity analysis indicates that the main average effect is not affected by the removal of any of the studies. However, the removal of the Lonergan (2016) study does substantially reduce heterogeneity (I^2 = 00.00%).

Figure 9: Forest plot of effects on trust among intergroup dialogue interventions

Lonergan 2016	-2.10 [-2.45, -1.75]
Cilliers 2018	0.10 [-0.03, 0.23]
Svensson and Brouneus 2013	0.16 [-0.07, 0.40]
RE Model	-0.60 [-2.05, 0.84]

Observed effect size

Sense of belonging
Two intergroup dialogue studies reported effects on the salience of ethnic and national identities (Lonergan 2016; Svensson and Brouneus 2013). A third study reported an index measure of attitudes about intergroup social tensions (Cilliers et al. 2018). However, this measure was substantially different from the others, and as such, a useful meta-analysis was not feasible.

The two studies reporting on salience of ethnic and national identities evaluated intergroup dialogue interventions targeting university students in Sri Lanka and Ethiopia. The study of students in Sri Lanka reported a large, significant effect on the salience of national identity (g = 2.85, [2.47, 3.2]) compared with ethnic identities (Lonergan 2016). However, the high risk of bias identified in that study means the findings should be interpreted with caution. The study of students in Ethiopia identified a small negative

effect on a measure of the salience of ethnic identities but the confidence interval for that effect crosses the line of no effect (g = -0.04, [-0.28, 0.2]) (Svensson and Brouneus 2013).

The study reporting on intergroup social tensions evaluated an intervention in Sierra Leone of a truth and reconciliation process in a post-conflict context. They found no significant effect of the intervention on that measure (g = 0.05, [-0.02, 0.12]) (Cilliers et al. 2018).

Willingness to participate

Only the study of the truth and reconciliation intervention in Sierra Leone reported a measure of willingness to participate outcomes. They reported an index measure of participation in community groups (Cilliers et al. 2018). They found a positive and significant effect of the intervention on that measure (g = 0.12, [0.05, 0.20]).

Willingness to help

Only the same study of the truth and reconciliation process in Sierra Leone reported a measure of willingness to help outcomes. They reported an index measure of public good contributions (Cilliers et al. 2018). Again, they found a small, positive and significant effect of the intervention on that measure (g = 0.07, [0.00, 0.14]).

Acceptance of diversity

Four studies contributed effect sizes to the meta-analysis of the average effect of intergroup dialogue interventions on acceptance of diversity outcomes. These four studies took place in Ethiopia, Indonesia, Liberia and Sri Lanka, and the outcome measures included in the analysis focused on measures of intergroup bias.

As can be seen from Figure 10, the meta-analysis suggests a negative average effect, but the confidence interval crosses the line of no effect (g = -0.42, [-1.52, 0.67]). Three of the studies reported positive effects; however, the confidence intervals for all of those cross the line of no effect (Hartman et al. 2018; Schiller 2012; Svensson and Brouneus 2013). The considerable heterogeneity in the analysis (I^2 = 98.88%) is driven by the effect of one study. Again, it is the sizeable significant negative effect of the dialogue intervention in Sri Lanka that differs from the trend of the other studies (Lonergan 2016). The study further identified differential effects by population. The study reported a large negative impact on students from the ethnic minority (g = -2.78, [-3.29, -2.25]), but not on majority students (g = 0.16, [-0.29, 0.61]).

Our sensitivity analysis indicates that the main average effect is not affected by the removal of any of the studies. However, the removal of the Lonergan (2016) study does substantially reduce heterogeneity (I^2 = 34.04%).

Figure 10: Forest plot of effects on acceptance of diversity among intergroup dialogue interventions
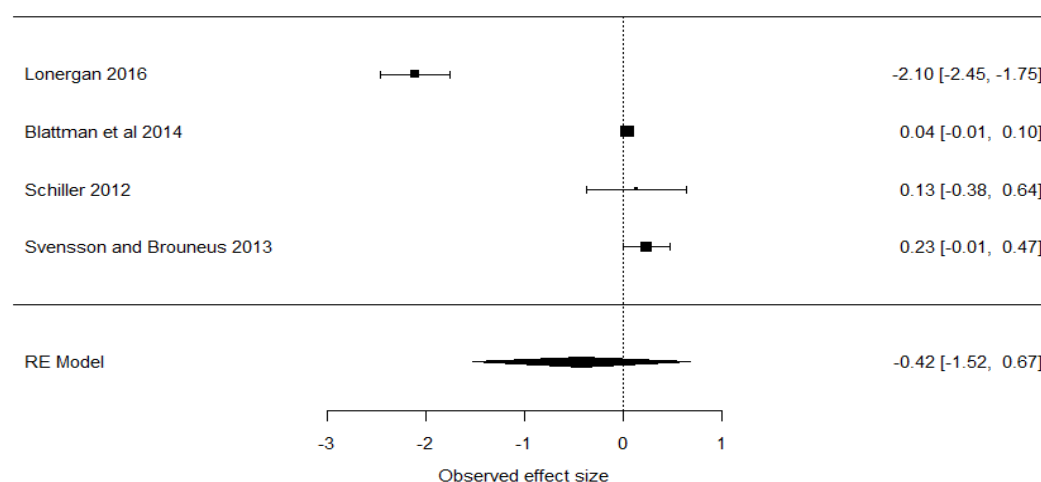
4.4.5 Workshop-based peace education with intergroup contact and economic support

Five studies on workshop-based peace education with intergroup contact and economic support interventions (workshops-contact-econ) reported effect sizes on social cohesion outcomes. Due to the greater overlap in outcome constructs reported by studies in this category, we were able to conduct meta-analyses for three of the social cohesion outcome groups (trust, sense of belonging and willingness to participate). We rated four of the five studies as having high risk of bias (Causal Design 2016; Dawop et al. 2019; Finkel et al. 2018; IMPAQ International 2017). As such, the results of the analysis should be treated with particular caution. The results of the meta-analyses indicate that these interventions have a postive effect on trust; for sense of belonging and willingness to participate outcomes, the meta-analysis finds effects in a positive direction, but the 95 per cent confidence interval crosses the line of no effect. Individual effect sizes related to the outcomes of willingness to help and an acceptance of diversity also reveal no significant effect of these interventions, although two studies found a positive effect on index measures of social cohesion outcomes (Ferguson 2019; IMPAQ International 2017).

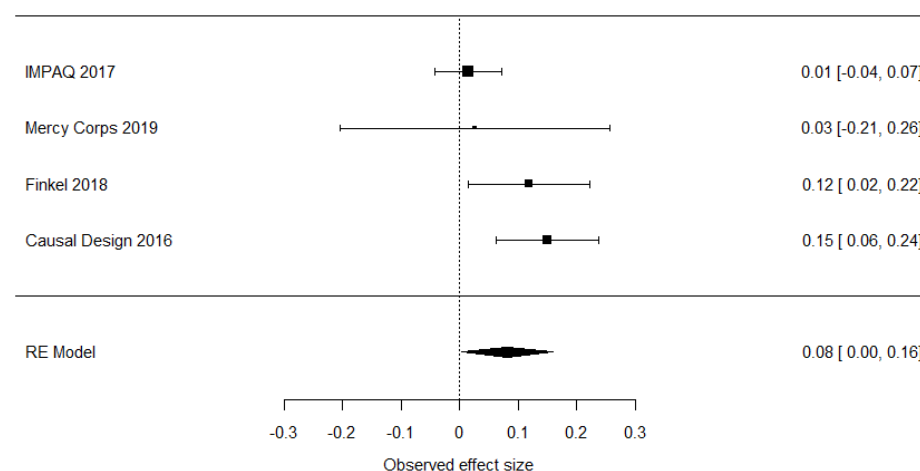**Trust**

Four studies contributed effect sizes to the meta-analysis of the average effect on trust outcomes of workshops-contact-econ interventions. One of these interventions took place in Burkina Faso (Finkel et al. 2018), one in Bosnia and Herzegovina (IMPAQ International 2017), and two in Nigeria (Causal Design 2016; Dawop et al. 2019). The outcomes included in the analysis are all measures of participants' levels of trust for others. Three of them measured their trust for out-groups, while the study in Burkina Faso measured generalised trust for others (Finkel et al. 2018).

As can be seen from Figure 11, the meta-analysis suggests a small positive effect (g = 0.08, [0.00, 0.16]). Although all of the studies report effects in a positive direction, there is moderate heterogeneity (I^2 = 58.59%). Our sensitivity analysis indicates that the heterogeneity is driven by the study from Bosnia and Herzegovina (IMPAQ International 2017). When that study is dropped, the effect size increases (g = 0.13, [0.06, 0.19]) and the statistical test for heterogeneity shows a substantial drop (I^2 = 0.00%).

Figure 11: Forest plot of effects on trust among workshops-contact-econ interventions

Sense of belonging

Three studies contributed effect sizes to the meta-analysis of the average effect of workshops-contact-econ interventions on sense of belonging outcomes. The interventions took place in Burkina Faso, Jordan and Nigeria. The outcomes included in the analysis are all measures of intergroup divisions.

As can be seen from Figure 12, the meta-analysis suggests a small positive effect but the confidence interval crosses the line of no effect (g = 0.10, [-0.01, 0.21]). All of the studies reported small positive effects but the effect sizes are only statistically significant for interventions in Burkina Faso and Nigeria (Finkel et al. 2018; Casual Design 2016). The effect size for the study in Jordan is very small in magnitude but is given substantial weight in the model due to its relative precision versus the other two studies (Ferguson 2019). This creates substantial heterogeneity in the model (I^2 = 82.18%).

Our sensitivity analysis indicates that the main average effect is affected by the removal of Ferguson (2019) and leads to a small, positive and significant main average effect (g = 0.15, [0.07, 0.23]). This removal, however, reduces the number of studies in the analysis to only two, which is too low for placing confidence in the results of the meta-analysis.

Figure 12: Forest plot of effects on sense of belonging outcomes among workshops-contact-econ interventions

Willingness to participate
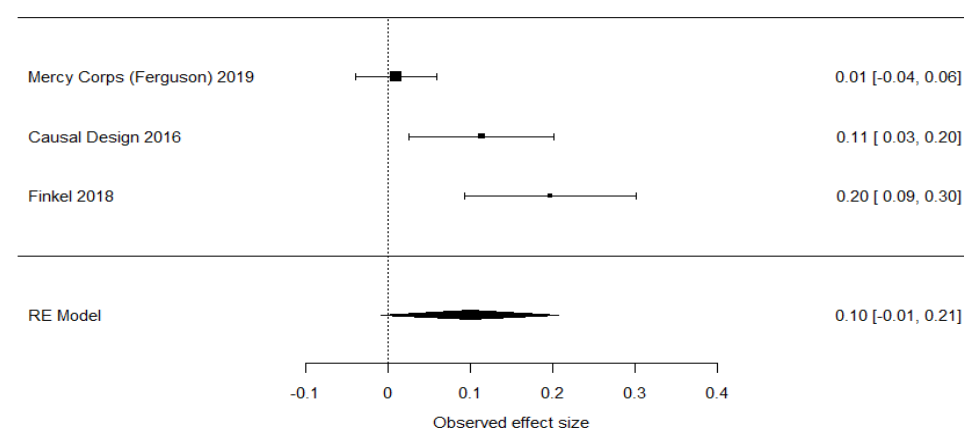
Four studies contributed effect sizes to the meta-analysis of the average effect of workshops-contact-econ interventions on willingness to participate outcomes. The interventions took place in Burkina Faso, Jordan and Nigeria, with two occurring in Nigeria. The outcomes included in the analysis are three behavioural measures of intergroup interactions and one measure of perceptions of whether ordinary community members participate in local decision-making (Finkel et al. 2018).

As can be seen from Figure 13, the meta-analysis suggests a small positive effect, but the confidence interval crosses the line of no effect (g = 0.06, [-0.05, 0.16]). Most of the studies reported small positive effect sizes but only one of those positive effect sizes does not cross the line of no effect. This is one of the studies in Nigeria (Causal Design 2016). The study in Jordan is the only one to report a small negative effect, but it also crosses the line of no effect (Ferguson 2019). The effect size for the study in Jordan is very small in magnitude but is given substantial weight in the model due to its relative precision versus the other two studies (Ferguson 2019). There is a substantial amount of heterogeneity in the model (I^2 = 77.77%).

Our sensitivity analysis indicates that the main average effect is not affected by the removal of any of the studies.

Figure 13: Forest plot of effects on willingness to participate among workshops-contact-econ interventions
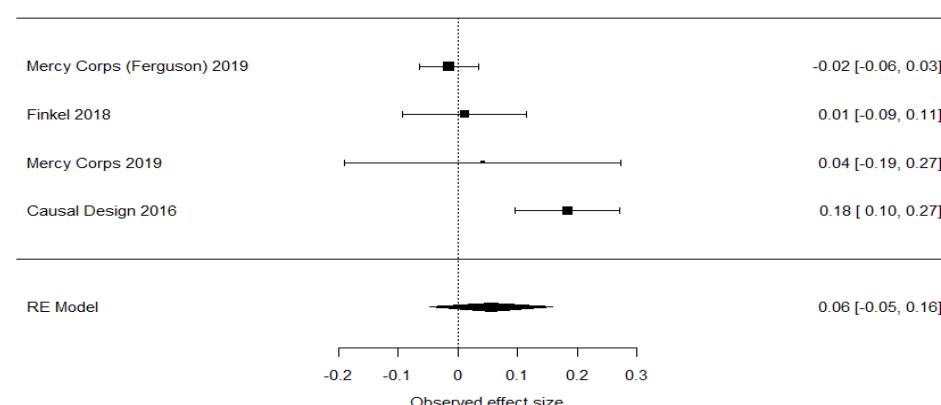
Willingness to help

Only one study within this intervention group measured effects on outcomes related to willingness to help (Dawop et al. 2019). The study of the workshops-contact-econ intervention in Nigeria found only very small, statistically insignificant effects on an intergroup measure of respondents' willingness to help others (g = 0.01, [-0.22, 0.24]).

Acceptance of diversity

One study in Nigeria measured intergroup bias and feelings of negativity towards the out-group through an indirect list experiment, and found no impact of the intervention (g = 0.02, [-0.08, 0.12]) (Dawop et al. 2019). One study in Burkina Faso reported a measure of acceptance of diversity (Finkel et al. 2018). The study found no effect of the intervention on acceptance of inter-ethnic marriage (g = 0.03, [-0.08, 0.13]).

Combined social cohesion outcomes measured as an index

One study of a workshops-contact-econ intervention in Bosnia and Herzegovina also reported an effect size for an index that combined several of the social cohesion outcome categories (IMPAQ International 2017). The measure was an index of inter-ethnic attitudes, including measures of willingness to participate, acceptance of diversity, forgiveness and empathy. There was no effect of the intervention on that index measure (g = 0.01, [-0.04, 0.07]).

4.4.6 Media for peace interventions

Five studies on media for peace interventions reported effect sizes on social cohesion outcomes. Due to the lack of overlap in outcome constructs reported by studies in this category, we were only able to conduct meta-analyses for two of the social cohesion outcome groups (trust and acceptance of diversity). There were, however, several different types of acceptance of diversity constructs reported across different studies that allowed us to conduct several separate meta-analyses of that outcome category. We rated three studies as having high risk of bias (Bilali et al. 2016; Bilali and Vollhardt 2015; Bilali and Vollhardt 2013). We rated the two others as having some concerns related to risk of bias (Paluck 2009; Vicente and Vilela 2020). As such, the results of the analysis including the first three studies should be treated with particular caution.

The results of the meta-analyses indicate that there is a small, positive and significant average effect of media for peace interventions on trust outcomes. However, there is no effect of these interventions on any of the four different acceptance of diversity measures that we examined. The individual effect sizes analysed on other social cohesion outcomes are mixed and inconclusive. For sense of belonging outcomes, the effect sizes report contradictory effects of the interventions; both effects are small and significant but one is positive and the other is negative. For the willingness to participate outcomes that are reported, one effect size indicates no effect of these interventions, while another suggests a negative effect of the intervention. In contrast, for willingness to help outcomes, one effect size indicates no effect of these interventions, while another suggests a positive effect of the intervention.

Trust

Three studies contributed effect sizes to the meta-analysis of the average effect of media for peace interventions on trust outcomes. One of these interventions took place in Burundi and the other two were implemented in Rwanda. The outcomes included in the analysis are all measures of mistrust. Two of those measures focus on intergroup mistrust (Bilali et al. 2016; Bilali and Vollhardt 2013) and one measures generalised trust (Paluck 2009).

As can be seen from Figure 14, the meta-analysis suggests a small, positive and statistically significant average effect (g = 0.10, [0.02, 0.18]). The positive effect is driven by two of the studies. All three of the studies reported positive effects but only one of those effects does not cross the line of no effect. The study with the statistically significant effect is the one that took place in Burundi and measures levels of intergroup mistrust (Bilali et al. 2016). As can be deduced from a visual inspection, heterogeneity in the model is minimal (I^2 = 0.00%).

Our sensitivity analysis indicates that the main average effect is affected by the removal of one of the studies. When the study by Bilali and colleagues (2016) is removed from the analysis, the main average effect remains positive but the confidence interval now crosses the line of no effect (g = 0.07, [-0.03, 0.18]).

Figure 14: Forest plot of effects on trust outcomes for media for peace interventions

Bilali and Vollhardt 2013	0.04 [-0.09, 0.18]
Paluck 2009	0.13 [-0.05, 0.31]
Bilali et al 2016	0.14 [0.02, 0.26]
RE Model	0.10 [0.02, 0.18]

Observed effect size
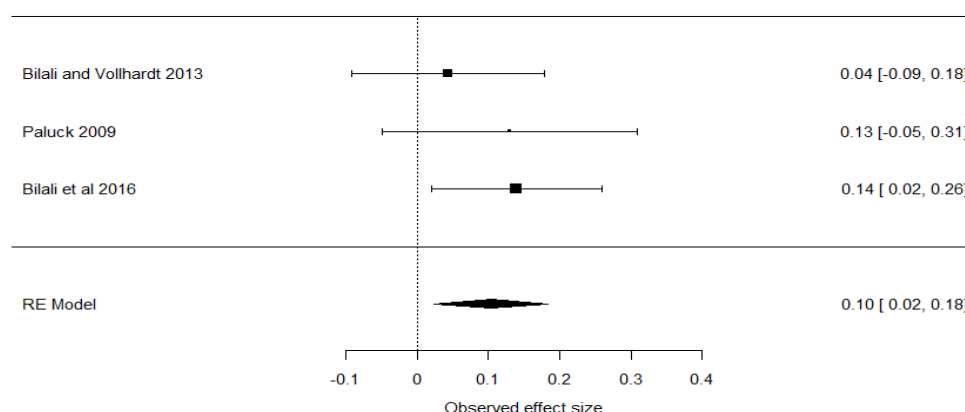
Sense of belonging

Only two studies reported effect sizes of media for peace interventions on sense of belonging outcomes, so a useful meta-analysis could not be conducted. These two studies focused on interventions that took place in the DRC and Rwanda. They each measured intergroup belonging differently. The study of the intervention in the DRC found a negative and significant effect corresponding to an increase in support for political exclusion of out-groups (g = -0.15, [-0.25, -0.05]) (Bilali and Vollhardt 2015). In contrast, the study of the intervention in Rwanda reported a positive and significant effect of the intervention on a measure of empathy for other Rwandans (g = 0.19, [0.01, 0.37]) (Paluck 2009).

Willingness to participate

Only two studies reported effect sizes of media for peace interventions on willingness to participate outcomes, so a useful meta-analysis could not be conducted. These two studies focused on interventions that took place in the DRC and Rwanda. They each measured intergroup participation differently. The study of the intervention in the DRC found a small, negative and significant effect of the intervention on a measure of support for political exclusion in terms of people from different groups getting together to discuss politics (g = -0.15, [-0.25, -0.05]) (Bilali and Vollhardt 2015). In contrast, the study of the intervention in Rwanda reported no effect of the intervention on a measure of willingness to affiliate with other groups (g = -0.04, [-0.21, 0.13]) (Paluck 2009).

Willingness to help

Only two studies reported effect sizes of media for peace interventions on two very different constructs of willingness to help outcomes, so a useful meta-analysis could not be conducted. These two studies focused on interventions that took place in Burundi and Mozambique. They each measure intergroup willingness to help differently. The study of the intervention in Burundi found a positive and significant effect of the intervention on a measure of active bystandership (g = 0.12, [0.00, 0.24]) (Bilali et al. 2016). In contrast, the study of the intervention in Mozambique reported no effect of the intervention on a

measure of destroying another person's endowment in a lab game (g = -0.09, [-0.21, 0.04]) (Vicente and Vilela 2020).

Acceptance of diversity

We conducted four different meta-analyses to examine the effect of media for peace interventions on the acceptance of diversity. Among the acceptance of diversity outcomes reported by the studies for media for peace interventions, there were several different outcome indicators that were reported in a comparable way across enough studies to contribute to separate meta-analyses. Several of these analyses were sensitive to the removal of the same study, which examined the effect of an intervention in the DRC where there is ongoing conflict (Bilali and Vollhardt 2015). In contrast, the other studies in the meta-analysis focus on more historical conflicts.

One such grouping was for measures of *intergroup tolerance*, which all examined attitudes towards cross-group marriage. Three studies contributed effect sizes to that meta-analysis. The interventions took place in Burundi, the DRC and Rwanda.

As can be seen from Figure 15, the meta-analysis suggests a positive average effect but the confidence interval crosses the line of no effect (g = 0.26, [-0.13, 0.66]). The positive effect is driven by two of the studies. The effect sizes reported by the studies in Burundi and Rwanda are both positive and their confidence intervals do not cross the line of no effect (Bilali et al. 2016; Paluck 2009). The study focused on the DRC reported a negative effect but the confidence interval crosses the line of no effect (Bilali and Vollhardt 2015). As can be deduced from a visual inspection, heterogeneity in the model is considerable (I^2 = 96.45%).

Our sensitivity analysis indicates that the main average effect is affected by the removal of one of the studies. When Bilali and Vollhardt (2015) is removed from the analysis, the main average effect becomes positive and significant (g = 0.42, [0.02, 0.83]).

Figure 15: Forest plot of effects on intergroup tolerance for media for peace interventions
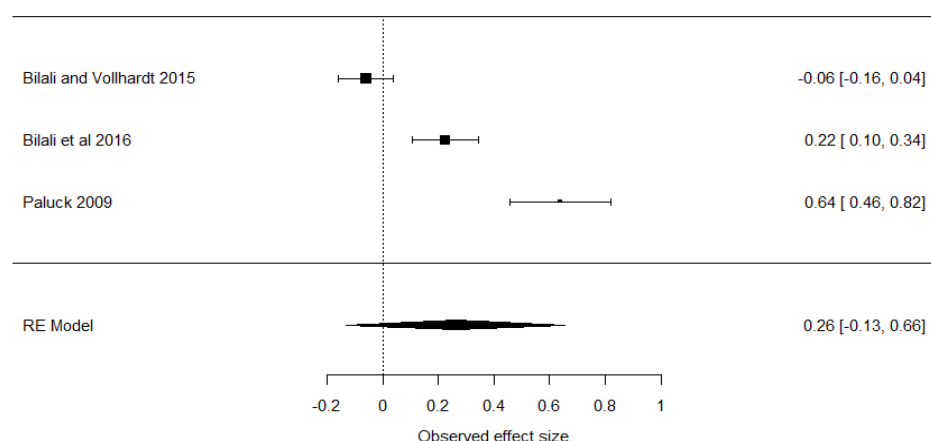
The second grouping was for measures of *rejection of multiple perspective*s, related to perceptions of the dangers or confusion of allowing multiple points of view to be voiced. Three studies contributed effect sizes to that meta-analysis. The interventions took place in Burundi, the DRC and Rwanda.

As can be seen from Figure 16, the meta-analysis suggests a small positive average effect but the confidence interval crosses the line of no effect (g = 0.07, [-0.02, 0.17]). The positive direction is driven by two of the studies. The effect sizes reported by the studies in Burundi and Rwanda are both positive; however, only the effect from the study in Burundi is significant (Bilali et al. 2016). The study focused on the DRC reported a negative effect but the confidence interval crosses the line of no effect (Bilali and Vollhardt 2015). Heterogeneity in the model is moderate ($I^2 = 44.80\%$).

Our sensitivity analysis indicates that the main average effect is affected by the removal of one of the studies. When Bilali and Vollhardt (2015) is removed from the analysis, the main average effect becomes positive and significant (g = 0.13, [0.04, 0.22]).

Figure 16: Forest plot of effects on rejection of multiple perspectives among media for peace interventions
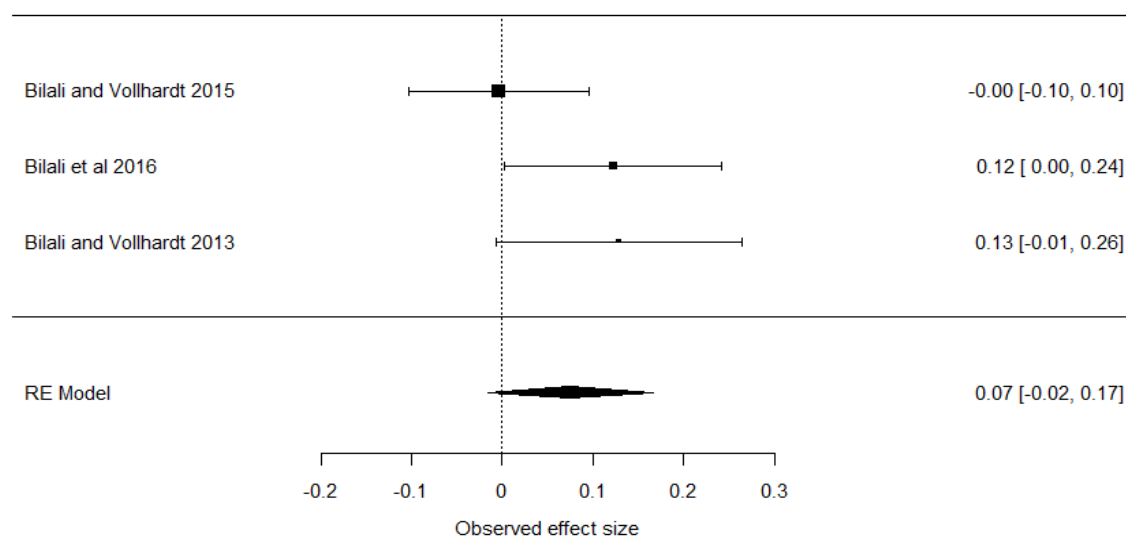
The third grouping was for measures of *exclusive victimhood*, which relates to the perception that one's own group has suffered more than other groups. Three studies contributed effect sizes to that meta-analysis. The interventions took place in Burundi, the DRC and Rwanda.

As can be seen from Figure 17, the meta-analysis suggests a small, positive average effect but the confidence interval crosses the line of no effect (g = 0.07, [-0.06, 0.21]).The positive direction is driven by two of the studies. The effect sizes from the studies in Burundi and Rwanda are both positive; however, only the effect from the study in Burundi is significant (Bilali et al. 2016). The study from the DRC reported a negative effect but the confidence interval crosses the line of no effect (Bilali and Vollhardt 2015). Heterogeneity in the model is substantial ($I^2 = 74.13\%$).

Our sensitivity analysis indicates that the main average effect is affected by the removal of one of the studies. When Bilali and Vollhardt (2015) is removed from the analysis, the main average effect becomes positive and significant (g = 0.14, [0.05, 0.23]).

Figure 17: Forest plot of effects on feelings of exclusive victimhood among media for peace interventions
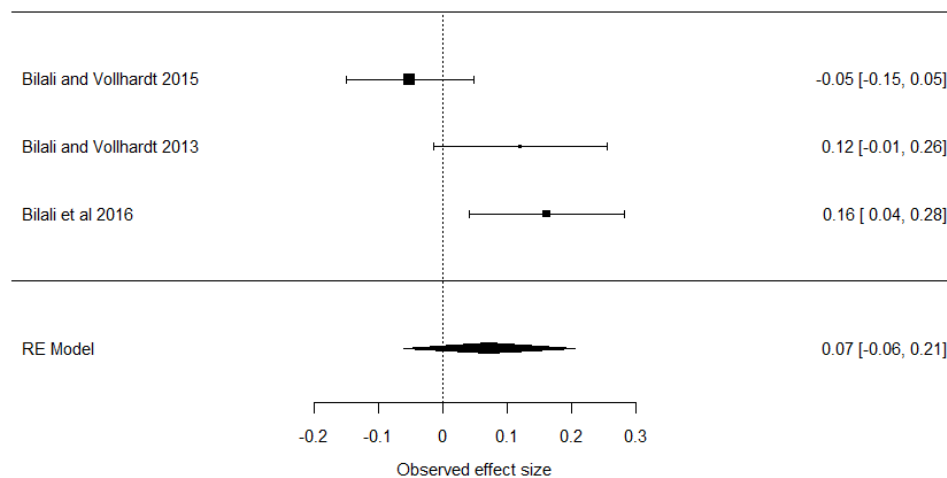
The fourth and final grouping was for measures of *inclusive victimhood*, which is similar to the previous construct but refers to the perception that other groups have similar amounts of persecution to one's own group. Three studies contributed effect sizes to that meta-analysis. The interventions took place in Burundi, the DRC and Rwanda.

As can be seen from Figure 18, the meta-analysis suggests a small, positive average effect but the confidence interval crosses the line of no effect (g = 0.07, [-0.02, 0.17]). The positive direction is driven by two of the studies. The effect sizes reported by the studies in the DRC and Rwanda are both positive; however, only the effect from the study in the DRC is significant (Bilali and Vollhardt 2015). The study focused on Burundi reported a negative effect but the confidence interval crosses the line of no effect (Bilali et al. 2016). Heterogeneity in the model is moderate (I^2 = 52.68%).

Our sensitivity analysis indicates that the main average effect is affected by the removal of one of the studies. When Bilali et al. (2016) is removed from the analysis, the main average effect becomes positive and significant (g = 0.12, [0.04, 0.20]).

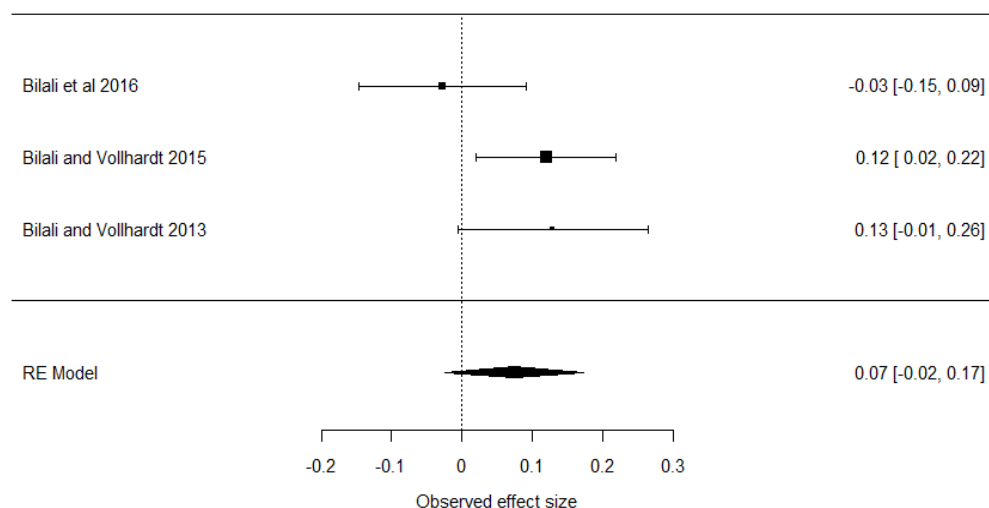**Figure 18: Forest plot of effects on feelings of inclusive victimhood among media for peace interventions**

4.5 Quantitative analysis of additional outcomes

In addition to effects on social cohesion outcomes, we further extracted data on all other outcomes reported by the included studies to answer Review Question 2. Our aim was to understand the extent to which studies analysed the relationship between intergroup social cohesion interventions and outcomes of resilience and sustainable peace. We identified three broad categories of additional outcomes reported in included studies: intermediate outcomes supporting resilient social cohesion; vertical cohesion outcomes; and human security outcomes.

The first category includes outcomes that are not social cohesion outcomes themselves, but are expected to serve as either intermediate or mutually reinforcing complementary outcomes along the causal chain.[22] The second category covers outcomes related to the relationships between citizens and the state, referred to as vertical cohesion.[23] Finally, the third category comprises indirectly related outcomes of different dimensions of human security. In this context, human security refers to an approach to security that posits that quality of life relies on freedom from threats across diverse dimensions, including economic threats, environmental threats, political threats, as well as threats to physical security (Giessmann et al. 2019). Given the inconsistency with which these outcomes are reported across studies, analysis is undertaken at a review level, incorporating studies from all intervention groups together.

4.5.1 Effects on intermediate outcomes supporting resilient social cohesion
This section reports the findings for five different outcomes that are theorised to form part of the causal chain towards building resilient intergroup social cohesion. They are roughly grouped in order of their placement along the causal chain, although it is important to note that individual outcomes may occur earlier or later depending on the context and intervention activities. Broadly, we move here from knowledge outcomes, to outcomes of key skills and behaviours, to changes in attitudes and beliefs. For each section, we report findings for outcomes where at least two studies reported effects to enable comparison.[24]

---

[22] In order to code these outcomes, we drew primarily on a database of peacebuilding indicators developed by the Alliance for Peacebuilding (Baumgardner-Zuzik et al. 2020). The theory is that these outcomes contribute to building resilient intergroup social cohesion. Given the relevance of these outcomes to possibly explaining the findings of Review Question 1, we focus in this section on the outcomes in this group.

[23] We include in this review interventions that, in addition to their primary goal of fostering intergroup social cohesion, also engage local government in activities or encourage citizens to engage more with the government. As a result, there are some studies reporting outcomes for vertical cohesion. As with the horizontal social cohesion outcomes above, these are grouped first by the dimension of social cohesion to which they referred (e.g. trust, willingness to participate), and then into specific constructs drawing primarily on the typology used for the Freedom in the World reports (Freedom House 2020). While we summarise the findings from this analysis below, full results are reported in Online appendix B.6.

[24] Intermediate outcomes for which only one study reported findings include: sense of community capacity to realise changes (Paluck 2009); social inclusion (Alan et al. 2020); social capital (Cilliers et al. 2018); and outcomes forming a reconciliation index (Lonergan 2016).

Summary of effects on intermediate outcomes

The effects of intergroup social cohesion interventions on the intermediate outcomes assessed in this section are typically small and statistically insignificant. There is some evidence of positive impacts on knowledge of intergroup sociocultural awareness when a shorter-term collaborative contact intervention is dropped from the analysis. There is no evidence of effects on understandings of concepts of peace and violence. Notably, in two contexts where the radio drama interventions and outcome constructs regarding understanding violence as a continuum were very similar due to being implemented by the same organisation, the effects were in opposite directions (Bilali et al. 2016; Paluck 2009).

Overall, we find limited to no effects on measures of social and emotional skills that may strengthen the conditions for intergroup cohesion. In one study, forgiveness of perpetrators increased, yet it was marked by a negative effect on mental health outcomes (discussed below) (Cilliers et al. 2018). In another instance, a negative effect on confidence was identified, particularly for girls (Aladysheva et al. 2017). The authors suggest that this may be indicative of a nonlinear causal chain for such interventions; as students learned more about bias and learned to question their assumptions, a decrease in self-assurance may be expected. The project was relatively short in duration, and this brevity is noted as a further potential explanation: more time may be necessary to build the skills needed to help students comfortably engage with nuance and uncertainty.

We further identify limited or negative effects on dispute resolution practices. However, this finding should be treated with caution, as it is derived from measures of outcomes at different levels. Finally, we similarly find no effects on outcomes of support for violence or extremism. Again, however, the interventions and measures that feed into this analysis vary substantially.

Sociocultural awareness

We extracted effect sizes from four studies of impacts on sociocultural knowledge, which we synthesised through meta-analysis. These included: knowledge of diverse historical perspectives measured in a radio drama intervention in Rwanda (Bilali and Vollhardt 2013); knowledge of aspects of the out-group culture from a radio drama intervention in Burundi (Bilali et al. 2016) and from a collaborative contact intervention in Nigeria (Scacco and Warren 2018); and a measure of knowledge of current affairs related to out-group contexts from another collaborative contact intervention in Nigeria (Okunogbe 2018). In this sense, these outcomes likely represent an initial step in the causal chain towards acceptance of diversity. Although being aware of the culture, politics or history of another group may not in and of itself be a measure of acceptance of diversity, it is logical that ignorance of the same could serve as a barrier to accepting diversity.[25]

---

[25] For the purposes of this review, we tried to differentiate between knowledge outcomes that are measures of social cohesion from those that may form the first step in the causal chain. For example, we coded outcomes of basic knowledge of out-group culture as measures of sociocultural awareness, but outcomes that measured recognition of the validity of others' culture as acceptance of diversity, such as understanding and recognising why out-groups would want the right to religious education or prayer.

In many ways, all of the studies included in this review aimed to build knowledge of the other group, directly aiming to trigger the mechanism described above, 'seeing the other', to help participants move towards recognising in the other group a shared sense of identity as compatriots or even simply as humans. In most cases, this was explicitly done through the training components included within the interventions. In the case of the radio dramas, the training was not in person, but a core goal of the 'edutainment' interventions was to familiarise citizens with each other (and indeed, studies of radio dramas report two of the outcomes on sociocultural awareness identified in the review). Even where interventions did not explicitly aim to teach participants about the other group, which was the case in three collaborative contact interventions, the theories of change were based on an assumption that participants would build knowledge of each other by working collaboratively together, without needing to be taught. Therefore, knowledge still forms a key part of the causal chain. As such, it is striking that only four studies measured outcomes related to knowledge of the other group.

As can be seen from Figure 19, the meta-analysis suggests a positive effect, but the confidence interval crosses the line of no effect (g = 0.11, [-0.02, 0.24]), and there is substantial heterogeneity (I^2 = 70.56%) in the estimation. Our sensitivity analysis indicates that this is partially driven by a single study (Scacco and Warren 2018) reporting a small negative effect, and the average affect becomes more precise when this study is dropped (g = 0.15, [0.01, 0.29]), although heterogeneity remains substantial (I^2 = 70.52%). The impact appears particularly positive for the two radio drama programmes (Bilali et al. 2016; Bilali and Vollhardt 2013).

Figure 19: Forest plot of effects on sociocultural awareness
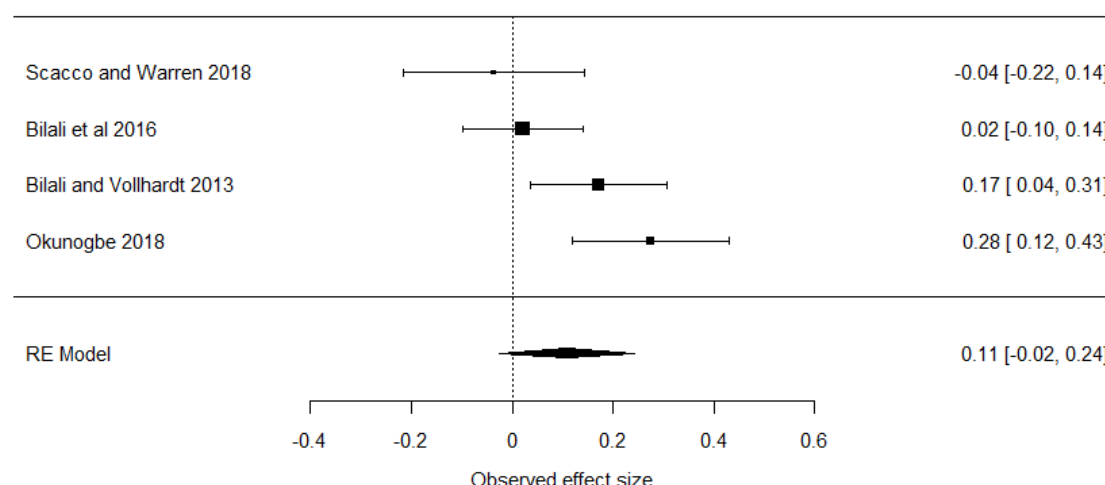
Knowledge of peace and conflict
We extracted effect sizes from four studies of impacts on participants' knowledge of concepts around violence, peace and mediation, which we synthesised through meta-analysis. These included: understanding violence as a continuum, measured in radio drama interventions in Rwanda (Paluck 2009) and Burundi (Bilali et al. 2016); whether students in Kyrgyzstan could correctly identify and define mediation skills (Aladysheva et al. 2017); and an index measure of students' understanding of positive peace in Palestine – i.e. that peace is more than the absence of conflict (Biton and Salomon 2006).

As can be seen from Figure 20, the meta-analysis suggests a small negative effect that is not statistically different from zero (g = -0.03, [-0.14, 0.08]). There is minimal heterogeneity (I^2 = 37.98%) in the estimation, and our sensitivity analysis indicates that the main average effect is not affected by the removal of any of the studies.

Figure 20: Forest plot of effects on knowledge of peace and conflict concepts

Bilali et al 2016   -0.14 [-0.26, -0.02]

Aladysheva 2017   0.00 [-0.21, 0.21]

Bilton & Salomon 2006   0.03 [-0.20, 0.25]

Paluck 2009   0.07 [-0.11, 0.25]

RE Model   -0.03 [-0.14, 0.08]

-0.3   -0.2   -0.1   0   0.1   0.2   0.3

Observed effect size
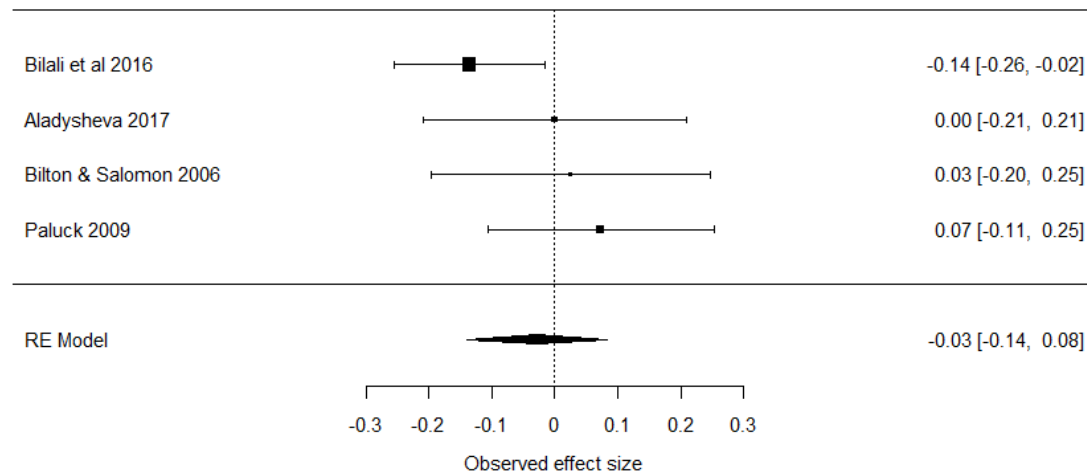
Social and emotional skills

Six studies reported effects on outcomes theorised to strengthen an individual's socio-emotional well-being and openness to intergroup cohesion, including empathy, forgiveness and perceived self-efficacy (confidence).

Two studies reported measures of empathy for out-groups. An intergroup dialogue intervention in Liberia (Hartman et al. 2018) reported a small, statistically insignificant but positive effect (g = 0.03, [-0.03, 0.09]). A collaborative contact intervention in Nigeria (Scacco and Warren 2018) reported a similarly small, insignificant but negative effect (g = -0.02, [-0.2, 0.15]). The small number of studies reporting effects on empathy is striking given how many of the studies noted building empathy for the other as a core part of the causal chain. Along with perspective-taking, it is frequently noted as a cornerstone of peace education (Harris and Morrison 2013).

Two studies evaluating intergroup dialogue interventions measured effects on feelings of forgiveness towards the other group. The study of an intergroup dialogue intervention bringing together individuals affiliated with groups on both sides of the conflict in Indonesia identified a small positive effect, but the confidence interval crosses the line of no effect (g = 0.13, [-0.38, 0.64]). The study from Sierra Leone (Cilliers et al. 2018) reported the effects of a truth and reconciliation process and found a small and statistically significant positive effect on forgiveness of perpetrators (g = 0.09, [0.00, 0.18]).

Finally, two studies reported measures of self-efficacy in youth. The school-based peace education programme in Kyrgyzstan found a statistically significant and negative effect on self-reported feelings of confidence (g = -0.2, [-0.34, -0.05]). Heterogeneity analysis suggested that this effect was driven by a negative, statistically significant impact on confidence among girls who participated in the intervention (g = -0.28, [-0.46, -0.09]); although the finding for boys was also negative, it was statistically insignificant (g = -0.11, [-0.34, 0.13]). There were no significant differences in effects by ethnicity. The collaborative

contact study of intercommunity volunteering in Lebanon found a small, positive but insignificant effect on youth's self-reported feelings of confidence (g = 0.07, [-0.07, 0.21]).

Dispute resolution

Five studies reported effects on respondents' perceptions of dispute resolution practices and success in their communities, including three workshops-contact-econ studies and two intergroup dialogue studies. The workshops-contact-econ studies were each implemented by Mercy Corps, two in Nigeria's Middle Belt (Causal Design 2016; Dawop et al. 2019) and one in Jordan (Ferguson 2019). One of the intergroup dialogue studies was of the Sierra Leone truth and reconciliation process (Cilliers et al. 2018), while the other was based in fragile communities in Liberia (Hartman et al. 2018). Both the Liberia study and the studies from Nigeria's Middle Belt focused on conflicts over land and resources. Some measures were at the individual level, including personal experiences with disputes and satisfaction with dispute resolution outcomes (Cilliers et al. 2018) and assessing self-reported behaviours of dispute resolution (Hartman et al. 2018). Other measures focused on perceptions of the capabilities of one's community to successfully resolve disputes (Causal Design 2016; Ferguson 2019; Dawop et al. 2019).

As reported in Figure 21, the meta-analysis suggests a very small positive effect that is not statistically different from zero (g = 0.01, [-0.11, 0.13]). Further, heterogeneity is substantial (I^2 = 86.60%), which is confirmed from a visual analysis of the forest plot. Our sensitivity analysis suggests that this is primarily driven by the effect from Causal Design (2016). When this study is removed, the average effect becomes negative and statistically significant (g = -0.04, [-0.08, -0.001]), and heterogeneity is substantially reduced (I^2 = 0.00%).

Figure 21: Forest plot of effects on dispute resolution practices
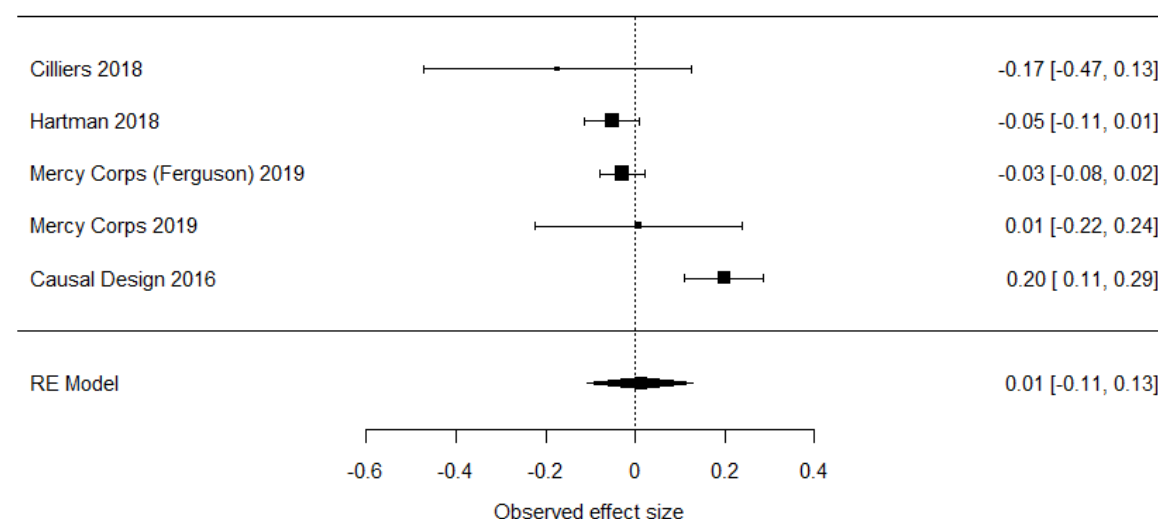
Support for violence and extremism

Five studies contributed effects to a meta-analysis of expressions of support for violence or extremism. These include two workshops-contact-econ studies, in Jordan (Ferguson 2019) and Burkina Faso, Chad and Niger (Finkel et al. 2018), which reported perceptions of acceptability of using violence against out-groups or against civilians to defend one's religion, respectively. Two intergroup dialogue studies from Sri Lanka (Lonergan 2016;

Cilliers et al. 2018) reported measures of acceptability of violence against out-groups and openness to returning to fight in violent conflict again, respectively. Finally, one study of media campaigns reported an index measure of support for violent extremism (Vicente and Vilela 2020).

As shown in Figure 22, the meta-analysis found no impacts on these outcomes (g = -0.00, [-0.04, 0.04]). Heterogeneity is minimal (I^2 = 0.20%), and our sensitivity analysis indicates that the average effect did not change significantly with the removal of any of the studies.

Figure 22: Forest plot of effects on support for violence and extremism

Lonergan 2016	-0.37 [-0.72, -0.02]
Vicente and Vilela 2019	-0.04 [-0.29, 0.22]
Cilliers 2018	-0.03 [-0.10, 0.04]
Mercy Corps (Ferguson) 2019	0.02 [-0.03, 0.06]
Finkel 2018	0.02 [-0.08, 0.13]
RE Model	-0.00 [-0.04, 0.04]

Observed effect size
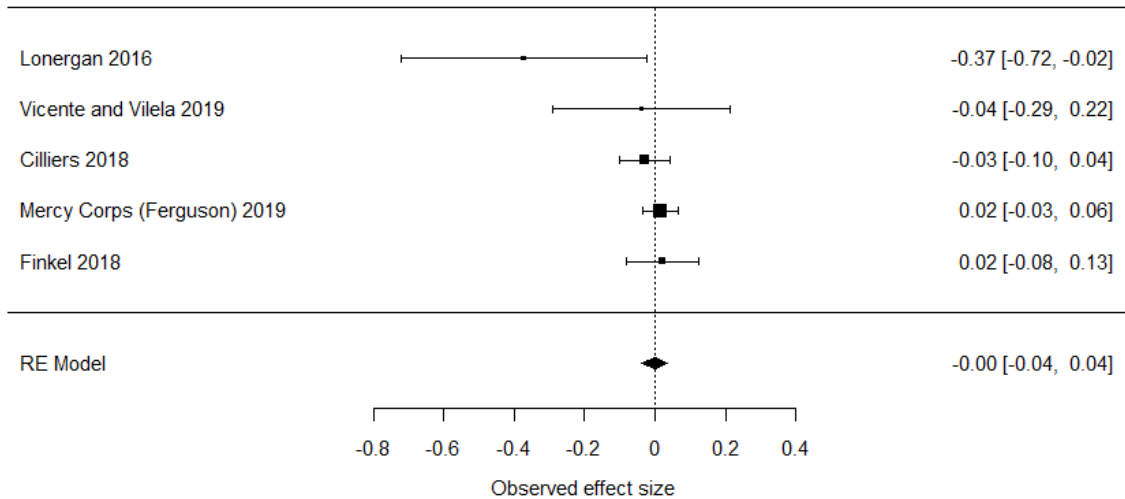
4.5.2 Effects on vertical cohesion

Seven included studies reported effects on outcomes related to vertical cohesion, studying changes in the relationships between citizens and the state that can be attributed to the intergroup social cohesion interventions implemented. Extending the framework applied to the horizontal, intergroup social cohesion outcomes, these are grouped into measures of trust, willingness to participate, and acceptance of diversity. No vertical cohesion outcomes related to sense of belonging or willingness to help were reported.

Overall, we find limited effects on outcomes of vertical social cohesion. It is important to note that few studies reported such effects, and outcomes related to vertical cohesion were a core part of the integrated programme in only one study (Finkel et al. 2018). The workshops-contact-econ intervention implemented across Burkina Faso, Chad and Niger had positive effects on an index measure of institutional trust, but no effects on civic engagement. Among other studies, there is some evidence of small positive effects on civil liberties, including lowering support for authoritarianism and greater tolerance for dissent, but it is not consistent across contexts.

4.5.3 Effects on human security outcomes

This section reports findings on outcomes related to different dimensions of human security. The intention is to understand the extent to which included studies identified relationships between intergroup social cohesion interventions and outcomes of resilience and sustainable peace, in line with Review Question 4. Included studies reported effects on measures of physical security, health security, economic security and

education security.[26] We primarily find no effects on outcomes of violence, employment or educational attainment that could be attributed to intergroup social cohesion interventions.

The most common way in which included studies incorporated efforts to address threats to a dimension of human security was through providing economic support, either through vocational training or grants for community projects or infrastructure. Seven studies incorporated elements of economic support in the interventions, but only three of them reported outcomes related to economic well-being, and, with one exception, the findings were insignificant. Most of these studies explicitly recognised that the purpose of the economic support was to facilitate participants' engagement in building intergroup social cohesion.

Few studies measured effects on mental health from participating in intergroup social cohesion interventions, and even within these, the specific outcomes measured varied widely. We found some evidence of positive effects of media for peace interventions on perceptions of the value of talking about trauma, and some evidence from a single study in Indonesia of positive effects on PTSD of an intergroup dialogue intervention. However, a different intergroup dialogue intervention appears to have worsened PTSD, anxiety and depression symptoms among participants in Sierra Leone, a worrisome finding. In the case in Indonesia, participants were individuals associated with two different parties to the conflict; neither group was purely victim or purely perpetrator. However, in Sierra Leone, the intervention comprised a truth and reconciliation process that brought together victims and perpetrators. The analysis is not disaggregated by group, but it is not unreasonable to expect that victims bore the brunt of this negative impact. As noted earlier, this study identified positive effects on forgiveness of perpetrators, suggesting that, while such interventions may help to heal relationships between groups, at an individual level there is a risk that they may harm or worsen the healing process for victims. This risk is also supported by exploratory evidence of a different truth and reconciliation process in Rwanda, where the authors identified positive outcomes at the intergroup level, but negative outcomes at the individual level for victims (Rime et al. 2011).[27]

4.6 Barriers and facilitators analysis

Although the pattern overall is one of small positive effects on social cohesion, we observe substantial heterogeneity in results. In this section, we present the findings of our analysis of the barriers and facilitators that might help to explain these variations in reported results.

As described in Section 3.3.6, we undertook a search for additional literature about each of the included studies. This search identified 25 additional documents, including implementation documentation and other texts such as case studies or communications

---

[26] As before, we only report outcomes for which at least two studies reported effects. Cilliers et al. (2018) reported effects on an index of gender equality outcomes. Paluck (2009) reported effects on public health practices, but these are used as a test of the unrelated treatment provided in control communities, which focused on public health messages.

[27] Unfortunately, we were unable to calculate effect sizes for the data reported in this study.

materials. Surprisingly, we did not identify any stand-alone qualitative evaluations or process evaluations of included studies. One study (Bilali and Vollhardt 2015) included a qualitative evaluation within the impact evaluation paper that was undertaken in an effort to understand the disappointing quantitative findings. One other study (Alaref et al. 2019) mentioned the existence of a process evaluation, but, even though all other documents relating to the programme were accessible on the World Bank's project page, the process evaluation was not available. This limited the richness of the qualitative synthesis, as qualitative components of impact evaluations tend to be limited and rarely integrated with the quantitative findings. Our analysis thus draws more heavily on synthesis of intervention and implementation characteristics than on qualitative research, and is thus more exploratory in nature. Table 7 provides an overview of the documents included in the barriers and facilitators analysis.

Table 7: Overview of evidence included in the barriers and facilitators analysis

Group	Study	Country	Included evidence
School-based peace education	Aladysheva et al. 2017	Kyrgyzstan	IE + 2 success stories / communications pieces
	Alan et al. 2020	Turkey	IE + pre-analysis plan (PAP) or protocol
	Biton and Salomon 2006	Palestine	IE only
	Cleven 2020	Bosnia and Herzegovina	IE only
Collaborative contact	Alaref et al. 2019	Lebanon	IE + 3 implementation documents
	Mousa 2019	Iraq	IE + pilot IE + protocol
	Okunogbe 2018	Nigeria	IE only
	Scacco and Warren 2018	Nigeria	2 IE papers + protocol
Intergroup dialogues	Cilliers et al. 2018	Sierra Leone	2 IE papers + protocol + 2 communications pieces
	Hartman et al. 2018	Liberia	4 IE papers
	Lonergan 2016	Sri Lanka	2 IE papers
	Rime et al. 2011	Rwanda	3 IE papers
	Schiller 2012	Indonesia	IE + 1 policy / lessons article
	Svensson and Brouneus 2013	Ethiopia	2 IE papers + policy peer-reviewed article
Workshop-based peace education with intergroup contact and economic support	Causal Design 2016	Nigeria	IE only
	Ferguson 2019	Jordan	IE + protocol + 3 implementation documents
	Finkel et al. 2018	Burkina Faso, Chad and Niger	2 IE papers + 3 communications pieces + 1 implementation document
	IMPAQ International 2017	Bosnia and Herzegovina	IE + 1 communications piece
	Dawop et al. 2019	Nigeria	IE + protocol + 1 communications piece
Media for peace	Bilali and Vollhardt 2013	Rwanda	IE + 1 implementation document + 1 meta-evaluation
	Bilali et al. 2016	Burundi	IE + 1 meta-evaluation
	Bilali and Vollhardt 2015	DRC	IE (including process / qualitative evaluation) + 1 meta-evaluation
	Paluck 2009	Rwanda	4 IE papers
	Vicente and Vilela 2020	Mozambique	IE only

To structure the analysis, we developed a framework of intervention and implementation characteristics that may influence effects, comprising: the specific components of each intervention; the process and mechanisms each aimed to trigger; the duration, size and scope of the intervention; and the population for whom effects were reported. We then reviewed each of the 37 included impact evaluation papers and additional identified literature for barriers and facilitators identified by the authors related to intervention design, implementation and analysis. Given the thinness of the additional literature identified, this step drew primarily on text from the impact evaluation papers. These were then coded and added into the framework. To enable us to differentiate between common challenges faced across programmes from those that are associated with better or worse programme effects, we also added into the framework summary findings for social cohesion outcomes for each study. We conducted both vertical (across-case) and horizontal (within-case) synthesis to identify trends in factors that may explain the results identified in Section 4.4. The framework and full set of characteristics and factors identified for each study is included in Online appendix B.5. Figure 23 presents an overview of the findings from this analysis. The blue-outlined boxes denote the basic theory of change adopted to structure the analysis: intergroup social cohesion interventions are designed and implemented in fragile contexts, triggering one or more mechanisms to realise intermediate outcomes on social and emotional skills and knowledge, ultimately leading to effects on social cohesion outcomes measured through the five dimensions. The shaded shapes present high-level summaries of the meta-analyses, with dotted lines for findings that are more exploratory due to limited data. The non-shaded shapes comprise the factors identified through the qualitative synthesis that may explain the quantitative findings. The analysis below focuses on these factors, roughly following the causal chain. We first discuss the factors related to programme design and intermediate outcomes, followed by those associated with social cohesion outcomes overall, and finally those related to specific dimensions of social cohesion. We identify at least one explanatory factor for each of the five dimensions except for willingness to help, where the limited data and heterogeneity of outcomes reported meant clear trends could not be identified.

Figure 23: Barriers and facilitators identified along the causal chain

Programmes that accurately identified local bottlenecks to intergroup social cohesion may have larger effects.

A lack of conflict assessments may be a barrier to better targeting of participants and intervention strategies

A lack of effects on intermediate outcomes may explain a lack of effects on social cohesion – but few studies measured intermediate outcomes

On average, we identified small, positive effects of intergroup social cohesion interventions on measures of social cohesion

Structural changes addressing intergroup inequalities and threats to human security may be required for larger, sustainable effects on social cohesion

Workshops-contact-econ and media for peace interventions have a positive average effect on trust

Interventions that incorporate peace education may trigger mechanisms of 'seeing the other' that facilitate effects on trust

Trust outcomes

But intergroup dialogue interventions do not

Sense of belonging outcomes

Causal chains may be non-linear; negative effects may be identified before positive effects have time to emerge

Interventions aiming to build intergroup social cohesion in fragile contexts are designed and implemented

Mechanisms of seeing, meeting, talking and/or working with 'the other' trigger intermediate outcomes (e.g. changes in social or emotional knowledge & skills)

Willingness to participate outcomes

Collaborative contact interventions have a positive average effect on willingness to participate

But workshops-contact-econ interventions do not

Levels of intergroup interaction at baseline may influence the type and scale of possible impacts

Small-scale interventions may have effects on direct participants only

Impacts on intergroup relationships among participants may emerge sooner than impacts on wider intergroup relationships, which may need more time to shift

Willingness to help outcomes

Acceptance of diversity outcomes

No intervention type on average has an effect on acceptance of diversity

Across all intervention types, we identified a small positive effect on willingness to help – but few studies measured such outcomes

Effects may be highly context-dependent

Key:
White-background shapes:
- Blue boxes correspond to high-level theory of change
- Circles correspond to factors that may facilitate effects
- Hexagons correspond to factors that may hinder effects
- Triangles correspond to factors that may moderate effects

Shaded shapes:
- Circles correspond to positive impact findings from meta-analyses
- Hexagons correspond to null impact findings from meta-analyses, i.e. breaks in the causal chain
- Dotted lines correspond to findings based on limited evidence

Source: Authors

4.6.1 Programme design and implementation factors

Programmes that accurately identified local bottlenecks to intergroup social cohesion tended to have larger and more positive effects. We identified multiple instances in which the bottlenecks to social cohesion targeted by the interventions appear to have been misaligned either with the context or with the population.

The three radio dramas present an example of the importance of context in identifying the appropriate bottleneck to intergroup relations. The three programmes based on radio dramas included in this study were each implemented by the same organisation, Radio La Benevolencija, and, as a result, were based on very similar designs and theories of change. In each case, the organisation made substantial efforts to ensure the radio drama was tailored to the local context. However, whereas the dramas in Burundi and Rwanda had positive effects on multiple measures of trust and acceptance of diversity, many of these effects were not replicated when the programme was implemented in the DRC. Bilali and Vollhardt (2015) conducted a qualitative study to understand the more disappointing findings from the DRC. While the programme was widely listened to, and interviews suggested that people found the characters relevant to their lives, it failed to trigger many of the attitude changes expected. In contrast, in Rwanda and Burundi, the drama was able to create shifts in perceptions of norms; specifically, the studies found increasing support for expressing dissent, talking about multiple perspectives and allowing intergroup marriage (Bilali and Vollhardt 2013; Bilali et al. 2016).

Qualitative evidence from interviews and focus group discussions in the DRC suggests that the different outcomes may be due to the ways in which people's interpretation of the drama interacted with the context in which they were listening to the show. The intended mechanism was that, by listening to a good example of how things 'could' work played out by characters they can relate to, listeners would see how they could apply the characters' actions to real-life situations and/or change their perceptions based on learning from the characters' experiences. In the DRC, where the context was marked by ongoing violent conflict and high levels of corruption, the actions taken by characters in the drama were not seen as realistc. Because the way in which things played out in the drama was so different from how similar situations were playing out in people's lives, listeners experienced it as a nice drama about an idealised world, rather than a practical example of change they could bring about in their own communities. In comparing the overall effects of the radio dramas in Burundi and Rwanda with those in the DRC, we find some evidence that suggests that such programmes may be more effective in contexts of latent conflict rather than in contexts of ongoing violent conflict. This suggests that the key bottleneck to intergroup social cohesion was not accurately diagnosed in the case of the DRC, and, as such, the intervention was not appropriately tailored to the context.

A lack of conflict assessments may be a barrier to better targeting of programme participants and key intervention strategies. Conflict assessments are designed to identify key actors in a context and the ways in which they interact. By identifying the relationships between actors within a context, it may be easier to identify appropriate entry points for building intergroup social cohesion. While conflict assessments are regularly used to inform the design and targeting of programming in fragile contexts, only one of the interventions included in this review noted that it was based on an in-depth conflict mapping assessment (Ferguson 2019), although one other study noted that conflict assessment training was incorporated as part of the intervention (Schiller 2012).

Conflict assessments can help to identify the underlying drivers of tensions between groups, particularly those that are not related to prejudice, experience of violence or differences in belief systems. In the case of Ferguson (2019), the conflict mapping identified perceptions of inadequate and insufficient public service delivery as a key driver of tensions between refugees and host communities in Jordan (Mercy Corps 2013). Although addressing systemic issues in public services was beyond the scope of the workshops-contact-econ project, the intervention nonetheless was designed to create perceptions of improved services through its community infrastructure component. The study authors hypothesised that significant effects on individual indicators were unlikely in the implementation and follow-up period of one year, but expected measurable impacts on index measures. The findings supported this hypothesis: the intervention had a positive effect on a combined index of perceptions around local services, violence and intergroup relationships (g = 0.07, [0.02, 0.11]).

In contrast, analysis of qualitative data from IMPAQ International (2017) found that participants in the workshops-contact-econ intervention in Bosnia Herzegovina felt the programme was 'preaching to the choir', providing actors willing to engage in intergroup social cohesion activities with skills and opportunities they already had and failing to effectively engage leaders and organisations that were seen as driving divisions across groups. Both of these programmes relied on participants' willingness to engage in intergroup activities; the case of Ferguson (2019) shows that self-selection into programmes need not be a barrier to effects, as the conflict mapping identified a salient driver of tensions that the intervention was subsequently designed to alleviate.

A lack of substantive changes in intermediate social cohesion outcomes may be a barrier to larger improvements in final social cohesion outcomes. As noted in Section 4.5.1, few studies measured effects on intermediate outcomes that may be necessary for realising impacts on social cohesion outcomes. For example, few studies measured outcomes of changes in knowledge, although many incorporated elements of peace education. Limited reporting of intermediate outcomes, combined with infrequent mapping of evaluation results along the causal chain, make it difficult to identify where in the causal chain breaks occur and interpret whether the break was due to a failure of implementation or programme design, or a failure of the theory.

Two studies from Nigeria demonstrate the value of identifying and measuring effects on early-stage outcomes. In one case, Scacco and Warren (2018) evaluated a collaborative contact intervention, where the basic theory of change was that by bringing young, at-risk Christian and Muslim men together in a very collaborative vocational training course, they would get to know each other through the social and collaborative interactions, which would then encourage them to change their attitudes and behaviours. They found primarily no effects on social cohesion outcomes, but also measured effects on sociocultural knowledge and empathy, potential early steps in the causal chain for building intergroup social cohesion – where they also found no effects. This may suggest that the intervention duration or intensity were insufficient to realise the outcomes.

In contrast, Okunogbe (2018) also measured the effects of a collaborative contact intervention in Nigeria, yet one in which treated participants spent a year immersed and volunteering in a province with a different ethnic majority from their own. The study found both a positive effect on sociocultural awareness and positive effects on willingness to

participate and social cohesion index measures. The four-month intervention evaluated by Scacco and Warren (2018) was perhaps not long enough or intense enough to ensure participants actually got to know each other through the vocational training activities. The only social cohesion outcome on which the intergroup treatment arm of this short-term intervention identified effects was openness to giving help, measured through in-game behaviour (g = 0.12, [0.06, 0.17]). The authors interpret this outcome as an intermediate outcome, which in turn leads them to suggest that their findings lend some support to the 'feet-first' theory that behaviours may change before attitudes.

4.6.2 Factors associated with social cohesion outcomes overall
Intergroup social cohesion interventions alone may not be sufficient for large-scale, sustainable social cohesion. The studies included in this review evaluated interventions conducted in fragile contexts, characterised by inequalities across social groups and moderate to severe threats to individuals' human security. A limitation of our review inclusion criteria is that the interventions included are those that directly address intergroup social cohesion, and not those that aim to do so via targeting underlying drivers of tensions. However, when it comes to building sustainable peace, without broader structural changes there are limits to the effects that can be expected from intergroup social cohesion interventions. The effects identified in this review overall are small. Threats to human security, such as poor access to services, threats from natural disasters, high levels of crime and violence, or limited economic opportunities, are often seen as drivers of fragility and intergroup tensions. These tensions often erupt across social cleavages, particularly where one group is comparatively better off. Even where the interventions in this review incorporated elements of economic support, almost all of them recognised that this was insufficient to address the economic needs. For example, Alaref et al. (2019) interpreted finding null effects on labour market outcomes of the collaborative contact intervention in Lebanon as being due to the programme's primary focus on instilling social cohesion values. This aligns with the findings of the limited quantitative analysis of human security outcomes, which overall found no effects on outcomes of violence, employment or educational attainment. The exception is the case of Finkel et al. (2018), who found positive effects on employment outcomes – yet this study evaluated the largest-scale programme included in the review, a five-year, US$59 million intervention implemented across three countries.

Rather than relying on the mechanisms targeted by the interventions in this review – of seeing, talking and collaborating with the other – to 'fix' the relationships between groups in fragile contexts, there may be a need to integrate such mechanisms into programmes that address the systemic barriers to human security in the given context. For example, Ferguson (2019) noted that the theory of change relied on the economic support component of the intervention to change *perceptions* of inadequate service delivery to accommodate both host and refugee households, without trying to address (or measure) if access to services changed. To support sustainable social cohesion, long-term efforts that do address such systemic barriers regarding inadequate service delivery may be necessary.

Smaller-scale interventions may not provide sufficient intensity of treatment to have effects beyond direct participants. Many of the studies in this review that found nil or negative effects were those evaluating the effects of smaller-scale intergroup social cohesion interventions on indirect participants. Diffusion of effects to those who do not

participate directly in any of the intervention activities tend to be smaller than effects on people who receive a higher 'dosage' of the intervention. This phenomenon has been called the 'funnel of attrition' (White 2013). For such interventions to reach a wider population, a longer-term and more intensive intervention may be required. For example, Cleven (2020) measured effects of a school-based peace education programme on parents whose children attended the targeted school, most of whom did not participate in the intervention activities. The study found generally nil or even negative effects, a starkly different picture than that found by the school-based peace education studies measuring effects on the children who participated. Aladysheva and colleagues (2017) and Alan and colleagues (2020) found positive effects of school-based peace education interventions on treated children's trust and willingness to help, while Alan and colleagues (2020) further found positive effects on willingness to participate and acceptance of diversity. However, studies of larger-scale programmes were able to identify effects on indirect participants. Finkel et al. (2018) measured effects of a workshops-contact-econ intervention on a mixture of direct and indirect participants, and found positive effects on trust and sense of belonging. Yet, as noted above, the intervention was very large scale, targeting over a hundred different 'core' and 'non-core' programme areas. This suggests that smaller-scale interventions, such as the one studied in Cleven (2020), may not provide sufficient intensity of treatment to have effects beyond direct participants.

Impacts on intergroup relationships among participants may emerge sooner than impacts on wider intergroup relationships, which may need more time to shift.
Social cohesion is a complex outcome, both to shift and to measure. As noted above, at times indicators related to intergroup social cohesion may be more akin to intermediate outcomes. Particularly, outcomes that measure changes in relationships among participants from different social groups, or that directly relate to the intergroup contact triggered by the intervention, may emerge sooner than effects on participants' perceptions of the out-group more generally.

What constitutes appropriate early-stage outcomes will likely vary by context and intervention. For example, Mousa (2019) evaluates the effects of a two-month collaborative contact intervention in Iraq, wherein young Muslim men joined a predominantly Christian soccer league. The study found positive effects on social cohesion outcomes where those outcomes related closely to the soccer league, such as willingness to continue training with the mixed team, vote for a Muslim player for a sportsmanship award or attend a celebratory iftar dinner for league participants. Although these are measures of intergroup social cohesion, they are very directly and immediately related to participation in the intervention and associated activities. The study did not find effects on social cohesion outcomes that were less directly related to the interactions of players on the field, such as on general measures of trust, willingness to donate to a mixed NGO, tolerance or a sense of exclusive victimhood.

Effects on broader measures of intergroup social cohesion are likely to take time to materialise and may also result from non-linear causal chains. For example, in one case, two endline surveys were undertaken, one at roughly a year following the start of a three-month-long intergroup dialogues intervention in Liberia and the second at roughly three years (Hartman et al. 2018). The longer-term follow-up identified effects on behavioural outcomes that had not yet emerged during the one-year follow-up. In another case,

staggered implementation of a one-year programme enabled exploratory analysis that suggested that larger effects were identified in communities where there was a longer follow-up period (Ferguson 2019).

Unfortunately, practical constraints may prevent longer-term follow-up periods. Particularly in complex, fragile contexts, realising positive impacts on social cohesion outcomes may require a longer duration or intensity of programme than is feasible. Evaluations based on follow-up periods of less than a year, particularly for shorter-term interventions (up to a year in duration), may not be able to show effects on some types of social cohesion outcomes, but may instead be able to identify effects on outcomes that occur earlier in the causal chains. For interventions that incorporate intergroup contact, this may include effects on relationships among participants related to one or more social cohesion dimensions or related to the specific type of contact introduced by the intervention.

4.6.3 Factors associated with specific social cohesion dimensions
Interventions that incorporate peace education may trigger mechanisms of 'seeing the other' that facilitate effects on trust. We were able to run meta-analyses of trust outcomes for three intervention groups: intergroup dialogues, workshops-contact-econ and media for peace. The latter two identified small positive effects on trust, yet the meta-analysis of intergroup dialogue interventions on trust outcomes found no effect. Among the other two intervention groups, both school-based peace education interventions that measured effects on children who participated found positive effects on trust, while the only collaborative contact intervention that measured trust outcomes found no effect. A common thread among interventions that had positive effects on trust outcomes is the incorporation of elements of peace education. This may be done in workshops or classrooms, but also through the 'edutainment' programmes implemented by the radio dramas that fed into the media for peace meta-analysis. By contrast, the intergroup dialogue interventions that fed into the meta-analysis did not incorporate either of the three types of peace education identified (conflict resolution skills-building, social and emotional skills-building, or edutainment). Further, the only collaborative contact intervention that reported measures of trust found no effect of the intervention. It also did not incorporate elements of peace education. Although two other intergroup dialogue interventions and one other collaborative contact intervention did incorporate elements of peace education, they did not report any measure of trust. As such, they do not feed into the analysis, and we cannot say whether they may have been effective or not at improving trust.

In trying to understand why incorporating peace education may facilitate effects on trust, we return to the typology of mechanisms identified earlier in the review. As discussed in Section 1.5, peace education aims to build the social and emotional skills required to trigger mechanisms of 'seeing the other'. This process aims to build familiarity and comfort with others' perspectives, and may focus on recognising similarities and respecting differences between and across people as individuals or social groups. This may create the necessary preconditions of comfort and respect for intergroup trust to emerge. However, it is worth reiterating that this analysis is exploratory, and as noted above, draws on incomplete data. Understanding how trust can be built is the subject of substantial ongoing and theoretical literature, which may or may not confirm this finding. Nonetheless, the framework synthesis of studies included in this review suggests that

interventions incorporating components of peace education may be associated with greater effects on outcomes of trust.

Impacts on sense of belonging may be non-linear; negative effects may be identified before positive effects have time to emerge. The impact evaluation of a school-based peace education intervention in Kyrgyzstan found a negative effect of the intervention on students' sense of belonging and confidence (Aladysheva et al. 2017). The other two school-based peace education interventions that reported effects on sense of belonging found no impact. The study in Kyrgyzstan further identified negative effects on students' sense of confidence and self-efficacy, particularly for girls. The authors interpreted these results as suggesting that the intervention was of insufficient duration. While there was time to teach students to recognise bias, inequality and prejudice in their communities, the treatment was too short to sufficiently build students' social and emotional skills that they would need in order to improve the situation. This in turn may have left students feeling less settled in their communities, and thus driven the negative impact on the outcome measuring whether they feel at home in Kyrgyzstan.

The findings of the meta-analysis of workshops-contact-econ interventions lend some support to this theory. Overall, the meta-analysis found null effects of these interventions on outcomes related to sense of belonging. However, of the three studies that fed into the meta-analysis, one was of only one year in duration (Ferguson 2019), while the other two were of four to five years (Causal Design 2016; Finkel et al. 2018). When the shorter-term programme was dropped from the meta-analysis, the effect became positive and significant. As this left only two studies in the meta-analysis, the finding is only exploratory, but it suggests that positive effects on sense of belonging are possible given sufficient time and sustained interaction with the target communities.

This finding is specifically related to interventions incorporating components of peace education, which requires participants to confront the negative consequences of intergroup prejudice and bias. Among other interventions in the review, two collaborative contact interventions reported effects on sense of belonging, and both found a positive impact that did not cross the line of no effect (Alaref et al. 2019; Mousa 2019). Unlike interventions incorporating peace education, however, collaborative contact-based approaches to building intergroup social cohesion do not require participants to directly confront their differences or the tensions across groups. This focus on a positive interaction between groups may facilitate effects on sense of belonging. However, these impact estimates were based on short-term follow-up periods of two- and three-month programmes; whether the impacts are sustainable is unknown.

Levels of intergroup interaction at baseline may influence the type and scale of possible impacts on willingness to participate. We find a significant positive impact of collaborative contact interventions on outcomes of willingness to participate, yet mixed findings for all other intervention types that report outcomes on this key dimension of social cohesion. The effect of the workshops-contact-econ interventions was in a positive direction but it crosses the line of no effect. One possible explanation for why this may be relates to the nature of participation measured across the different studies. In the collaborative contact interventions, authors of the studies presented evidence suggesting participants had minimal intergroup contact pre-intervention. By contrast, in the case of the workshops-contact-econ interventions, baseline conditions ranged from some to

significant regular contact between the groups outside of the intervention, tensions notwithstanding. In such cases where groups already interact regularly, a general willingness to participate may already exist, and thus evaluations may be unlikely to identify changes in openness to or frequency of interaction. However, they may be able to realise effects on the nature of the interaction. Within the measures feeding into the meta-analysis, effects on three outcomes from three studies relating to frequency of participation were null, whereas the effect on an outcome of the nature of participation – whether respondents felt shared resources were managed peacefully – was positive and significant (Causal Design 2016). This potential explanation is further supported by the positive and significant effects identified of a workshops-contact-econ intervention on a social cohesion index, which incorporated measures not only on the frequency of interactions, but also on behaviours during interactions, acceptance of diversity, and perceived tensions over interactions (Ferguson 2019).

Another potential explanation may be due to participants in workshops-contact-econ interventions knowingly self-selecting into social cohesion interventions. In the case of the collaborative contact interventions, participants in the programmes that feed into the meta-analysis did not self-select into explicitly intergroup social cohesion interventions. In one case, participants were blind to the true purpose of the intervention, which was masked as a vocational training programme (Scacco and Warren 2018), while in another, participation was required by national law (Okunogbe 2018). In the third, the draw of the intervention was the opportunity to play in a well-funded soccer league (Mousa 2019).

However, students who participated in the school-based peace education intervention in Turkey also did not self-select into the programme (their teachers did) (Alan et al. 2020). The study found no effect on willingness to participate with out-group students from another school, although it did identify an effect on willingness to participate with classmates. This is more akin to the aforenoted factor related to intermediate impacts on improved cohesion among participants emerging before effects on broader social cohesion can be identified. Indeed, this factor may also contribute to understanding the positive impacts of the collaborative contact interventions. The three studies feeding into the meta-analysis identified positive impacts on willingness to participate in activities similar to those carried out through the intervention. For example, Mousa (2019) identified positive effects of the intergroup soccer league on willingness to sign up for a future mixed-religion league; Okunogbe (2018) identified a positive impact of the volunteer exchange programme on willingness to migrate outside the respondents' ethnic region.

Effects on acceptance of diversity may be highly context-dependent. Throughout the interventions included in this review, findings on acceptance for diversity were mixed, and it was difficult to identify clear patterns that may explain the heterogeneity in effects. As we have noted, one of the challenges in synthesising effects on intergroup social cohesion interventions is that they frequently measure outcomes in very different ways. However, in part this is due to the need to identify outcome constructs that are relevant to the contexts in which they are being measured.

For example, we have noted that the pattern of findings among radio dramas shows some positive effects in Burundi and Rwanda, but mixed and even negative effects in the

DRC. The exception to this trend is when looking at measures of inclusive or exclusive victimhood. These measures assess the extent to which respondents feel that only (or primarily) their group has suffered in the conflict, as opposed to feeling that all groups have suffered. In Burundi and in the DRC, the studies identified positive effects on these outcomes – indeed, this was the only social cohesion outcome on which the study in the DRC identified a positive effect (Bilali et al. 2016; Bilali and Vollhardt 2013).

However, in Rwanda, the programme was unable to realise a significant effect. This may be due to the context of the history of the genocide in Rwanda; although people of all three main ethnic groups were murdered, by and large the genocide was perpetrated by one group against another group. Drawing inferences about acceptance for diversity via an outcome where a positive effect is understood to be a larger proportion of people feeling that all groups suffered equally in the conflict may not be appropriate in this context.

Nigeria, the country for which we identified the largest number of studies, offers an example of the extent to which appropriate measures of intergroup social cohesion must be highly contextualised. One study measured the effects of the national government's policy of requiring university graduates to spend a year undertaking public service in a randomly assigned place, focusing on an ethnicity from south-western Nigeria (Okunogbe 2018). A second study measured the effects of a highly targeted collaborative contact intervention masked as a vocational training programme in the north-western city of Kaduna (Scacco and Warren 2018). A third study measured the effects of a workshops-contact-econ intervention in Nigeria's middle belt (Dawop et al. 2019). These three studies, although all taking place in Nigeria, each identified a distinct social cleavage across which tensions flared. In the first, ethnicity was the primary salient identity; in the second, religion; and in the third, farmer versus pastoralist lifestyles. Each study presented evidence that supported the decision for which identity cleavage was most appropriate in terms of building intergroup social cohesion and mitigating tensions.

Different groups, and indeed different individuals, will have different starting points in terms of what constitutes acceptance of diversity. The study in the south-west measured acceptance of diversity by asking respondents whether any of their four closest friends were from a different ethnic group (Okunogbe 2018). This is quite different from the question posed in the north-western city, which was whether respondents had any friends they considered close from a different religion (Scacco and Warren 2018). The study in the Middle Belt did not ask about friendship specifically, but did include social interaction in its measures (Dawop et al. 2019). Unlike the choice of social cleavage of focus, the implications of these different measures of acceptance of diversity were not locally justified, and it is thus difficult to draw conclusions about the heterogeneity in findings.

4.7 Cost evidence

Overall, the studies provided very little information about the cost of the interventions. When costs were mentioned, the methodology for arriving at these costs or the breakdown of these costs by expense or outcome were usually not given. A full overview of cost data per study is included in Online appendix B.7 Cost evidence.

Nine of the 24 included studies provided some information related to programme costs. Six of the studies provided total programme costs (Alaref et al. 2019; Finkel et al. 2018; Hartman et al. 2018; IMPAQ International 2017; Dawop et al. 2019; Okunogbe 2018), ranging from less than US$2 million to a total cost of US$400 million per year for a national-level programme.

Two studies reported approximate costs per direct participant (Hartman et al. 2018; Dawop et al. 2019), with the costs ranging from US$60 (Dawop et al. 2019) to US$100–150 per direct participant (Hartman et al. 2018). There is no information available about direct participant attrition for either of these studies, and thus it is unclear if or how this factors into the cost. One study reported example or component costs, but not overall cost evidence (Cilliers et al. 2018), while two other studies reported the amount of funding they were allocated from certain sources for a certain year (Ferguson 2019; Svensson and Brouneus 2013). Finally, one study presented a breakdown of costs based on strategic outcome (Finkel et al. 2018).

The diversity of cost measures included and the lack of details about the methods used to arrive at the different estimates means that the usefulness of the existing data on costs is limited. Without this detail, we are unable to convert the cost data into comparable metrics for the purposes of cost-effectiveness analyses. This is a major gap in the existing literature that should be addressed in future studies.

5. Discussion and conclusion

This systematic review summarises the findings from 37 impact evaluations of 24 different studies, comprising 31 unique interventions or intervention arms, which aimed to build intergroup social cohesion in fragile contexts. Most of the studies were conducted in Sub-Saharan Africa, but we also identified studies carried out in the Middle East and North Africa, Europe and Central Asia, and East Asia and Pacific, but no studies were identified in Latin America and the Caribbean.

Based on an analysis of the different intervention components, we identified six distinct types of interventions: (1) school-based peace education interventions; (2) collaborative contact interventions; (3) intergroup dialogue interventions; (4) workshop-based peace education with intergroup contact and economic support; and (5) media for peace interventions.

While there are some exceptions, the overall pattern across intervention types and different measures of social cohesion suggests small positive effects. This trend is repeated when the analysis is disaggregated by intervention types and dimensions of social cohesion, although due to the low number of studies included for each analysis the majority of these estimates remain imprecise and we also observe high heterogeneity across many of the analyses.

We conducted a qualitative synthesis of barriers and facilitators that may help to explain these results. The factors identified through that synthesis are summarised here:

- **Programmes that accurately identified local bottlenecks to intergroup social cohesion tended to have larger and more positive effects.** We identified multiple instances in which the bottlenecks to social cohesion targeted by the interventions appear to have been misaligned either with the context or with the population.

- **A lack of conflict assessments may be a barrier to better targeting of programme participants and key intervention strategies.** While conflict assessments are regularly used to inform the design and targeting of programming in fragile contexts, this was only explicitly mentioned in a couple of contexts. A lack of such analysis, and the associated targeting that follows, represents a possible barrier to larger effects from the programmes included in this review.

- **A lack of substantive changes in intermediate social cohesion outcomes may be a barrier to larger improvements in final social cohesion outcomes.** The effects of intergroup social cohesion interventions on intermediate outcomes, such as sociocultural awareness, social and emotional skills, and dispute resolution practices, are limited, and small at best, with some examples of negative effects. A breakdown in the theorised causal chain, from intervention activities to improved social cohesion outcomes, may explain small effects on these final outcomes. However, as few studies report intermediate outcomes, it is not clear if this 'breakdown' is due to limitations in programme design (theory) or implementation (practice).

- **Intergroup social cohesion interventions may not be sufficient for sustainable social cohesion without structural changes addressing threats to human security outcomes.** Ultimately, when it comes to building sustainable social cohesion and peace, without broader structural changes that address structural drivers of conflict there may be limits to the effects that can be expected from intergroup social cohesion interventions.

- **Smaller-scale interventions may not provide sufficient intensity of treatment to have effects beyond direct participants.** Many of the studies in this review that found nil or negative effects were those evaluating the effects of smaller-scale intergroup social cohesion interventions on indirect participants. The diffusion of effects to those who do not participate directly in any of the intervention activities tends to be smaller than the effects on people who receive a higher 'dosage' of the intervention.

- **Impacts on intergroup relationships among participants may emerge sooner than impacts on wider intergroup relationships, which may need more time to shift.** Smaller-scale interventions that incorporate elements of intergroup contact in particular may be better able to identify effects on changes in relationships among participants from different groups who have interacted with each other through the intervention. Effects on intergroup relationships more broadly may require more intensive interventions and/or time to emerge.

- **Long and non-linear causal chains may be a barrier to substantive improvements in social cohesion.** Social cohesion is a complex outcome, both to shift and to measure. Effects are likely to take time to materialise, be difficult to measure and may also result from non-linear causal chains. Local conditions at baseline may influence the type and scale of possible impacts. Particularly, effects on acceptance of diversity may be highly context-dependent.

5.1 Implications for policy and practice

While there are some exceptions, the overall pattern across intervention types and different measures of social cohesion suggests small positive effects of the interventions assessed in this review. The results, and in particular the effects of specific interventions

on specific outcomes, should be interpreted with some caution, as they are often based on relatively few observations and many rely on studies with high risk of bias. Based on the findings from both our qualitative and quantitative analysis, we identify the following implications for policy and practice:

- While intergroup social cohesion outcomes are relatively difficult to change, the evidence suggests that it is possible to improve such outcomes through targeted interventions.
- However, intergroup social cohesion interventions are not sufficient on their own to address the drivers of intergroup tensions. For larger improvements in social cohesion outcomes, multi-component interventions that address intergroup social cohesion as well as structural barriers to human security may be needed.
- Realistic timeframes are needed to allow substantive changes to social cohesion to materialise, and for the potential for non-linear effects, taking into account the duration and intensity of the intervention. Evaluations with shorter-term follow-up periods may need to focus on identifying impacts on early stages of the causal chain.
- If the intervention aims to improve social cohesion outcomes at a population level, and beyond direct participants, more intensive intervention strategies are likely to be required.
- Conducting a rigorous conflict diagnostic assessment during intervention design may facilitate effectiveness, by identifying the key bottlenecks to social cohesion and sustainable peace in the intervention context, including identifying key stakeholders from all groups who may need to be engaged in the intervention to ensure effectiveness. This can ensure that intervention strategies are chosen to respond to locally relevant needs and context.
- Developing a theory of change for the proposed intervention, mapping out the theorised causal pathways, together with locally relevant indicators for intermediate and final outcomes, can improve intervention effectiveness and facilitate better learning and evaluation. The risk of negative effects also highlights the need to identify and mitigate any potential negative unintended effects.
- When existing evidence is available, consulting this report and relevant primary studies can help to improve outcomes by informing the design of new programmes.
- Finally, when developing and implementing interventions where there is a limited evidence base, consider whether it would be possible to include an impact evaluation within the new programme.

5.2 Implications for research

This review has identified a modest and growing body of literature assessing the effects of programmes aiming to improve intergroup social cohesion in fragile contexts. The existing evidence allows us to identify some preliminary findings, but our ability to identify strong and generalisable findings is somewhat limited by the characteristics of the available evidence. First, while the growth in impact evaluations is encouraging, the field is highly fragmented. This fragmentation spans the types of interventions being evaluated, how authors describe them, the outcome constructs being measured and how they are measured. As a result, the number of observations for any given combination of

intervention and outcome constructs is generally small, limiting our ability to compare and synthesise findings across studies to identify generalisable and context-specific findings.

Second, the geographical distribution of studies is uneven, with the majority of studies being conducted in Sub-Saharan Africa, with a handful of studies from the Middle East and North Africa, Europe and Central Asia, and East Asia and Pacific, and no studies from Latin America and the Caribbean identified. Third, few of the existing studies adopt mixed-method designs with causal chain analysis, limiting in particular the extent to which the existing evidence addresses questions related to programme design, implementation and context. Fourth, while the included studies meet a threshold for inclusion based on study design and analysis criteria, the vast majority of studies in this area suffer from high risks of bias, due to limitations in their study design and analysis.

Fifth, there are some important gaps in the way studies are conducted and the questions they address. Notably, over half of the included studies do not include any mention of ethics. For an area where there is also evidence to suggest that interventions can do harm, appropriate procedures for addressing ethics, including through formal review and ethics approval, is essential. Finally, relatively few studies address intermediate outcomes, costs and equity. Analyses of costs are especially important to ensure that study results are relevant for a policy and practice audience. We had intended to conduct cost-effectiveness analysis for this review, but we had to abandon these plans due to a lack of data and transparent reporting for the studies that do provide data on this.

Funders and researchers should consider the following when commissioning and designing new studies:
- Developing a standardised intervention taxonomy to facilitate the use of common terminology to describe the same interventions;
- Adopting a common framework across studies, including both intermediate and impact outcomes to which locally relevant indicators can be mapped, to enhance the value and potential for cross-study lessons and evidence synthesis – this should include measures of all five dimensions of social cohesion: trust, sense of belonging, willingness to participate, willingness to help, and acceptance of diversity;
- Recognising that effects on acceptance of diversity may be highly context-dependent, and ensuring that efforts are made to identify appropriate measures of change for the local context;
- Where participants knowingly self-select into intergroup social cohesion interventions, evaluating effects on the nature of intergroup interactions may be more appropriate than measuring frequency or openness to interactions;
- Ensuring analysis is structured along the causal chain, including identifying and evaluating effects on outcomes earlier in the chain before social cohesion outcomes, particularly for evaluations based on shorter follow-up periods;
- Ensuring new studies include data on costs, based on a clear and transparently reported approach;
- Employing study designs informed by a mixed-methods, theory-based approach to impact evaluations that considers a range of questions relevant to policy and practice, including intervention design, implementation, contextual factors and intermediate outcomes;

- Ensuring research designs and methods are sensitive to inequalities across different population groups – taking into account diverse experiences, power dynamics and gendered inequality in study design and conduct will ensure new studies are sensitive to the needs and effects of programmes with regard to vulnerable groups;
- Adopting best practice for ethical research conduct and protection of research participants, including undertaking and reporting review and approval of study protocols and procedures by relevant review boards; and
- Adhering to commonly accepted standards for research transparency and reporting, including pre-registration of all new studies (experimental and quasi-experimental, e.g. in the 3ie Research Transparency Policy 2018).

6. Acknowledgements

7. Contributions of authors

The review team comprises Ada Sonnenfeld (AS), Dr Jennifer Doherty (JD), Miriam Berretta (MB), Birte Snilstveit (BS), Dr Shannon Shisler (SS) and John Eyers (JE). AS and JD bring relevant field experience as practitioners in social cohesion in fragile contexts, while JD further adds expertise from her doctoral research on governance and stability in ethnically divided contexts. AS co-led the development of a previous systematic review looking at citizen engagement in governance, while JD and MB have significant experience supporting the development, data analysis and management of systematic review projects. BS is an expert in systematic review methods, with over a decade of synthesis experience. SS is a quantitative methods expert, with over a decade of experience designing, managing and analysing quantitative research, including meta-analyses. JE is an expert in information retrieval, with decades of experience in supporting information retrieval for systematic reviews.

Review responsibilities:
- Content: The content of the review has been developed by AS, JD and MB, with input and quality assurance from BS and SS.
- Systematic review methods: The review methods were drafted by AS, JD and MB, with input and quality assurance from SS and BS.

- Statistical analysis: The statistical analysis was overseen by SS. Effects data were extracted by SS and JD, with quality assurance from SS and BS, while outcome classification and grouping were undertaken by AS and JD.
- Information retrieval: The initial search strings were developed by MB, AS and JD, then quality-assured and carried out by JE. Grey literature review, snowballing and reference checks were developed by AS, JD and MB.

8. Declarations of interest

No conflicts of interest are present for any of the authors.

9. Characteristics of studies

9.1 Characteristics of included studies

Table 8: Characteristics of included studies

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
School-based peace education interventions								
Aladysheva et al. 2017 **Kyrgyzstan – Osh, Jalalabad and Batken oblasts (southern Kyrgyzstan)**	• Conflict resolution • Social and emotional skills • Cooperative contact • Dialogue sessions	Living Side By Side (LSBS): LSBS is an interactive peacebuilding training programme for youth aged 13–18. LSBS aims to provide motivation and the means to change attitudes and behaviour related to intolerance and conflict through: increasing awareness and knowledge, discovery of shared values and 'common humanity', and learning by applying new skills. The training was implemented as an extracurricular after-school workshop. At the end of each round of the training, the students were guided to develop and implement	• Ethics approval referenced • Subgroup analysis disaggregated by sex and ethnicity	17	Immediately after and at 12 months	RCT	Business as usual	• Trust (trusting: intragroup, intergroup, generalised) • Sense of belonging (shared: generalised) • Willingness to participate (willing: generalised) • Willingness to help (give: generalised) • Acceptance of diversity (intolerance: intergroup) • Acceptance

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
		a school or community project, working in multi-ethnic groups and serving multi-ethnic audiences.						of diversity (bias: intergroup)
Alan et al. 2020 **Turkey –** Sanliurfa and Mersin (southern provinces)	• Social and emotional skills	Understanding Each Other: The intervention is a curriculum designed to promote social cohesion, by improving perspective-taking ability among elementary school children. About three hours per week were dedicated to the programme. The programme was implemented through a variety of reading and visual materials, including animated videos and diaries of refugee and host children. They also played games: two versions of a trust and cooperation (prisoner's dilemma) game and one version of a dictator game with a follow-up survey.	• Subgroup analysis by conflict-affected, displaced population	9	Immediately after	Cluster RCT	Business as usual	• Trust (trusting: intergroup) • Trust (trustworthy: intergroup, intragroup) • Sense of belonging (shared: intergroup) • Willingness to participate (willing: intergroup, intragroup) • Willingness to help (give: intergroup) • Willingness to help (receive: intergroup, intragroup)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
Biton and Salomon 2006 **Palestine** – Major urban areas	• Social and emotional skills	Pathways into Reconciliation: The programme is a school-based curriculum working towards a more positive conception of peace among the beneficiaries. The curriculum concerns three spheres: awareness; understanding of and tolerance for others in one's immediate surroundings; and cultivating tolerance for Palestinians (or Jews). The programme is embedded in regular school curricula and carried out by regular classroom teachers who have been trained in the programme. In addition, activities included face-to-face meetings between Jewish and Palestinian students. The intervention aims to: impart values of tolerance and acceptance, and recognition of equal	• Does not address gender or equity	6–8	Immediately after	Cluster RCT	Business as usual	• Acceptance of diversity (tolerance: intergroup)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
		rights; develop awareness; and teach non-violent communication and conflict resolution skills.						
Cleven 2020 (NDC Mostar) **Bosnia and Herzegovina –** Canton of Herzegovina-Neretva Treatment: Prozor-Rama Control: Capljina	• Social and emotional skills • Collaborative contact • Discursive contact	The Nansen Dialogue Centre Mostar programme is focused on an ethnically segregated elementary school, where Croats and Bosnians run two schools under one roof, with separate teachers and principal, and using the building in shifts. The dialogue activities included seminars for local politicians, administrators, parents and teachers, and joint activities for school children. The seminars aimed to build and improve inter-ethnic relationships, by giving participants the opportunity to understand each other's perspectives and experiences.	• Equity not addressed	36	After the programme had been running for 2 years	Difference in difference (DID)	Business as usual (non-participants)	• Trust (trusting: intragroup and intergroup) • Sense of belonging (shared: intergroup) • Willingness to participate (active: generalised) • Acceptance of diversity (tolerance: intergroup)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
Collaborative contact								
Alaref et al. 2019 **Lebanon –** North, Mt Lebanon, Beirut, Bekaa, South	• Conflict resolution • Cooperative contact • Communal (intergroup) contact • Economic security support	National Volunteer Service Program (NVSP): The NVSP seeks to promote social cohesion among Lebanese youth aged 15–24 by financing volunteer activities and providing soft skills. The primary objective of the intervention is to defuse social tensions by bringing diverse youth together to work towards shared goals and find solutions to improve community assets and service delivery, as well as increase their employability.	• Subgroup analysis by education, place of residence, socio-economic status and sex	4.5	12 months	DID	Business as usual	• Sense of belonging (shared: intergroup) • Acceptance of diversity (tolerance: intergroup)
Okunogbe 2018 **Nigeria –** National	• Cooperative contact Communal (intergroup) projects • Economic security support	National Youth Service Corps (NYSC) is a compulsory one-year national service programme for Nigerian graduates aged under 30. The programme's goal is to reconstruct, reconcile and rebuild post-war Nigeria. It was designed to help ease labour	• Subgroup analysis based on age	12	Unclear	RCT	In-group intervention	• Acceptance of diversity (tolerance: intergroup) • Social cohesion index (mixed, intragroup, intergroup)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
		shortages and even out the distribution of skilled labour. Youth from different regions of the country serve together, which is intended to promote unity and reduce prejudice.						
Mousa 2019 **Iraq –** Ankawa, a Christian suburb of Erbil	• Cooperative contact	This intervention aims to reduce the role of identity in preventing intergroup interactions. By assigning Iraqi Christians displaced by ISIS to a mixed soccer team with Sunni Arabs (the same ethno-religious background as ISIS), the intervention seeks to increase tolerance and trust towards Muslims, while reducing bias.	• Evaluation design: heterogeneity analysis by religion	4	Immediately after/one day after/ three weeks after/4 months after	RCT	Business as usual	• Trust (trusting: intergroup, generalised) • Trust (mistrust: intergroup) • Sense of belonging (shared: intergroup) • Willingness to participate (willing: intergroup) • Acceptance of diversity (tolerance: intergroup)
Scacco and Warren 2018 **Nigeria** – Kaduna	• Proximity contact • Cooperative contact	Urban Youth Vocational Training (UYVT): The intervention is a computer course designed to reveal whether sustained contact in	• Ethics approval referenced • Heterogeneity analysis by	4	Not clear	RCT	Business as usual (participated in survey only) and	• Acceptance of diversity (bias: intergroup) • Willingness to

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
(north-western Nigeria)		an educational setting can improve communal relations in a conflict-prone environment. Approximately one third of the UYVT participants were assigned to religiously homogenous classrooms, and the others to heterogenous ones. Partners were randomly assigned to work in cooperation on course assignments and custom-designed activities. The main goal of the intervention was to understand whether structured social contact decreases prejudice and discrimination and improves communal relations in a conflict-prone environment.	religion, socio-economic status, language and education • Data disaggregated by religion				then between homogenous and non-homogenous classes, i.e. those with inter-religious contact and without	help (give: intergroup)
Intergroup dialogues								
Cilliers et al. 2018 (Cilliers et al. 2016) **Sierra Leone:**	• Dialogue sessions • Trust and reconciliation forums • Communal	Fambul Tok (FT): FT aims to deal with the effects of civil war and foster sustainable peace. The intervention involves the training of committee	• Ethics approval referenced • Conflict-affected population	3–4	12–18 months	Cluster RCT	Business as usual	• Trust (trusting: intergroup, generalised) • Sense of belonging

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
Bombali, Kailahun, Koinadugu, Kono and Moyamba districts	(intergroup) projects • Economic security support	participants in trauma healing and mediation. These committees then plan a reconciliation forum for broad-based participation. It comprises a ceremony of truth-telling and forgiveness-seeking, capped by a ceremony to 'cleanse' those who expressed remorse. FT also establishes local structures to further heal the community. These include: a Peace Tree as a focal point for dispute resolution, communal farms as a reconciliation pledge, and a Peace Mother to focus on women's issues.	targeted					(shared: intergroup) • Willingness to participate (willing: generalised) • Willingness to help (give: generalised)
Hartman et al. 2018 (Blattman et al. 2012; Blattman et al. 2011)	• Conflict resolution • Social and emotional skills • Peace messaging	Community Empowerment Program: The intervention focused on improving informal conflict resolution institutions by teaching skills and practices (i.e. how to resolve disputes) and	• Heterogeneity analysis by age, ethnicity, religion, socio-economic status, social capital, sex	21	1–18 months	Cluster RCT	Business as usual	• Acceptance of diversity (tolerance: intergroup)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
Liberia – Lofa, Nimba and Grand Gedeh counties	campaigns • Civic education	instilling norms (i.e. how disputes ought to be resolved). 15% of adults in target communities were invited to workshops led by two Justice and Peace Commission facilitators. The training emphasised: direct engagement in disputes, strategies for problem solving and negotiation, face-saving and positive-sum resolutions, and avoidance of forum shopping and the formal justice system. Lessons combined lectures, small group sessions, participatory dramas and experience-sharing opportunities. The programme also addressed civic education, human rights and community collective action.						
Lonergan 2016 (Lonergan	• Reconciliation dialogues	The programme was designed to reduce interpersonal atrocity risk,	• Ethics approval referenced	1 (5 weeks)	Immediate (assumed)	RCT	Business as usual	• Trust (mistrust: intergroup; trusting:

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
2017) **Sri Lanka –** University of Colombo		by creating ideal conditions for reconciliation and positive intergroup contact. This was achieved through five sessions that focused on: team-building activities, themes of identity and coexistence, trust, individual and intergroup empathy, and mutual acknowledgement of past wrongs.						intergroup) • Sense of belonging (shared: intergroup; separate: intergroup) • Willingness to participate (anxiety about: intergroup) • Acceptance of diversity (tolerance and intolerance: intergroup; acceptance and rejection of multiple perspectives: intergroup; bias: intergroup)
Rime et al. 2011 (Kanyangara et al. 2014; Kanyangara	• Truth and reconciliation forums	The government modified a traditional community-based conflict resolution system called gacaca for judging those still	• Subgroup analysis by exposure to violence (victim vs perpetrator)	Unclear	Immediately after intervention	DID	Waitlist control	• Effect sizes could not be calculated

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
et al. 2007) **Rwanda -** National		imprisoned for participation in the Rwandan genocide. The tribunals are made up of elected people of integrity. The prisoners, survivors and the wider community discuss the alleged acts. Participation and acceptance of the agreed punishment are mandatory. These tribunals are restorative in nature as they aim to reintegrate perpetrators so that they can coexist in the same community with victims.						
Schiller, 2012 **Indonesia –** Central Highlands region of Aceh	• Conflict resolution • Proximity contact • Dialogue sessions	Workshop for Peace – Together we can build, maintain and safeguard peace: The intervention aimed to test the effects of intergroup contact on reconciliation, and whether dialogue-based contact succeeds in: reducing intergroup prejudice, increasing intergroup empathy, trust, tolerance	• Conflict-affected population targeted	6	Immediately after	RCT	Waitlist control	• Acceptance of diversity (bias: intergroup)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
		and forgiveness, and increasing individual healing in a post-conflict context. The intervention applied three distinct treatments to determine the optimal conditions for facilitating reconciliation: dialogue-based contact workshops, training-based workshops, and mixed-method contact workshops that used a combination of dialogue and training techniques.						
Svensson and Brouneus 2013 **Ethiopia –** Addis Ababa University, Sidist Kilo University Campus	• Social and emotional skills • Dialogue sessions	Sustained Dialogue (SD): SD is a series of repeated small-group bi-weekly dialogue meetings on the issue of inter-ethnic tolerance on the university campus. They are held between groups with a history of conflict and tension led by trained moderators that aim to improve contact, end conflict and build peace.	• Equity sensitive analytical framework used	9	In the final weeks of intervention	RCT	Business as usual (non-participants)	• Trust (trusting: generalised, intergroup) • Trust (mistrust: generalised) • Trust (trustworthy: generalised) • Sense of belonging (separate:

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
		The intervention seeks to transform the relationships between conflicting parties by addressing the underlying issues that cause tensions.						intergroup) • Acceptance of diversity (bias: intergroup)
Workshop-based peace education with intergroup contact and economic support								
Causal Design 2016 **Nigeria –** Middle Belt states of Benue, Kaduna, Nasarawa and Plateau (CONCUR); Plateau state (IPNN)	Both projects: • Conflict resolution • Cooperative contact • Dialogue sessions • Communal (intergroup) projects CONCUR: • Community conflict prevention forums IPNN: • Peace messaging campaigns	Community-Based Conflict Management Cooperative Use of Resources (CONCUR) and Inter-Religious Peacebuilding in Northern Nigeria (IPNN): CONCUR's programme activities covered three categories: building capacity of local leaders for peaceful conflict resolution; increasing cooperation around economic activity and resource management; and generating support for long-term policy solutions among local and national leaders through research and advocacy. IPPN was designed to build on	• Conflict-affected population targeted	48	6 months	DID	Business as usual	• Trust (trusting: intergroup) • Willingness to participate (willing: intergroup)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
		CONCUR by building religious leaders' capacity for non-violent conflict resolution. Religious leaders disseminated messages of peace through radio adverts, while agreements on natural resource management were reached and a negotiator network was created.						
Ferguson 2019

Jordan – Irbid, Ajloun, Mafraq, Zarqa, Ma'an, Karak, Amman, Salt, Jerash, Tafilah and Balqa governorates | Software only: • Conflict resolution • Proximity contact • Communal (intergroup) projects

Hardware and software: • Conflict resolution • Proximity contact • Communal | Strengthening Social Capital and Reducing Tensions Between Jordanian Host Communities and Syrian Refugees (previously Conflict Management for Syrian Refugees, Host Communities and Municipal Actors in Jordan): The intervention aims to enhance stability and social cohesion by improving relationships between individuals in the displaced Syrian and host | • Ethics concerns discussed • All data were collected and research undertaken in line with ethical standards, and with the permission of International Security and Development Centre, | 76 | Immediately after 3 years of intervention | DID; IV-regression (2-stage least squares or bivariate probit) | Business as usual | • Willingness to participate (willing: intergroup) • Acceptance of diversity (tolerance: intergroup) |

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
	(intergroup) projects • Economic security support	Jordanian communities and improving relationships between the populace and local institutions via improved service delivery. This is achieved through a 'hardware' component that aims to provide infrastructure and reduce constraints to public and other services, and a 'software' component that aids community leaders to facilitate dispute resolution and increases cooperation between communities.	Leading Point, Mercy Corps, and the Jordanian government					
Finkel et al. 2018 **Burkina Faso, Chad and Niger**	• Conflict resolution • Social and emotional skills • Dialogue sessions • Peace messaging campaigns • Media	Peace through Development II (PDEV II): PDEV II aimed to increase resilience to violent extremism in at-risk communities in Burkina Faso, Chad and Niger. The programme had three subgoals: improved social cohesion, enhanced resilience to extremism, and	• Vulnerable population targeted • Age, culture, place of residence, socio-economic status and conflict-affected	42–49	Not given	DID	High and low treatment intensity area	• Trust (trusting: generalised)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
	capacity building • Economic security support • Civic education • Good governance support	improved civic outlook. It sought to achieve these through four strategic objectives: Youth More Empowered (via, among others, vocational and entrepreneurial skills training), Moderate Voices Increased (through integrated radio, social media, civic education and conflict resolution activities), Civil Society Capacity Increased (through training, strengthening advocacy and issue-based campaigns), and Local Government Strengthened (through organised community entities, citizen participation, advocacy and government outreach).	dimensions considered					
IMPAQ International 2017 (USAID) **Bosnia and**	• Social and emotional skills • Peace messaging campaigns	Trust, Understanding and Responsibility for the Future Activity (PRO-Future): PRO-Future aims to rebuild inter-ethnic trust by helping people address the past,	• Vulnerable population targeted • Heterogeneity analysis (other than	48	Immediately after 3 years of treatment from baseline	DID	Business as Usual	• Trust (trusting: intergroup)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
Herzegovina – 30 target municipalities across the country	• Media capacity building • Reconciliation dialogues • Economic security support: Communal (intergroup) projects	and understand and accept each other's narratives of the past. It seeks to engage key influencers to support these local reconciliation efforts. Programme activities included: trainings, peace camps, cultural events, roundtable discussions, online competitions and discussions, grants and community initiatives. By improving inter-ethnic relationships and trust, PRO-Future hopes to enable citizens to overcome division and advocate for institutional change, ultimately leading to social well-being and economic prosperity.	subgroup)					
Dawop et al. 2019 **Nigeria –** Middle Belt states of	• Conflict resolution • Cooperative contact • Dialogue sessions	Engaging Communities for Peace in Nigeria (ECPN): ECPN aims to prevent violence and conflict between farmer and pastoralist communities by	• Ethics approval referenced • Conflict-affected population targeted	48	3 years after baseline	Cluster RCT	Business as usual	• Trust (trusting: intergroup) • Willingness to participate (willing:

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
Benue and Nasarawa	• Early warning committees • Community conflict prevention forums • Communal (intergroup) projects	strengthening peaceful dispute resolution capacity, building trust through opportunities for collaboration, and fostering engagement among farmers and pastoralists, local authorities, and neighbouring communities. Committees for peace, early warning/response and joint projects were formed. Participants were trained in mediation and dispute resolution, and addressed tensions through the implementation of projects. The final stage was conflict prevention forums for the wider community to discuss important issues and policies. Government officials also attended these events.						intergroup) • Willingness to help (give: intergroup) • Willingness to help (give: generalised) • Acceptance of diversity (bias: intergroup)
Media for peace								
Bilali and Vollhardt 2013	• Radio drama	Musekeweya (New Dawn): The intervention is a radio soap opera designed to	• Equity-sensitive research methods	Not clear	6 months	RCT	Priming design	• Trust (mistrust: intergroup)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
Rwanda – all four regions including the capital		increase historical perspective-taking regarding intergroup conflict, reduce intergroup mistrust, and reduce construal of past victimisation (collective victimhood and inclusive victim consciousness). Educational messages based on psychological theories of intergroup conflict, reconciliation and trauma are embedded in a fictional story about intergroup conflict. The broadcast aims to prevent violence and promote reconciliation following the genocide.	• Age, education, sex, place of residence and victimisation were considered in sampling					• Acceptance of diversity (blindness to multiple perspectives: intergroup) • Acceptance of diversity (intolerance: generalised) • Acceptance of diversity (awareness of multiple perspectives: intergroup)
Bilali and Vollhardt 2015 **DRC –** North and South Kivu provinces (eastern DRC)	• Radio drama	Kumbuka Kesho (Remembering Tomorrow): Kumbuka Kesho is a violence prevention radio drama aimed at promoting peaceful intergroup relations in the eastern DRC. The entertainment education programme embeds educational	• Equity-sensitive research methods • Research team trained in ethical research practices for conflict-	Not clear	Not clear	RCT	Priming design	• Sense of belonging (separate: intergroup) • Willingness to participate (unwilling: intergroup) • Acceptance of diversity

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
		messaging aiming to raise awareness about factors that lead to mass violence, and provide role models to prevent violence and promote reconciliation. It seeks to reduce ingroup superiority, conformity aspects, perceived collective victimhood, social distance and political exclusion.	affected population					(blindness to multiple perspectives: generalised, intergroup) • Acceptance of diversity (awareness of multiple perspectives: intergroup) • Acceptance of diversity (intolerance: intergroup)
Bilali et al. 2016 **Burundi –** Bururi, Gitega, Makamba, Ngozi and Bujumbura Mairie provinces	• Radio drama	Murikira Ukuri (Shedding Light on the Truth): Murikira Ukuri is a radio drama designed with embedded educational messages that raise awareness about the roots and evolution of violence, and encourage behaviours that prevent violence and promote intergroup reconciliation. The intervention seeks to reduce social distance, in-	• Intervention design: vulnerable population targeted • Equity-sensitive research methods • Research team trained in ethical research	Not clear	Not clear	Ordered logit regression	Business as usual	• Trust (mistrust: intergroup) • Willingness to help (give: intergroup) • Acceptance of diversity (intolerance: intergroup, generalised) • Acceptance of diversity (blindness to

114

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
		group superiority, obedience and misaligned attributions of responsibility; and increase out-group trust, dissent, openness to talking about trauma, perspective-taking and participant understanding of the individuals' role in conflict, including the concept of active bystandership.	practices for conflict-affected population					multiple perspectives: intergroup) • Acceptance of diversity (awareness of multiple perspectives: intergroup) • Acceptance of diversity (bias: intergroup)
Paluck 2009 (Paluck 2007a; Paluck 2007b; Paluck and Green 2009) **Rwanda -** National	• Radio drama	Musekeweya (New Dawn): The intervention is an education entertainment radio soap opera designed to address the lack of trust, communication and interaction as well as trauma caused by the genocide. It features two Rwandan communities in a fictional conflict and conveys messages about reduction of intergroup prejudice, violence and trauma.	• Ethics approval referenced • Age, education, sex, socio-economic status were considered in sampling • Single-sex focus groups included • Subgroup analysis by ethnicity and survivor status	12	Immediately after	Cluster RCT	Encouragement design (listening groups for New Dawn vs unrelated programme)	• Trust (trusting: generalised) • Trust (mistrust: generalised) • Willingness to participate (willing: intergroup, intragroup) • Acceptance of diversity (tolerance: intergroup)

Papers and location	Intervention components	Intervention description	Equity	Intervention duration (months)	Follow-up period	Study design and analysis	Comparison group	Social cohesion outcomes reported
Vicente and Vilela 2020	• Peace messaging campaigns	The programme consists of a religious sensitisation campaign and a workshop on job searching, job opportunities and business management. The religious awareness session aimed to oppose extremist Islamic views, violence and insurgency. The professional training session aimed to facilitate employment in the local labour market, and thus increase the opportunity cost of joining a violent group or engaging in conflict.	• Ethics approval referenced • Intervention design: vulnerable population targeted • Evaluation design: subgroup analysis based on education, ethnicity and socio-economic status	3	0–2 months	RCT	Business as usual (no intervention)	• Willingness to help (reluctance: generalised; choosing not to help: generalised)

9.2 Characteristics of excluded studies

In total, 523 studies were excluded during the full-text review. Table 9 details the reasons for exclusion. The most common reason was that the intervention did not aim to build intergroup cohesion between groups across which tensions threatened to become, had been or were currently violent.

In total, 277 studies were excluded as they were classified as other social cohesion interventions. For instance, although Cyrus (2012) looked at social cohesion outcomes of relevance, such as tolerance, the intervention assessed representation within a military reform, therefore it was excluded as being based on vertical cohesion, the military being a structure of government. Additionally, 26 studies were excluded because the intervention did not take place in a fragile context, while a further 28 targeted a high-income country population. At full-text stage, we identified 19 that studied irrelevant interventions and 13 that did not evaluate interventions, though these were the most common reasons for exclusion during title and abstract screening.

In terms of exclusions for study design, 45 studies were excluded for not being primary studies, while 103 were excluded for not addressing questions of effects. A significant number of these papers followed the OECD Development Assistance Committee (DAC) evaluation guidelines; these studies were excluded as they were more focused on evaluating the relevance and sustainability of the interventions and descriptive data on outcomes, rather than attribution of impacts on participants. We excluded 68 studies on the grounds that they were theory-based evaluations whose primary method of analysis was qualitative. Inappropriate comparison conditions led to the exclusion of 119 studies, including primarily cases where there was no comparison group, and to a lesser extent, where the study only assessed the marginal effects of an intervention. For example, in Obiagu et al. (2020), the evaluation measured differences in impact between two similar social cohesion interventions; there was no 'pure' control group. This means that the evaluation measures the marginal effect of one approach compared to the other, rather than the effect of peace education compared to no peace education.

An additional 41 studies were excluded for not meeting minimum quantitative study design conditions. Malhotra and Liyange (2005) is an example of a study excluded because of this criteria, as there was no randomisation to participant allocation and the control group was not matched statistically to the treatment group. Six papers were excluded due to a lack of reporting on effects of any key social cohesion outcome of interest.

Finally, 12 studies were excluded as being only lab-in-the-field experiments, wherein the evaluation was designed primarily to evaluate immediate reactions to a short-term exposure rather than sustainably change intergroup social cohesion outcomes. These studies were excluded except where a lab-in-the-field experiment was used to measure the effects of a larger intervention. Nielsen et al. (2016) provides a good example of this as, despite the evaluation employing a longer-term follow-up, the outcome change is based purely on a 15-minute-long intergroup puzzle-solving exercise and no further intervention, and as such the paper was excluded under this code. Finally, we excluded 53 studies that were duplicate copies of included papers, and 13 studies because we were unable to access the full text of the study.

Table 9: Reasons for exclusion during full-text screening

Exclusion criteria	Number of papers excluded during full-text screening	Description
High-income country population	52,828	Based on the World Bank's Country and Lending Group classifications for the first year of intervention implementation.
Intervention not relevant	19	Study referred to an intervention that did not relate to any form of building social cohesion.
No intervention	13	An intervention is a set of activities implemented in a controlled, deliberative way, with the aim of creating a change in the participants. At title and abstract screening, it was not always clear whether there was an intervention or not; at the full-text screening stage, studies were mainly excluded on this code as they measured existing networks between groups.
Other social cohesion intervention	277	The largest number of studies excluded at full-text screening came under this code. Many studied were community-driven development projects which became excluded after title and abstract screening, many disarmament, demobilisation and reintegration programmes were also excluded here as they only included economic elements, criteria which excludes studies under this code.
Not a primary study	45	Many studies excluded here look at different relevant interventions and their evaluations of effectiveness, rather than making a judgement on effectiveness based on their own analysis.
Does not address impact	103	This code addressed whether the study evaluated the impact of the intervention on participants. In many cases here, the evaluations were based on the OECD DAC outlines, which did not address impact, but mainly addressed effectiveness and sustainability.
Not a fragile context	26	Country appears on the World Bank's List of Fragile Situations, or scored over 90 on the Fragile State Index, for the first year of the intervention's implementation. For pre-2005 studies, the state is fragile if the intervention took place in a situation where two or more armed groups had been in conflict leading to a minimum of 1,000 battle-related deaths.

Exclusion criteria	Number of papers excluded during full-text screening	Description
		The study defines a conflict between the groups in question as one of the motivating factors for their analysis.
Theory-based evaluation – Qualitative	68	When the primary source of analysis was qualitative, and there was a clear theory of change presented, studies were excluded on this basis.
Comparison conditions	119	Exclusion on comparison conditions came under two scenarios: (1) there was no control/comparison group; or (2) the study only assessed the marginal effects of an intervention. This could be when two similar interventions were assessed without a separate control group, or the comparison groups both participated in the same intervention but for differing lengths of time.
Fails minimum study design requirements	4,141	Most studies were excluded here as they were not RCTs or quasi-experimental designs.
Lab-in-field only	12	If an intervention was lab-in-field only, as opposed to lab-in-field to assess a larger intervention, it was excluded under this code.
Duplicate	53	
No access to full text	13	

9.3 Characteristics of ongoing studies

We identified five ongoing studies that may meet inclusion criteria, three of which focus on school-based peace education interventions. Table 10 presents the characteristics of ongoing studies.

Table 10: Characteristics of ongoing studies

Study short title	Start date	End date	Intervention group	Study design	Social cohesion outcomes
Brück and Nillesen (ongoing) – Kyrgyzstan	Feb 2014	Nov 2014	School-based peace education	RCT	Trust Sense of belonging
ECAS and Mercy Corps (ongoing) – Myanmar	Jan 2016	–	Workshops-based peace education with intergroup contact and economic support	–	Sense of belonging Acceptance of diversity
Spitler and Roessler (ongoing) – Liberia	Mar 2015	–	School-based peace education	Regression discontinuity design	Acceptance of diversity
CAF Pilar et al. (ongoing) – Colombia	Aug 2017	Nov 2019	School-based peace education	RCT	Sense of belonging Acceptance of diversity
Ferguson et al. (ongoing) – Colombia	Apr 2017	–	Intergroup dialogues	RCT	Trust Willingness to help Acceptance of diversity

Online appendices

Online appendix A: Further detail on inclusion criteria and study selection methods

https://www.3ieimpact.org/sites/default/files/2020-12/SR46-Online-appendix-A-Further-detail-on-inclusion-criteria-and-study-selection-methods.pdf

Online appendix B: Further analyses

https://www.3ieimpact.org/sites/default/files/2020-12/SR46-Online-appendix-B-Further-analyses.pdf

Online appendix C: Further detail of risk of bias and sensitivity analyses

https://www.3ieimpact.org/sites/default/files/2020-12/SR46-Online-appendix-C-Further-detail-of-risk-of-bias-and-sensitivity-analyses.pdf

Online appendix D: Data extraction tools

https://www.3ieimpact.org/sites/default/files/2020-12/SR46-Online-appendix-D-Data-extraction-tools.pdf

References to studies

Included studies[28]

Aladysheva, A, Kyzy, GA, Brück, T, Esenaliev, D, Karabaeva, J, Leung, W and Nillesen, E, 2017. *Impact evaluation of the Livingsidebyside peacebuilding educational programme in Kyrgyzstan*. 3ie Grantee Final Report. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: <https://developmentevidence.3ieimpact.org/search-result-details/impact-evaluation-repository/impact-evaluation-of-the-livingsidebyside-peacebuilding-educational-programme-in-kyrgyzstan/5926> [Accessed 18 December 2020].

Alan, S, Baysan, C, Gumren, M and Kubilay, E, 2020. *Building Inter-Ethnic Cohesion in Schools: An Intervention on Perspective-Taking.* HCEO Working Paper 2020-009. Chicago: Human Capital and Economic Opportunity Global Working Group. Available at: <https://hceconomics.uchicago.edu/research/working-paper/building-inter-ethnic-cohesion-schools-intervention-perspective-taking> [Accessed 18 December 2020].

Alaref, J, Solano, RL, Asenjo, C, Alawiye, M, El-Kik, M, 2019. *The Short-Term Impact of Inter-Community Volunteering Activities and Soft Skills Training on Self-Reported Social Cohesion Values: Quasi-Experimental Evidence from Lebanon*. World Bank Group Policy Research Working Paper, WPS no. 869. Washington, DC: World Bank Group. Available at: <http://documents.worldbank.org/curated/en/676801546874918189/The-Short-Term-Impact-of-Inter-Community-Volunteering-Activities-and-Soft-Skills-Training-on-Self-Reported-Social-Cohesion-Values-Quasi-Experimental-Evidence-from-Lebanon> [Accessed 18 December 2020].

Bilali, R and Vollhardt, J, 2013. Priming effects of a reconciliation radio drama on historical perspective-taking in the aftermath of mass violence in Rwanda. *Journal of Experimental Social Psychology*, 49(1), pp.144–151. Available at: <https://www.researchgate.net/publication/256752580_Priming_effects_of_a_reconciliation_radio_drama_on_historical_perspective-taking_in_the_aftermath_of_mass_violence_in_Rwanda> [Accessed 18 December 2020].

Bilali, R and Vollhardt, J, 2015. Do mass media interventions effectively promote peace in contexts of ongoing violence? Evidence from Eastern Democratic Republic of Congo. *Peace and Conflict: Journal of Peace Psychology*, August. Available at: <https://www.researchgate.net/publication/282324762_Do_Mass_Media_Interventions_Effectively_Promote_Peace_in_Contexts_of_Ongoing_Violence_Evidence_From_Eastern_Democratic_Republic_of_Congo> [Accessed 18 December 2020].

Bilali, R, Vollhardt, J and Rarick, J, 2016. Assessing the impact of a media-based intervention to prevent intergroup violence and promote positive intergroup relations in Burundi. *Journal of Community and Applied Social Psychology*, 26(3), pp.221–235. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/casp.2246> [Accessed 18 December 2020].

---

[28] The list is organised by the references for the 'top' copies used as the most recent or most complete impact evaluation, with additional impact evaluation papers associated with a top copy listed below on an indent.

Biton, Y, Salomon, G, 2006. Peace in the eyes of Israeli and Palestinian youths: Effects of collective narratives and peace education program. *Journal of Peace Research*, 43(2), pp. 167–180. Available at: <https://edu.haifa.ac.il/personal/gsalomon/Biton_Salomon_Proofs.pdf> [Accessed 18 December 2020].

Causal Design 2016. CONCUR/IPNN Impact Evaluation. Evaluation brief. Portland, OR: Mercy Corps. Available at: <https://www.mercycorps.org/research-resources/impact-evaluation-conflict-economy> [Accessed 18 December 2020].

Cilliers, J, Dube, O and Siddiqi, B, 2018. *Can the wounds of war be healed? Experimental evidence on reconciliation in Sierra Leone.* 3ie Impact Evaluation Report 75. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: doi: https://doi.org/10.23846/OW2.IE75.

Cilliers, J, Dube, O and Siddiqi, B, 2016. Reconciling after civil conflict increases social capital but decreases individual well-being. *Science*, 352(6287), pp.787–794. Available at: <https://science.sciencemag.org/content/352/6287/787> [Accessed 18 December 2020].

Cleven, E, 2020. Making Connections? A Study of Interethnic Dialogue in a Divided Community in Bosnia and Herzegovina. *Ethnopolitics*, 19:2, 125-149, DOI: 10.1080/17449057.2019.1647666. Dawop, DS, Grady, C, Inks, L and Wolfe, RJ, 2019. Does Peacebuilding Work in the Midst of Conflict? Impact Evaluation of a Peacebuilding Program in Nigeria. Portland, OR: Mercy Corps. Available at: <https://www.mercycorps.org/sites/default/files/2020-01/PRG_NigeriaImpactEvaluation_R_lo_FINv3_Web.pdf> [Accessed 18 December 2020].

Ferguson, N, 2019. *Final Report: Impact Evaluation of Mercy Corps' Programme: Strengthening Social Capital and Reducing Tensions Between Jordanian Host Communities and Syrian Refugees*. Portland, OR: Mercy Corps. Available at: <https://www.mercycorps.org/research-resources/social-stability-jordan> [Accessed 18 December 2020].

Finkel, S, Belasco, C, Gineste, C, Neureiter, M and McCauley, J, 2018. *Peace Through Development II, Burkina Faso, Chad, and Niger: Impact Evaluation Endline Report*. Washington, DC: USAID. Available at: <https://pdf.usaid.gov/pdf_docs/PA00SWPK.pdf> [Accessed 18 December 2020].

Finkel, S, Rojo-Mendoza, R, Shwartz, C, Belasco, C and Kreft, A, 2015. *Impact Evaluation of Peace through Development II (P-DEV II) Radio Programming in Chad and Niger.* Washington, DC: USAID.

Hartman, A, Blair, R, and Blattman, C, 2018. Engineering Informal Institutions: Long-run Impacts of Alternative Dispute Resolution on Violence and Property Rights in Liberia. NBER Working Paper 24482. Cambridge, MA: National Bureau of Economic Research. Available at: <https://www.nber.org/papers/w24482> [Accessed 18 December 2020].

Blattman, C, Hartman, A and Blair, R, 2011. *Can We Teach Peace and Conflict Resolution?: Results from a randomized evaluation of the Community Empowerment Program (CEP) in Liberia: A Program to Build Peace, Human Rights, and Civic Participation.* Evidence from Randomized Evaluations of Peacebuilding in Liberia, Policy Report 2011.2. New Haven, CN: Innovations for Poverty Action. Available at: <https://www.poverty-action.org/sites/default/files/publications/blattman_hartman_blair_can_we_teach_peace_ipa_liberia_0.pdf> [Accessed 18 December 2020].

Blattman, C, Hartman, A and Blair, R, 2012. *Building Institutions at the Micro-Level: Results from a field experiment in property dispute and conflict resolution*. HiCN Working Papers 128. Households in Conflict Network. Available at: <https://ideas.repec.org/p/hic/wpaper/128.html> [Accessed 18 December 2020].

Blattman, C, Hartman, A and Blair, R, 2014. How to promote order and property rights under weak rule of law? An experiment in changing dispute resolution behavior through community education. *The American Political Science Review*, 108(1), 100-120. Available at: <https://www.jstor.org/stable/43654050> [Accessed 18 December 2020].

IMPAQ International, 2017. *Rebuilding interethnic trust in Bosnia and Herzegovina*: Impact and performance evaluation of USAID/BiH Trust, Understanding and Responsibility for the Future (Pro-Future) activity. Prepared under the USAID Monitoring and Evaluation Support Activity (MEASURE-BiH), Contract Number AID-168-C-14-00003. Sarajevo: USAID/Bosnia and Herzegovina.

Lonergan, K, 2016. *Atrocity prevention through reconciliation? Testing the impact of interpersonal reconciliation in Sri Lanka*. Department of Peace and Conflict Research, Uppsala University. Master's thesis, Supervisor: Ralph Sundberg.

Lonergan, K, 2017. *Does reconciliation prevent future atrocities? Evaluating practice in Sri Lanka*. Peaceworks No. 132. Washington, DC: United States Institute of Peace. Available at: <https://www.usip.org/sites/default/files/PW132-Does-Reconciliation-Prevent-Future-Atrocities-Evaluating-Practice-in-Sri-Lanka.pdf> [Accessed 18 December 2020].

Mousa, S, 2019. Building tolerance: Intergroup contact and soccer in post-ISIS Iraq. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab. Available at: < https://static1.squarespace.com/static/59a360bacd0f681b1cb27caa/t/5d7c278ade5a2c25a51fe1a6/1568417682088/mousa-jmp-1.pdf> [Accessed 18 December 2020].

Mousa, S, 2018. *Overcoming the trust deficit: An Experiment on Intergroup Contact and Soccer in Post-ISIS Iraq*. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab. Available at: <https://www.povertyactionlab.org/sites/default/files/publications/Mousa_SoccerIraq_Dec2018.pdf> [Accessed 18 December 2020].

Okunogbe, O, 2018. Does exposure to other ethnic regions promote national integration? Evidence from Nigeria. Policy Research Working Paper 8606. Washington, DC: The World Bank Group.

Paluck, E, 2009. Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology,* 96(3), pp.574–587. Available at: <https://www.researchgate.net/publication/24170533_Reducing_Intergroup_Prejudice_and_Conflict_Using_the_Media_A_Field_Experiment_in_Rwanda> [Accessed 18 December 2020].

Paluck, E, 2007a. *Reducing intergroup prejudice and conflict with the media: A field experiment in Rwanda*. HiCN Working Papers 34. Households in Conflict Network. Available at: <https://ideas.repec.org/p/hic/wpaper/34.html> [Accessed 18 December 2020].

Paluck, E, 2007b. *Reducing intergroup prejudice and conflict with the mass media: A field experiment in Rwanda.* Yale University, ProQuest Dissertations Publishing. Available at: <https://search.proquest.com/openview/3e50057933a6caa1db1526bdedd492ef/1?pq-origsite=gscholar&cbl=18750&diss=y> [Accessed 18 December 2020].

Paluck, E and Green, D, 2009. Deference, dissent, and dispute resolution: An experimental intervention using mass media to change norms and behavior in Rwanda. *American Political Science Review*, 103(4), pp.622–644. Available at: <https://www.researchgate.net/publication/51992793_Deference_Dissent_and_Dispute_Resolution_An_Experimental_Intervention_Using_Mass_Media_to_Change_Norms_and_Behavior_in_Rwanda> [Accessed 18 December 2020].

Rime, B, Kanyangara, P, Yzerbyt, V and Paez, D, 2011. The impact of Gacaca tribunals in Rwanda: Psychosocial effects of participation in a truth and reconciliation process after a genocide. *European Journal of Social Psychology*, 41, 695–706. DOI: 10.1002/ejsp.822.

Kanyangara, P, Rime, B, Philippot, P and Yzerbyt, V, 2007. Collective rituals, emotional climate and intergroup perception: Participation in "gacaca" tribunals and assimilation of the Rwandan genocide. *Journal of social issues*, 63(2), 387-403.

Kanyangara, P, Rime, B, Paez, D and Yzerbyt, V, 2014. Trust, individual guilt, collective guilt and dispositions toward reconciliation among Rwandan survivors and prisoners before and after their participation in postgenocide gacaca courts in Rwanda. *Journal of Social and Political Psychology*, 2(1), 401-416. doi:10.5964/jspp.v2i1.299.

Scacco, A and Warren, S, 2018. Can social contact reduce prejudice and discrimination? Evidence from a field experiment in Nigeria. *American Political Science Review*, 112(3), pp. 654–677. Available at: doi: https://doi.org/10.1017/S0003055418000151.

Scacco, A and Warren, S, 2016. Youth vocational training and conflict mitigation: An experimental test of social contact theory in Nigeria. Working paper. Available at: <https://s18798.pcdn.co/cesspolicon2016/wp-content/uploads/sites/3319/2016/02/Scacco_Warren_UYVT.pdf> [Accessed 18 December 2020].

Schiller, R, 2012. *Talking to the other: The effects of intergroup contact on reconciliation in post-conflict societies.* Student Scholarship, Fletcher School of Law and Diplomacy. Available at: <https://dl.tufts.edu/concern/pdfs/gx41mw53d> [Accessed 18 December 2020].

Svensson, I and Brounéus, K, 2013. Dialogue and interethnic trust: A randomized field trial of 'sustained dialogue' in Ethiopia. *Journal of Peace Research*, 50(5), pp.563–575. Available at: doi: https://doi.org/10.1177/0022343313492989.

Vicente, P and Vilela I, 2020. Preventing Violent Islamic Radicalization: Behavioral Evidence from Northern Mozambique. NOVAFRICA Working Paper Series, Working Paper No 2008. Lisbon: Universidade Nova de Lisboa. Available at: <https://novafrica.org/research/preventing-violent-islamic-radicalization-experimental-evidence-on-anti-social-behavior/> [Accessed 18 December 2020].

Ongoing studies

Brück, T and Nillesen, E, 2014. *Impact Evaluation of peace building educational programme in Kyrgyzstan [Protocol].* Registry for International Development Impact Evaluations, RIDIE-STUDY-ID-5310e1448f95b. Available at: <https://ridie.3ieimpact.org/index.php?r=search/advancedSearchDetailView&id=176> [Accessed 18 December 2020].

CAF, Pilar, Maris, L and Rodríguez, C, 2018. *Building Citizen Competencies from the Classroom: Experimental Evidence from Colombia.* American Economic Association RCT Registry. Available at: doi: https://doi.org/10.1257/rct.3116-1.0.

ECAS Consulting and Mercy Corps 2016. Mercy Corps: Baseline Evaluation - PROSPER. Available at: <https://www.dmeforpeace.org/resource/baseline-context-assessment-promoting-sustainable-peace-and-resilience-prosper-in-kayah-state/> [Accessed 18 December 2020].

Ferguson, L, Ibanez, AM, Larreguy, H and Moya, A, 2017. *Promoting Social Cohesion after Conflict: Should We Inform or Engage Emotions? [Protocol].* EGAP Registry. Available at: <http://egap.org/sites/default/files/Larreguy_EGAP20.pdf> [Accessed 18 December 2020].

Spitler, A and Roessler, P, 2018. *Evaluating the Effectiveness of Sport as a Pedagogical Tool: A Year Two Impact Evaluation of the Monrovia Football Academy [Protocol].* Registry for International Development Impact Evaluations, RIDIE-STUDY-ID-5b55455f20bcb. Available at: <https://ridie.3ieimpact.org/index.php?r=search/advancedSearchDetailView&id=715> [Accessed 18 December 2020].

Other references

Allport, GW, 1954. *The nature of prejudice*. Cambridge: Addison-Wesley Publishing Company.

Barron, P, Diprose, R and Woolcock, M, 2006. *Local conflict and community development in Indonesia: assessing the impact of the Kecamatan development program* (English). Indonesian social development paper no. 10. Washington, D.C. : World Bank Group. Available at: <http://documents.worldbank.org/curated/en/487771468269091358/Local-conflict-and-community-development-in-Indonesia-assessing-the-impact-of-the-Kecamatan-development-program> [Accessed 18 December 2020].

Baumgardner-Zuzik, J, DeYoung, S, Myers, E, Thabit, K and Vetsch, C, 2020. Eirene Peacebuilding Database [database]. Washington, DC: Alliance for Peacebuilding. Available at: <https://allianceforpeacebuilding.org/our-work/about-our-work/peacebuilding-evaluation/eirene-peacebuilding-database> [Accessed 20 February 2020].

Berghof Foundation, 2019. *Berghof Glossary on Conflict Transformation: 20 notions for theory and practice.* Berlin: Berghof Foundation. Available at: <https://berghof-foundation.org/library/berghof-glossary> [Accessed 11 December 2020].

Blair, G and Imai, K, 2012. Statistical analysis of list experiments. *Political Analysis*, 20(1), pp.47–77.

Borenstein, M, 2009. Effect sizes for continuous data. In H Cooper, LV Hedges and JC Valentine (Eds.), *The handbook of research synthesis and meta-analysis*, 221–235. Russell Sage Foundation.

Chan, J, To, H-P and Chan, E, 2006. Reconsidering social cohesion: developing a definition and analytical framework for empirical research. *Social Indicators Research*, 75, pp.273–302.

Colletta, NJ and Cullen, ML, 2000. *The Nexus between Violent Conflict, Social Capital and Social Cohesion: Case Studies from Cambodia and Rwanda.* World Bank, Social Development Family, Environmentally and Socially Sustainable Development Network.

Czuba, K, O'Neill, TJ and Ayala, AP, 2017. *The Impact of Food Assistance on Pastoralist Livelihoods in Humanitarian Crises: An evidence synthesis*. Humanitarian Evidence Programme. Oxford: Oxfam GB. Available at: <https://fic.tufts.edu/publication-item/impacts-of-in-kind-food-assistance/> [Accessed 10 February 2020].

Dayal, AK and Christien, A, 2020. Women's participation in informal peace processes. *Global Governance*, 26, pp.69–98.

De Berry, JP and Roberts, AJ, 2018. *Social cohesion and forced displacement: a desk review to inform programming and project design (English).* Washington, DC: World Bank Group. Available at: <http://documents.worldbank.org/curated/en/125521531981681035/Social-cohesion-and-forced-displacement-a-desk-review-to-inform-programming-and-project-design> [Accessed 18 December 2020].

Deneckere, M and Hauck, V, 2018. *Supporting peacebuilding in times of change.* Maastricht: ECDPM. Available at: <https://ecdpm.org//wp-content/uploads/ECDPM-September-2018-Supporting-peacebuilding-times-change-case-study-Germany.pdf> [Accessed 25 April 2020].

Department for International Development (DFID), 2019. DFID Annual Report and Accounts 2018 to 2019. Available at: <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/815787/ARA-2019.pdf> [Accessed 18 December 2020].

Dragalov, G, Ignácz, Z, Lorenz, J, Delhey, J and Boehnke, K, 2013. *Social Cohesion Radar. Measuring Common Ground.* Gütersloh: Bertelsmann Stifting.

Dugarova, E, 2015. *Social inclusion, poverty eradication and the 2030 agenda for sustainable development*. UNRISD Working Paper No. 2015-15. Geneva: United Nations Research Institute for Social Development.

Durkheim, É, 1897. *Le Suicide: Étude de Sociologie.* Paris: Félix Alcan.

ECAS Consulting and Mercy Corps, 2016. Mercy Corps: Baseline Evaluation - PROSPER. Available at: <https://www.dmeforpeace.org/resource/baseline-context-assessment-promoting-sustainable-peace-and-resilience-prosper-in-kayah-state/> [Accessed 18 December 2020].

Egger, M, Davey Smith, G, Schneider, M and Minder, C, 1997. Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315, 629–634.

Ellis, PD, 2010. *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge: Cambridge University Press.

Fearon, JD, Humphreys, M and Weinstein, JM, 2009. Can development aid contribute to social cohesion after civil war? Evidence from a field experiment in post-conflict Liberia. *American Economic Review: Papers and Proceedings*, 99(2), pp.287–291.

Federal Ministry for Economic Cooperation and Development (BMZ), 2015. OECD DAC Peer Review of Germany 2015 - Memorandum by Germany, BMZ, Bonn. Available at: <https://www.oecd.org/dac/peer-reviews/Memorandum%20for%20DAC%20Peer%20Review%20of%20Germany%202015%20ENG.pdf> [Accessed 18 December 2020].

Federal Ministry for Economic Cooperation and Development (BMZ), 2020. *Strategy on transitional development assistance*: Overcoming crises - strengthening resilience - creating new prospects. BMZ Document 02 | 2020. Bonn: Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung. Available at: https://www.bmz.de/en/publications/topics/development_policy/Strategiepapier505_01_2020.pdf [Accessed 18 December 2020].

Ferguson, K, 2016. *Countering violent extremism through media and communication strategies: A review of the evidence.* Norwich, UK: Partnership for Conflict, Crime and Security Research. Available at: <https://www.paccsresearch.org.uk/wp-content/uploads/2016/03/Countering-Violent-Extremism-Through-Media-and-Communication-Strategies-.pdf> [Accessed 18 August 2020].

Freedom House 2020, *Freedom in the World Research Methodology*. Available at: <https://freedomhouse.org/reports/freedom-world/freedom-world-research-methodology> [Accessed 18 December 2020].

Fund for Peace (FFP), 2019a. Fragile States Index – Global Data. Washington, DC: FFP. Available at: <https://fragilestatesindex.org/data> [Accessed 12 April 2019].

Giddens, A, 1990. *The Consequences of Modernity*. Stanford: Stanford University Press.

Giessmann, HJ, Schädel, A and Feda, B, 2019. Fostering human security. In: B. Austin, H.J. Giessmann, A. Schädel and A. Annan (Eds.), *Berghof Glossary on Conflict Transformation and Peacebuilding: 20 essays on theory and practice*. Berlin: Berghof Foundation. Available at: <https://berghof-foundation.org/library/berghof-glossary> [Accessed 18 December 2020].

GIZ, n.d. *Open Fund for Social Cohension* [project page]. Available at: <https://www.giz.de/projektdaten/projects.action;jsessionid=763189195FBEB8A2458F385BEF83DFC1?pn=201720226&request_locale=en_GB> [Accessed 18 December 2020].

GIZ, 2020. GIZ@AidEx: Showcasing Transitional Development Assistance in Chad and Turkey. Available at: <https://www.giz.de/en/worldwide/81368.html> [Accessed 18 December 2020].

Goldacre, B, Drysdale, H, Powell-Smith, A, Dale, A, Milosevic, I, Slade, E, Hartley, P, Marston, C, Mahtani, K and Heneghan, C, 2016. The COMPare Trials Project. <www.COMPare-trials.org> [Accessed 15 April 2020].

Harris, IM and Morrison, ML, 2013. *Peace Education*. 3rd edition. Jefferson, NC: McFarland and Co., Inc.

Higgins, J and Green, S, 2011. *Cochrane handbook for systematic reviews of interventions.* (version 5.0.2, updated September 2009). The Cochrane Collaboration.

Higgins, JPT, Sterne, JAC, Savović, J, Page, MJ, Hróbjartsson, A, Boutron, I, Reeves, B and Eldridge, S, 2016. A revised tool for assessing risk of bias in randomized trials. In: J Chandler, J McKenzie, I Boutron and V Welch (eds). *Cochrane Methods. Cochrane Database of Systematic Reviews*, Issue 10 (Suppl 1). dx.doi.org/10.1002/14651858.CD201601.

Hedges, LV, Tipton, E and Johnson, MC, 2010. Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), pp.39–65.

Hewstone, M and Swart, H, 2011. Fifty-odd years of inter-group contact: From hypothesis to integrated theory. *British Journal of Social Psychology*, 50(3), 374–386. Available at: doi: https://doi.org/10.1111/j.2044-8309.2011.02047.x.

Hombrados, JG and Waddington, H, 2012. A tool to assess risk of bias for experiments and quasi-experiments in development research. Mimeo. New Delhi: The International Initiative for Impact Evaluation.

Hymel, S and Darwich, L, 2018. Building peace through education. *Journal of Peace Education*, 15(3), pp.345–357. Available at: doi: https://doi.org/10.1080/17400201.2018.1535475.

Idris, I, 2016. *Building Social Cohesion in Post-Conflict Situations.* GSDRC Helpdesk Research Report 1332. Birmingham, UK: GSDRC, University of Birmingham.

Jensen, J, 2019. *Intersections of Pluralism and Social Cohesion.* Intersections: Practicing Pluralism, 3. Ottawa: Global Centre for Pluralism.

Kaplan, S, 2009. Identity in fragile states: social cohesion and state building. *Development*, 52(4), pp.466–472.

Kim, J, Sheely, R and Schmidt, C, 2020. *Social Capital and Social Cohesion Measurement Toolkit for Community-Driven Development Operations.* Portland, OR: Mercy Corps. Available at: <https://collaboration.worldbank.org/content/usergenerated/asi/cloud/attachments/sites/collaboration-for-development/en/groups/community-driven-development-global-solutions-group/documents/jcr:content/content/primary/blog/social_capital_ands-OhaN/Social%20Capital%20and%20Social%20Cohesion%20Measurement%20Toolkit%20(Final).pdf> [Accessed 3 January 2020].

King, E, Samii, C and Snilstveit, B, 2010a. Interventions to promote social cohesion in sub-Saharan Africa. *Journal of Development Effectiveness*, 2(3), pp.336–370. Available at: doi: https://doi.org/10.1080/17449057.2010.504552.

King, E, Samii, C and Snilstveit, B, 2010b. *Interventions to promote social cohesion in sub-Saharan Africa.* 3ie Synthetic Review 002. Available at: doi: https://doi.org/10.23846/sr1012 [Accessed 10 February 2020].

King, E and Samii, C, 2014. Fast-track institution building in conflict-affected countries? Insights from recent field experiments. *World Development*, 64,740-754.

La Ferrara, E, 2016. *Media as a tool for institutional change in development.* EDI Working Paper Series WP16/10.III.4. Oxford: Economic Development and Institutions (EDI). Available at: <https://edi.opml.co.uk/wpcms/wp-content/uploads/2017/06/EDI-PF-PAPER-10.III_.4-La-Ferrera.pdf> [Accessed 11 December 2019].

Lockwood, D, 1999. Civic Integration and Social Cohesion. In: I Gough and G Olofsson (Eds.), *Capitalism and Social Cohesion*, 63-84. New York: Palgrave Macmillan.

Malhotra, D and Liyange, S, 2005. Long-term effects of peace workshops in protracted conflicts. *Journal of Conflict Resolution*, 49(6), 908-924. DOI: 10.1177/0022002705281153.

Marc, A, Willman, A, Aslam, G, Rebosio, M and Balasuriya, K, 2013. *Societal Dynamics and Fragility. Engaging Societies in Responding to Fragile Situations.* Washington, DC: World Bank Group.

McCauley, C, 2002. Head-first versus feet-first in peace education. In: G Salomon and B Nevo, eds. *Peace Education: The concept, principles, and practices around the world*. Mahwah, NJ: Erlbaum, pp. 247–257.

Mercy Corps, 2013. Mapping of host community-refugee tensions in Mafraq and Ramtha, Jordan. Aman: Mercy Corps.

Moayed, T, 2019. *Recognising Communities: Local Level Responses to the Pathfinders' Grand Challenge.* Conference Discussion Paper. World Bank Group.

Moehler, D, 2014. *Democracy, Governance and Randomised Media Assistance.* London: BBC Media Action. Available at: <http://downloads.bbc.co.uk/rmhttp/mediaaction/pdf/democracy-governance-research-report.pdf> [Accessed 10 February 2020].

Moher, D, Liberati, A, Tetzlaff, J, Altman, DG and The PRISMA Group, 2009. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA Statement. *PLoS Medicine*, 6(7), e1000097. Available at: doi: https://doi.org/10.1371/journal.pmed.1000097 [Accessed 12 December 2019].

Munafò, M and Smith, GD, 2018. Repeating experiments is not enough. *Nature*, 553, pp.399–401.

Nielsen, I, Gu, J, Shachat, J, Smyth, R and Peng, Y, 2016. An experimental study of the effect of intergroup contact on attitudes in urban China. *Urban studies*, 53(14), 2991-3006. Available at: doi: http://dx.doi.org/10.1177/0042098015598730.

Nolan, C and Knox, A, 2019. *Governance, Crime and Conflict Initiative*. Boston: Abdul Latif Jameel Poverty Action Lab (J-PAL) and Innovations for Poverty Action (IPA). Available at: <https://www.povertyactionlab.org/sites/default/files/documents/gcci-evidence-review-web-july-2019.pdf> [Accessed 12 December 2019].

Obiagu, AN, Mezieobi, DI, Aroh, PN and Akubue, FN, 2020. The effect of cooperative concept mapping on misconceptions, knowledge achievement, and transfer of learning in peace education. *The Social Studies*, 111(1), 18-38, DOI: 10.1080/00377996.2019.1652139.

Office for the Coordination of Humanitarian Affairs (OCHA), 2019. *Global Humanitarian Overview 2019: United Nations Coordinated Support to People Affected by Disaster and Conflict*. New York: United Nations Office for the Coordination of Humanitarian Affairs. Available at: <https://www.unocha.org/sites/unocha/files/GHO2019.pdf> [Accessed 18 March 2019].

OECD, 2012. Conflict analysis and its use in evaluation. Chapter in: *Evaluating Peacebuilding Activities in Settings of Conflict and Fragility: Improving Learning for Results*. Paris: OECD Publishing. Available at: doi: https://doi.org/10.1787/9789264106802-9-en [Accessed 10 February 2020].

OECD, 2018. *States of Fragility 2018*. Paris: OECD Publishing. Available at: doi: https://doi.org/10.1787/9789264302075-en [Accessed 10 February 2020].

OECD, 2019. Data warehouse, OECD.Stat (database). Available at: doi: <https://doi.org/10.1787/data-00900-en> [Accessed 13 September 2019].

O'Neill, J, Tabish, H, Welch, V, Petticrew, M, Pottie, K, Clarke, M, Evans, T, Pardo Pardo, J, Waters, E, White, H and Tugwell, P, 2014. Applying an equity lens to interventions: using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *Journal of Clinical Epidemiology*, 67(1), 56–64. Available at: doi: https://doi.org/10.1016/j.jclinepi.2013.08.005.

O'Reilly, M, O' Suilleabhain, A and Paffenholz, T, 2017. *Reimagining Peacemaking: Women's Roles in Peace Processes.* New York: International Peace Institute. Available at: <https://www.ipinst.org/wp-content/uploads/2015/06/IPI-E-pub-Reimagining-Peacemaking-rev.pdf> [Accessed 15 June 2020].

Paluck, EL, Green, SA and Green, DP, 2019. The contact hypothesis re-evaluated. *Behavioural Public Policy*, 3(2), 129–158.

Peters, J L, Sutton, AJ, Jones, DR, Abrams, KR and Rushton, L, 2008. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. Journal of clinical epidemiology, 61(10), 991–996. Available at: doi: https://doi.org/10.1016/j.jclinepi.2007.11.010.

Pettigrew, T and Hewstone, M, 2017. The Single Factor Fallacy: Implications of Missing Critical Variables from an Analysis of Intergroup Contact Theory. *Social Issues and Policy Review*, 11(1), 8–37. Available at: doi: https://doi.org/10.1111/sipr.12026 [Accessed 20 April 2020].

Pettigrew, TF and Tropp, LR, 2005. Allport's Intergroup Contact Hypothesis: Its History and Influence. In: Dovidio, J, Glick, P and Budman, L, eds. *On the Nature of Prejudice: Fifty Years After Allport*. Oxford: Blackwell Publishing.

Pettigrew, TF and Tropp, LR, 2006. A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90(5), pp.751–783.

Pettigrew, TF and Tropp, LR, 2008. How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology,* 38, 922-934. Available at: doi: https://doi.org/10.1002/ejsp.504.

Quinn, JR, 2020. Tractionless transitional justice in Uganda: The potential for thin sympathetic interventions as ameliorating factor. In: El-Masri, S, Lambert, T and Quinn, JR, eds. *Transitional Justice in Comparative Perspective*. Cham: Springer Nature Switzerland AG, pp.19–48.

R Development Core Team (2010). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at: <http://www.R-project.org/>.

Sanchez-Meca, J, Marín-Martínez, F and Moscoso, S, 2003. Effect-Size Indices for Dichotomized Outcomes in Meta-Analysis. *Psychological methods,* 8(4), 448-467. DOI:10.1037/1082-989X.8.4.448.

Sangalang, CC and Vang, C, 2017. Intergenerational Trauma in Refugee Families: A Systematic Review. *Journal of Immigrant Minority Health*, 19, 745–754. Available at: doi: https://doi.org/10.1007/s10903-016-0499-7.

Schrader, PG and Lawless, KA, 2004. The knowledge, attitudes, and behaviors approach how to evaluate performance and learning in complex environments. *Performance Improvement*, 43, pp.8–15. Available at: doi: https://doi.org/10.1002/pfi.4140430905.

Siddique, S, 2001. Social cohesion and social conflict in Southeast Asia. In N. T. Colletta, T. G. Lim, and A. Kelles-Viitanen (Eds.), *Social cohesion and conflict prevention in Asia*. Washington DC: The World Bank.

Snilsveit, B, Stevenson, J, Langer, L, da Silva, N, Rabat, Z, Nduku, P, Polanin, J, Shemilt, I, Eyers, J and Ferraro, PJ, 2019. *Incentives for climate mitigation in the land use sector – the effects of payment for environmental services (PES) on environmental and socio-economic outcomes in low- and middle-income countries: a mixed-method systematic review*. 3ie Systematic Review 44. London: International Initiative for Impact Evaluation (3ie). Available at: https://doi.org/10.23846/SR00044.

Sonnenfeld, A, Chirgwin, H, Berretta, M, Longman, K, Krämer, M and Snilstveit, B, 2020. *Building peaceful societies: an evidence gap map.* 3ie Evidence Gap Map Report 15. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: doi: https://doi.org/10.23846/EGM015 [Accessed 30 April 2020].

Sönsken, S, Kruck, A and El-Nahel, Z, 2019. Building and sustaining peace. In: B. Austin, H.J. Giessmann, A. Schädel and A. Annan (Eds.), *Berghof Glossary on Conflict Transformation and Peacebuilding: 20 essays on theory and practice*. Berlin: Berghof Foundation Operations GmbH. Available at: https://berghof-foundation.org/files/publications/Berghof_Glossary_2019_eng.pdf

State and Peacebuilding Fund (SPF) (2018). State and Peacebuilding Fund annual report. Washington DC: The World Bank. Available at http://pubdocs.worldbank.org/en/510271559335344163/SPF-Annual-Report-2018.pdf. [Accessed 19 July 2019].

Sterne, JAC, Hernán, MA, Reeves, BC, Savović, J, Berkman, ND, Viswanathan, M, Henry, D, Altman, DG, Ansari, MT, Boutron, I, Carpenter, JR, Chan, AW, Churchill, R, Deeks, JJ, Hróbjartsson, A, Kirkham, J, Jüni, P, Loke, YK, Pigott, TD, Ramsay, CR, Regidor, D, Rothstein, HR, Sandhu, L, Santaguida, PL, Schünemann, HJ, Shea, B, Shrier, I, Tugwell, P, Turner, L, Valentine, JC, Waddington, H, Waters, E, Wells, GA, Whiting, PF and Higgins, JPT, 2016. ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. *British Medical Journal,* 355; i4919; doi: 10.1136/bmj.i4919.

Tankard, M and Paluck, E, 2016. Norm perception as a vehicle for social change. *Social Issues and Policy Review*, 10(1), pp. 181–211. Available at: <https://socialbyselection.files.wordpress.com/2016/10/a0394-tankardpaluck2016.pdf> [Accessed 13 February 2020].

UK Aid, 2015. Strategic Defence and Security Review and UK Aid Strategy. London: HM Government. Available at <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/555607/2015_Strategic_Defence_and_Security_Review.pdf> [Accessed 16 July 2019].

United Nations Department of Economic and Social Affairs (UNDESA), n.d. SDG 16. Available at: <https://sdgs.un.org/goals/goal16> [Accessed 11 December 2020].

United Nations, Department of Public Information, 2016. General Assembly Adopts Resolution on Improving United Nations Peacebuilding Architecture Ahead of Day-long Debate. GA/11780, 27 April 2016, Available at: <https://www.un.org/press/en/2016/ga11780.doc.htm> [Accessed 18 April 2019].

United National Peacebuilding Fund (PBF) (2020). PBF Investments. <https://www.un.org/peacebuilding/fund/documents/investments> [Accessed 18 December 2020]

United Nations (UN) and World Bank, 2018. *Pathways for Peace: Inclusive Approaches to Preventing Violent Conflict.* Washington, DC: World Bank. Available at: <https://openknowledge.worldbank.org/handle/10986/28337> [Accessed 18 April 2019].

Urdal, H, 2006. A Clash of Generations? Youth Bulges and Political Violence. *International Studies Quarterly* 50(3), pp.607–629.

Valentine, JC and Aloe, AM, 2016. How to communicate effect sizes for continuous outcomes: a review of existing options and introducing a new metric. *Journal of Clinical Epidemiology,* 72:84-89. doi:10.1016/j.jclinepi.2015.10.017

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. *Journal of statistical software*, 36(3), 1-48.

Waddington, H, White, H, Snilstveit, B, Hombrados Garcia, J, Vojtkova, M, Davies, P, Bhavsar, A, Eyers, J, Perez Koehlmoos, T, Petticrew, M, Valentine, JC and Tugwell, P, 2012. How to do a good systematic review of effects in international development: a toolkit. *Journal of Development Effectiveness*, 4(3), 359-387.

White, H, 2013. Using the causal chain to make sense of the numbers. *Evidence Matters* blog, 12 February. International Initiative for Impact Evaluation (3ie). Available at: <https://www.3ieimpact.org/blogs/using-causal-chain-make-sense-numbers> [Accessed 18 December 2020].

White, H, Menon, R and Waddington, H, 2018. *Community-driven development: Does it build social cohesion or infrastructure? A mixed-methods evidence synthesis*. 3ie Working Paper 30. New Delhi: International Initiative for Impact Evaluation (3ie).

World Bank, n.d. *World Bank Country and Lending Groups.* Available at: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups> [Accessed 16 December 2020].

Zembylas, M, Charalambous, C and Charalambous, P, 2016. *Peace Education in a Conflict-Affected Society.* Cambridge: Cambridge University Press.

Other publications in the 3ie Systematic Review Series

The following reviews are available at
http://www.3ieimpact.org/evidence-hub/publications/systematic-reviews

Effects of electricity access interventions on socio-economic outcomes in low- and middle-income countries. 3ie Systematic Review 45. Moore, N, Glandon, D, Tripney, J, Kozakiewicz, T, Shisler, S, Eyers, J, Zalfou, R, Leon, MDA, Kurkjian, V, Snilstveit, B and Perdana, A. 2020.

Incentives for climate mitigation in the land use sector – the effects of payment for environmental services (PES) on environmental and socio-economic outcomes in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 44. Snilstveit, B, Stevenson, J, Langer, L, da Silva, N, Rabat, Z, Nduku, P, Polanin, J, Shemilt, I, Eyers, J, Ferraro, PJ, 2019.

Does incorporating participation and accountability improve development outcomes? Meta-analysis and framework synthesis. 3ie Systematic Review 43. Waddington, H, Sonnenfeld, A, Finetti, J, Gaarder, M and Stevenson, J, 2019.

Impact of financial inclusion in low- and middle-income countries: a systematic review of reviews. 3ie Systematic Review 42. Duvendack, M and Mader, P, 2019.

Agricultural input subsidies for improving productivity, farm income, consumer welfare and wider growth in low- and middle-income countries: a systematic review. 3ie Systematic Review 41. Hemming, DJ, Chirwa, EW, Ruffhead, HJ, Hill, R, Osborn, J, Langer, L, Harman, L, Coffey, C, Dorward, A and Phillips, D, 2018.

Vocational and business training to improve women's labour market outcomes in low- and middle-income countries: a systematic review. 3ie Systematic Review 40. Chinen, M, De Hoop, T, Balarin, M, Alcázar, L, Sennett, J, and Mezarina, J, 2018.

Interventions to improve the labour market for adults living with physical and/or sensory disabilities in low- and middle-income countries: a systematic review. 3ie Systematic Review 39. Tripney, J, Roulstone, A, Vigurs, C, Hogrebe, N, Schmidt, E and Stewart, R, 2017.

The effectiveness of contract farming in improving smallholder income and food security in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 38. Ton, G, Desiere, S, Vellema, W, Weituschat, S and D'Haese, M (2017)

Interventions to improve the labour market outcomes of youth: a systematic review of training, entrepreneurship promotion, employment services and subsidized employment interventions. 3ie Systematic Review 37. Kluve J, Puerto S, Robalino D, Romero JM, Rother F, Stöterau J, Weidenkaff F and Witte M (2017)

Promoting handwashing and sanitation behaviour change in low- and middle-income countries: a mixed-method systematic review. 3ie Systematic Review 36. Buck, ED, Remoortel, HV, Hannes, K, Govender, T, Naidoo, S, Avau, B, Veegaete, AV, Musekiwa, A, Lutje, V, Cargo, M, Mosler, HJ, Vandekerckhove, P and Young T (2017)

Incorporating the life cycle approach into WASH policies and programmes: A systematic review. 3ie Systematic Review 35. Annamalai, TR, Narayanan, S, Devkar, G, Kumar, VS, Devaraj, R, Ayyangar, A and Mahalingam, A (2017)

Effects of certification schemes for agricultural production on socio-economic outcomes in low- and middle-income countries: a systematic review 34. Oya, C, Schaefer, F, Skalidou, D, McCosker, C and Langer, L (2017)

Short-term WASH interventions in emergency response: a systematic review. 3ie Systematic Review 33. Yates, T, Allen, J, Joseph, ML and Lantagne, D (2017)

Community monitoring interventions to curb corruption and increase access and quality of service delivery in low- and middle-income countries. 3ie Systematic Review 32. Molina E, Carella L, Pacheco A, Cruces, G and Gasparini, L (2016)

Effects and mechanisms of market-based reforms on access to electricity in developing countries: a systematic review. 3ie Systematic Review 31. Bensch, G, Sievert, M, Langbein, J, Kneppel, N (2016)

Youth gang violence and preventative measures in low- and middle-income countries: a systematic review (Part II), 3ie Systematic Review 30. Higginson, A, Benier, K, Shenderovich, Y, Bedford, L, Mazerolle, L, Murray, J (2016)

Youth gang membership and violence in low- and middle-income countries: a systematic review (Part I), 3ie Systematic Review 29. Higginson, A, Benier, K, Shenderovich, Y, Bedford, L, Mazerolle, L, Murray, J (2016)

Cash-based approaches in humanitarian emergencies: a systematic review, 3ie Systematic Review Report 28. Doocy, S and Tappis, H (2016)

Factors affecting uptake of voluntary and community-based health insurance schemes in low-and middle-income countries: a systematic review, 3ie Systematic Review 27. Panda, P, Dror, IH, Koehlmoos, TP, Hossain, SAS, John, D, Khan, JAM and Dror, DM (2016)

Parental, community and familial support interventions to improve children's literacy in developing countries: a systematic review, 3ie Systematic Review 26. Spier, E, Britto, P, Pigott, T, Roehlkapartain, E, McCarthy, M, Kidron, Y, Song, M, Scales, P, Wagner, D, Lane, J and Glover, J (2016)

Business support for small and medium enterprises in low- and middle-income countries: a systematic review, 3ie Systematic Review 25. Piza, C, Cravo, T, Taylor, L, Gonzalez, L, Musse, I, Furtado, I, Sierra, AC and Abdelnour, S (2016)

Interventions for improving learning outcomes and access to education in low- and middle- income countries: a systematic review, 3ie Systematic Review 24. Snilstveit, B, Stevenson, J, Phillips, D, Vojtkova, M, Gallagher, E, Schmidt, T, Jobse, H, Geelen, M, Pastorello, M, and Eyers, J (2015)

Economic self-help group programmes for improving women's empowerment: a systematic review, 3ie Systematic Review 23. Brody, C, De Hoop, T, Vojtkova, M, Warnock, R, Dunbar, M, Murthy, P and Dworkin, SL (2016)

The identification and measurement of health-related spillovers in impact evaluations: a systematic review, 3ie Systematic Review 22. Benjamin-Chung, J, Abedin, J, Berger, D, Clark, A, Falcao, L, Jimenez, V, Konagaya, E, Tran, D, Arnold, B, Hubbard, A, Luby, S, Miguel, E and Colford, J (2015)

The effects of school-based decision-making on educational outcomes in low- and middle-income countries: a systematic review, 3ie Systematic Review Report 21. Carr-Hill, R, Rolleston, C, Pherali, T and Schendel, R, with Peart, E, and Jones, E (2015)

Policing interventions for targeting interpersonal violence in developing countries: a systematic review, 3ie Systematic Review 20. Higginson, A, Mazerolle, L, Sydes, M, Davis, J, and Mengersen, K (2015)

The effects of training, innovation and new technology on African smallholder farmers' wealth and food security: a systematic review, 3ie Systematic Review 19. Stewart, R, Langer, L, Rebelo Da Silva N, Muchiri, E, Zaranyika, H, Erasmus, Y, Randall, N, Rafferty, S, Korth, M, Madinga, N and de Wet, T (2015)

Community based rehabilitation for people with disabilities in low- and middle-income countries: a systematic review, 3ie Systematic Review 18. Iemmi, V, Gibson, L, Blanchet, K, Kumar, KS, Rath, S, Hartley, S, Murthy, GVS, Patel, V, Weber, J and Kuper H (2015)

Payment for environmental services for reducing deforestation and poverty in low- and middle-income countries: a systematic review, 3ie Systematic Review 17. Samii, C, Lisiecki, M, Kulkarni, P, Paler, L and Chavis, L (2015)

Decentralised forest management for reducing deforestation and poverty in low- and middle- income countries: a systematic review, 3ie Systematic Review 16. Samii, C, Lisiecki, M, Kulkarni, P, Paler, L and Chavis, L (2015)

Supplementary feeding for improving the health of disadvantaged infants and young children: a systematic and realist review, 3ie Systematic Review 15. Kristjansson, E, Francis, D, Liberato, S, Greenhalgh, T, Welch, V, Jandu, MB, Batal, M, Rader, T, Noonan, E, Janzen, L, Shea, B, Wells, GA and Petticrew, M (2015)

The impact of land property rights interventions on investment and agricultural productivity in developing countries: a systematic review, 3ie Systematic Review Report 14. Lawry, S, Samii, C, Hall, R, Leopold, A, Hornby, D and Mtero, F, 2014.

Slum upgrading strategies and their effects on health and socio-economic outcomes: a systematic review, 3ie Systematic Review 13. Turley, R, Saith, R., Bhan, N, Rehfuess, E, and Carter, B (2013)

Services for street-connected children and young people in low- and middle-income countries: a thematic synthesis, 3ie Systematic Review 12. Coren, E, Hossain, R, Ramsbotham, K, Martin, AJ and Pardo, JP (2014)

Why targeting matters: examining the relationship between selection, participation and outcomes in farmer field school programmes, 3ie Systematic Review 11. Phillips, D, Waddington, H and White, H (2015)

The impact of export processing zones on employment, wages and labour conditions in developing countries, 3ie Systematic Review 10. Cirera, X and Lakshman, R (2014)

Interventions to reduce the prevalence of female genital mutilation/cutting in African countries, 3ie Systematic Review 9. Berg, RC and Denision, E (2013)

Behaviour change interventions to prevent HIV among women living in low and middle income countries, 3ie Systematic Review 8. McCoy, S, Kangwende, RA and Padian, NS (2009)

The impact of daycare programs on child health, nutrition and development in developing countries, 3ie Systematic Review 7. Leroy, JL, Gadsden, P and Guijarro, M (2011)

Willingness to pay for cleaner water in less developed countries: Systematic review of experimental evidence, 3ie Systematic Review 6. Null, C, Hombrados, JG, Kremer, M, Meeks, R, Miguel, E and Zwane, AP (2012)

Community-based intervention packages for reducing maternal morbidity and mortality and improving neonatal outcomes, 3ie Systematic Review 5. Lassi, ZS, Haider, BA and Langou, GD (2011)

The effects of microcredit on women's control over household spending: a systematic review, 3ie Systematic Review 4. Vaessen, J, Rivas, A, Duvendack, M, Jones, RP, Leeuw, F, van Gils, G, Lukach, R, Holvoet, N, Bastiaensen, J, Hombrados, JG and Waddington, H, (2013).

Interventions in developing nations for improving primary and secondary school enrolment of children: a systematic review, 3ie Systematic Review 3. Petrosino, A, Morgan, C, Fronius, T, Tanner-Smith, E, and Boruch, R, 2016.

Interventions to promote social cohesion in Sub-Saharan Africa, 3ie Systematic Review 2. King, E, Samii, C and Snilstveit, B (2010)

Water, sanitation and hygiene interventions to combat childhood diarrhoea in developing countries, 3ie Systematic Review 1. Waddington, H, Snilstveit, B, White, H and Fewtrell, L (2009**)**