Jacobus Cilliers
Oeindrila Dube
Bilal Siddiqi

Can the wounds of war be healed? Experimental evidence on reconciliation in Sierra Leone

May 2018

Impact Evaluation Report 75

Governance

3ie 10 YEARS
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 impact evaluations

3ie-supported impact evaluations assess the difference a development intervention has made to social and economic outcomes. 3ie is committed to funding rigorous evaluations that include a theory-based design, and use the most appropriate mix of methods to capture outcomes and are useful in complex development contexts.

About this report

3ie accepted the final version of the report, *Can the wounds of war be healed? Experimental evidence on reconciliation in Sierra Leone,* as partial fulfilment of requirements of grant OW2.253 awarded through Open Window 2. The content has been copy-edited and formatted for publication by 3ie. Despite best efforts in working with the authors, some figure and tables could not be improved or references fully corrected. We have copy-edited the content to the extent possible.

The 3ie technical quality assurance team for this report comprises Tara Kaul, Radhika Menon, an anonymous external impact evaluation design expert reviewer and an anonymous external sector expert reviewer, with overall technical supervision by Jyotsna Puri. The 3ie editorial production team for this report comprises Angel Kharya and Akarsh Gupta, with Beryl Leach providing overall editorial supervision.

Funding for this impact evaluation was provided by 3ie's donors, which include UK aid through the Department of International Development, the Bill & Melinda Gates Foundation and the Hewlett Foundation. A complete listing of all of 3ie's donors is available on the 3ie website.

Suggested citation: 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

Cover photo: Lee Karen Stow / Alamy Stock Photo posed by model

Can the wounds of war be healed? Experimental evidence on reconciliation in Sierra Leone

Jacobus Cilliers
Georgetown University

Oeindrila Dube
New York University

Bilal Siddiqi
World Bank

3ie Impact Evaluation Report 75

May 2018

Acknowledgements

We would like to thank John Caulker from Fambul Tok and Libby Hoffman from Catalyst for Peace for allowing us to evaluate the Fambul Tok programme. Many helpful comments were provided by Eric Dixon, James Fearon, Michael Gilligan, David Laitin, Sendhil Mullainathan, David Stasavage and participants at the Michigan Development Seminar, the Princeton-Warwick Political Economy Conference, the Princeton Political Economy Seminar, Inter-American Development Bank, the Villars Conflict and Development Conference, American Political Science Association annual meeting 2014 and the Econometric Society 2015 Winter Meetings, the Universidad de los Andes seminar, International Food Policy Research Institute and the PEGNet 2015 Conference. We are also grateful to Theresa Betancourt for providing valuable input on the psychometric survey measures. For outstanding research assistance in the field, we thank Anthony Mansaray, Ali Ahmed, Jessica Creighton and Nadia Hasham (in chronological order of involvement). This project was funded by 3ie and the J-PAL Governance Initiative, and received extensive support from Innovations for Poverty Action. We thank them, without implicating them, for making this project possible. All errors and omissions are our own.

Summary

Wars destroy more than physical capital. When neighbours fight neighbours, this severs social ties, undermining social capital within the community. To recover from wars and rebuild social capital, many countries undertake truth and reconciliation efforts. We examine the consequences of one such programme in post-conflict Sierra Leone, which was designed and implemented by a non-governmental organisation called Fambul Tok. The programme sets up forums in sections, or groups of 10 villages, where victims detail war atrocities and perpetrators confess to war crimes. We use random assignment to study its impact across 100 sections, surveying 2,200 individuals in these areas. The short-run effects are measured nine months after the forums are held, as part of the intervention. In a subsample of sections, we are also able to gauge longer-run effects, 31 months after the forums take place.

We find that the reconciliation programme had both positive and negative consequences. On the one hand, the programme led to greater forgiveness of those who perpetrated violence during the civil war. Furthermore, respondents were found to be more trusting of ex-combatants, and social networks became stronger as people sought more help and advice from each other. Individuals residing in treated villages also became more community oriented in their behaviour: they joined more organisations like parent–teacher associations and contributed more to public goods.

On the other hand, the reconciliation process *undermined* psychological well-being and left lasting psychological damage. In particular, individuals in treated communities scored worse on three psychological measures: anxiety, depression and post-traumatic stress disorder. Moreover, all of these effects, both psychological and societal, persisted for nearly three years after the intervention. These results suggest that confronting past war experiences may prove traumatic by reopening old wounds.

Taken together, our findings show that reconciliation can promote forgiveness and bring community members together, while also reducing psychological well-being. Our findings suggest that reconciliation processes should be redesigned in ways that reduce their psychological costs, while retaining their positive societal benefits.

Contents

List of figure and tables

Abbreviations and acronyms

ANCOVA	Analysis of covariance
CDD/R	Community-driven development or reconstruction.
CDF	Civil defence force
CPTSD-RI	Child Posttraumatic Stress Disorder Reaction Index
DSM-IV	Fourth Diagnostic and Statistical Manual of Mental Disorders
FDR	False discovery rate
FT	Fambul Tok
NGO	Non-governmental organisation
PAP	Pre-analysis plan
PTA	Parent–teacher association
PTSD	Post-traumatic stress disorder
RUF	Revolutionary United Front

1. Introduction

Recovering from conflict is a formidable challenge. Wars destroy more than physical infrastructure, severing social ties between individuals. As a consequence, wars leave behind not just physical damage, but also fractured social networks, weaker provision of public goods and reduced trust among community members. Recovering from wars, therefore, also means rebuilding social capital.

Many posit that the key to renewing societies in the aftermath of war lies in truth and reconciliation efforts, which aim to air wartime grievances by bringing victims face to face with the perpetrators of war. As part of these efforts, victims describe their experiences, while perpetrators admit to crimes, seeking forgiveness for their actions. The stated objectives of these programmes are ambitious: talking about the past is held to promote personal healing, improve interpersonal relations, promote economic development and generate lasting peace. Moreover, this approach is very common: in fact, almost every country emerging from conflict in the last two decades has used some type of reconciliation process.[1] Yet, there is little rigorous evidence surrounding their impact. In a review of the evidence, Mendeloff (2004) concludes that these claims 'rest far more on faith than on sound logic or empirical evidence'.

Our study seeks to fill this gap. We experimentally evaluate a reconciliation programme in Sierra Leone, which experienced a brutal civil war during 1991–2002. The programme was designed and implemented by a non-governmental organisation (NGO) called Fambul Tok (meaning Family Talk in Krio). When Fambul Tok (FT) began expanding its programme to new areas of the country in 2011, we used random assignment to evaluate the impact of its work. Our evaluation is completely independent, and we provided no input to any aspect of its programme.

As part of the programme, forums are held, during which war survivors detail their experiences, and perpetrators admit to their crimes and seek forgiveness for their actions. Importantly, no one is prosecuted or otherwise punished for participating. Typically, these forums are held in sections, which are groups of up to 10 villages. In addition, after the forums are held, FT also installs additional institutional structures in treated villages: a Peace Tree serves as a focal point for promoting dispute resolution; a Peace Mothers group facilitates discussion of gender-based violence and promotes female economic activities; and in some places land is set aside for communal farms as a pledge to work towards reconciliation.

The programme was administered in waves, enabling the NGO to work within its capacity for the purposes of the evaluation. The first wave included 40 sections and the second wave had 60 sections. Of these 100 sections, 50 were assigned to treatment and 50 were assigned to control; 2,200 individuals were surveyed. The endline surveys took place at least nine months after the forums were held. In the first wave, the respondents were also resurveyed 31 months after the intervention. This timeframe allows us to examine longer-run effects than is typically feasible in evaluations of this type.

[1] Some examples include: Chad, Colombia, Congo, El Salvador, Fiji, Ghana, Guatemala, Kenya, Liberia, Morocco, Nigeria, Peru, Sierra Leone, Solomon Islands, South Africa, South Korea, Sri Lanka, Sudan, East Timor and Uganda.

We chose to evaluate the intervention on a broad set of dimensions related to societal and individual healing. Thus, we ask questions related to forgiveness and psychological well-being, attitudes towards ex-combatants and women, trust, the strength of social networks, conflict incidence and resolution, economic activity, public goods provision and household socioeconomic welfare. Given the large set of hypotheses and outcomes of interest, we pre-registered our evaluation design to tie our hands and resist the follies of 'fishing' (Casey et al. 2012; Humphreys et al. 2013). Our pre-analysis plan is registered with the Evidence in Governance and Politics design registry.

Our analysis shows that the reconciliation process led to greater forgiveness of those who perpetrated violence during the civil war. Furthermore, respondents are found to be more trusting of ex-combatants, and social networks became stronger as people sought more help and advice from each other. Individuals residing in treated villages also became more community oriented in their behaviour: they joined more organisations such as parent–teacher associations (PTAs), and contributed more to public goods. We also find a significant improvement in gender-related attitudes, including views on women's rights to express their opinions. This is likely to reflect the success of women's groups within treated communities. However, we find no conclusive evidence that economic activity, economic outcomes, conflict incidence or conflict resolution changed due to the programme.

Our analysis also shows that the reconciliation process *undermined* psychological well-being along several dimensions. In particular, individuals in treated communities score worse on three psychological measures: anxiety, depression and post-traumatic stress disorder (PTSD). Moreover, all these effects, both psychological and societal, persisted over the longer term, 31 months after the intervention. These results suggest that confronting war experiences may reopen old wounds, and are in line with psychology studies reporting risk of *retraumatisation* from one-time debriefing (van Emmerik et al. 2002; Rose et al. 2002). Our results are also consistent with non-experimental evidence that national-level participation in the Truth and Reconciliation Commission did not improve psychological health (Kaminer et al. 2001) and that participation in the *gacaca* courts in Rwanda may have worsened it (Brounéus 2008; 2010).

Taken together, our findings show that reconciliation can promote forgiveness and bring community members together, while also reducing psychological well-being. Thus, the societal benefits came at substantial psychological cost.

These findings have direct implications for the design of reconciliation programmes. Policymakers need to consider the potential psychological costs when deciding whether to implement a reconciliation programme. Furthermore, the findings suggest that in the long run these efforts need to be redesigned in a way that minimises their psychological cost, while retaining their societal benefit. For example, combining these processes with ongoing therapeutic counselling may be one potential way forward. This is an important avenue for future research.

2. Conflict and reconciliation in Sierra Leone

Sierra Leone had a civil war between 1991 and 2002. The war had no overt ethnic or religious dimensions (Humphreys and Weinstein 2006). Rather, discontent over corruption and authoritarian rule laid the groundwork for rebellion. Kleptocratic rulers enriched themselves with illicit diamond mining throughout the 1970s and 1980s, but few public services were provided over this time (Reno 1995). A one-party state was declared in 1978, and this persisted into the 1990s.

The rebellion was launched in 1991 by rebels from the Revolutionary United Front (RUF). Besides discontent over government ineffectiveness, control of the country's diamond wealth also played an important motivating role (Keen 2005; Bellows and Miguel 2009). Diamonds financed – and thereby prolonged – the conflict.

The war was brutal. More than 50,000 people were killed and over half the population was displaced. Thousands were also raped or had limbs amputated (Human Rights Watch 1999). The vast majority of atrocities were committed by the RUF (Conibere et al. 2004; Smith et al. 2004). However, the Sierra Leonean Army gained notoriety for colluding with the rebels – sometimes to share diamond profits, at other times to avoid direct battles. In the process, they also terrorised civilians.

Local militias, called civil defence forces (CDF), also emerged during the conflict. Towards the end of the war, some CDF factions may also have targeted civilians, but generally they were revered for defending the local population.

The violence was largely neighbour on neighbour, committed by people from the same community who knew one another (Keen 2005). Although the RUF was fighting to overthrow the state, the violence was often personal and motivated by grievances over the local abuse of power, in what is considered to be a strongly gerontocratic society. Chiefs – typically older men – held considerable power over resource allocation, with direct consequences for development outcomes (Acemoglu, Reed and Robinson 2014). Disenfranchised youths, who gained access to guns during the war, would often target chiefs or elders from their home village (Keen 2005).

The nature of this violence underscored the need for reconciliation when the war ended. Following the conflict, the Sierra Leonean government and international community created a special court to try the most notorious, high-profile perpetrators. It indicted 13 such individuals over the next decade. The special court was an expensive endeavour, costing US$125 million in its first four years (Cassese 2006); but the government also set up a national-level Truth and Reconciliation Commission, which heard over 7,000 statements. However, this covered only a small fraction of all atrocities committed. Overstretched, the Truth and Reconciliation Commission did not include broad-based participation by the rural population, and therefore did not reach many of the individuals most adversely affected by the war. This created a gap in the reconciliation process that FT was created to fill. Its aim was to reach out to the communities whose wartime grievances remained unaddressed.

3. Fambul Tok's reconciliation programme

FT was started by a Sierra Leonean in 2007 and ultimately came to operate in 5 of Sierra Leone's 13 districts: Bombali, Kailahun, Koinadugu, Kono and Moyamba. The NGO designed its reconciliation programme with the aim of airing wartime grievances and unifying the community.

Its intervention occurs at the level of sections, which are clusters of up to 10 contiguous villages. The NGO places great emphasis on the process being community driven. At point of entry, FT holds a consultation meeting with all village chiefs in a section to obtain consent and support for the project. Two groups are established: a reconciliation committee consisting of village chiefs, religious and youth leaders, as well as some war survivors and former combatants; and an outreach committee, consisting mostly of youths. The reconciliation committee is trained by FT staff in trauma healing and mediation, and is tasked with reaching out to victims and perpetrators to participate in the truth-telling process. The content of the training varies according to the circumstances of the community, and responds to the needs identified by FT staff as they work in the communities. The initial training was designed and facilitated by international NGO the Center for Justice and Peacebuilding.

The outreach committee helps in publicising and planning, to ensure broad-based participation from all villages in the section. The publicising is done through existing structures in the community, such as community meetings, and is most often communicated via the village chief.

After several months of organisation by these two committees, the reconciliation forum is held. This consists of a two-day bonfire ceremony, at which victims share their experiences and perpetrators ask for forgiveness for their war crimes. Participants from all the neighbouring villages congregate around the bonfire. The ceremony is preceded by a speech from the section chief and/or some religious leaders. Individuals are then invited to come forward in turn and share how they were hurt by someone during the war. The perpetrator, or a member of the perpetrator's family, is then invited to come forward and speak. He or she sometimes provides some justification, and generally also asks for forgiveness. The victim is then granted an opportunity to forgive the perpetrator. This process is punctuated by occasional singing and dancing.

It is capped by a 'cleansing' ceremony, where perpetrators who have expressed remorse are cleansed of their sins. These are traditional ceremonies conducted by Sierra Leone's 'secret societies' and are often accompanied by the pouring of libations.

The bonfire ceremonies, in their entirety, are relatively inexpensive, costing US$150–200.

After the ceremony, FT also establishes a series of local institutional structures to further heal the community. It sets up a Peace Tree, as a symbolic gesture, which provides a focal point for resolving disputes. Conflict mediation using this Peace Tree relies on elders and chiefs, who are traditionally responsible for exercising justice. In some treated areas, it also creates communal farms on land set aside as a pledge towards reconciliation. Finally, it helps establish a Peace Mothers group to promote women's economic activities and discuss gender-targeted atrocities perpetrated during the war.

Some examples of Peace Mothers activities include fish farming and interactive workshops to discuss their wartime experiences.

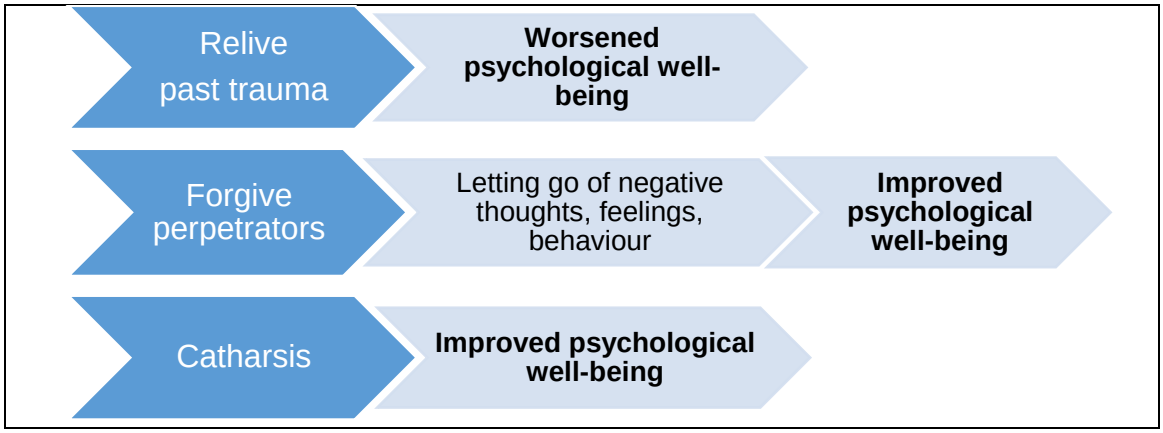
This intervention may have impacts other than reconciliation. For example, the farms may generate economic activity and the Peace Mothers groups can have a galvanising effect on female participation in the community. We will argue below that most of the effects observed are probably due to reconciliation, rather than to these other impacts.

4. Theory of change

Reconciliation, the restoration of fractured relationships, can lead to changed attitudes and behaviour at personal, interpersonal and group levels.

First, on a personal level, the reconciliation process may influence individual levels of forgiveness and psychological well-being, as survivors talk about their past experiences and perpetrators are given the opportunity to express remorse and request forgiveness. On the one hand, reconciliation may improve psychological well-being, if talking about the past is cathartic (Huyse 2003; Cobban 2002; Hamber 2003); or else the act of forgiving may itself boost psychological well-being (Esterling et al. 1999; Smyth et al. 1998; West et al. 2001; Excline et al. 2000; Reed 2004). On the other hand, the reconciliation process may undermine psychological well-being, if victims find it traumatic to talk about the past, if perpetrators feel shame and guilt about admitting their crimes and if third parties learn of new atrocities or experience vicarious retraumatisation (Hayner 2001). The graph below indicates the possible positive and negative relationships between confronting past traumatic events and psychological well-being.

Figure 1: Positive and negative impacts of confronting past traumatic events on psychological health



Second, on an interpersonal level, the reconciliation process may alter community members' attitudes towards ex-combatants. For example, if ex-combatants express remorse, others may distinguish between a bad person and a bad act.[2] As a result, trust towards ex-combatants could increase. This, in turn, could help to reintegrate perpetrators into the community, as more community members are willing to interact with those individuals.

[2] Rigby (2006), for example, quotes a human rights worker who 'is able to justify forgiveness by reference to the fact that they were not fully responsible – "they were drugged"'.

Third, the reconciliation process may restore social ties and boost social capital more broadly. For example, people may be more willing to interact with one another after the community has acknowledged their suffering, and after individuals have forgiven each other. In this regard, reconciliation may strengthen social networks. In addition, these very factors – acknowledgement and forgiveness – may also improve individuals' outlook towards their community and thereby boost their willingness to participate in community organisations and to contribute to public goods.

Finally, altered attitudes to the network structure can lead to further changes in interpersonal relations, evident in increased reciprocal economic activity and relationships involving less conflict (Fehr et al. 1997).

5. Research hypotheses and pre-analysis plan

Given the novelty of the experiment and the wide range of potential impacts of the programme, our pre-registered pre-analysis plan (PAP)[3] included a broad range of hypotheses: forgiveness and psychological well-being, war-related beliefs and attitudes, incidences of conflict and conflict-resolution mechanisms, social capital indicators (such as network structures or public goods provision) and economic indicators.

The hypotheses are:

- **Programme implementation.** The programme was implemented according to stated objectives.
- **Individual attitudes.** FT affects: (1) levels of forgiveness among respondents who were hurt during the war; (2) attitudes towards ex-combatants; (3) attitudes towards war; and (4) gender attitudes.
- **Social capital.** FT affects: (1) level of trust among community members; (2) social divisions and the inclusion of groups that have traditionally been marginalised; (3) strength of networks within the community; (4) strength of group membership; and (5) contribution to the public good.
- **Psychological health:** FT affects individual psychological well-being.
- **Conflict and conflict resolution.** FT affects: (1) incidence of conflict and crime within the community and between communities; and (2) satisfaction and nature of conflict resolution.
- **Economic activity and outcomes.** FT affects: (1) economic activity and cooperation; and (2) economic welfare.

For each hypothesis, we listed the specific indicators to be analysed separately and as a group, using mean effects. For the mean effects analysis, we constructed an aggregate index of the indicators within each hypothesis, as in Anderson (2008) and Kling and others (2007). We also adjusted the standard errors to multiple comparisons by controlling for the false discovery rate (FDR) using the method detailed in Benjamini and Hochberg (1995). These approaches help to mitigate problems of over-rejection of the

[3] Our PAP can be found here: http://bit.ly/1AIG5b0. The plan was registered on 12 December 2012. As in Casey and others (2012), we finalised the plan while the endline data were being entered and cleaned, and before any of the authors had access to the data. In particular, we asked our field staff to password-protect the data and give us access to the password only after the analysis plan was posted.

null hypothesis due to multiple inferences. In the data section below, we discuss in detail the constituent indicators for each hypothesis.

Owing to differences in data collection across the two waves, we developed a modified PAP for the second wave. But in pooling data across the two waves for the results in this paper, we followed the aggregation specified in the plan for the first wave, which was developed before we had access to any endline data from either wave.

5.1 Changes to the pre-analysis plan

We were forced to make some modifications to our analysis ex post due to challenges in aggregating conditional and unconditional outcomes, and slight changes in the design of the social network between waves.

Our PAP specified how particular indicators would be aggregated in our examination of various hypotheses. As in Casey and others (2012) and Humphreys and others (2012), some of our original hypotheses combined 'conditional' outcomes that relate to a subsample of respondents with 'unconditional' outcomes that relate to the full sample. However, this could create a bias in the Anderson (2008) index, due to sample selection, and may induce false rejection of the null in the index of Kling and others (2007), if the two sample sizes differ substantially. For example, in the measurement of trust, all respondents stated how much they trusted people in general, but only respondents who knew ex-combatants personally stated how much they trusted them – and this latter group of respondents made up less than half the sample.

The Anderson (2008) approach would create a composite trust index restricted to this latter half. But the selected sample of individuals who knew ex-combatants could have various levels of generalised trust, and their trust levels may also respond differently to treatment. And so this approach could produce a biased treatment effect that is not representative of the full sample. The approach of Kling and others (2007) would create a composite trust index by imputing values to half the sample, for whom one of three major indicators would be missing. Imputation at this scale could artificially reduce the standard errors, because the sample size increases without increasing true underlying variation.

We therefore made changes to three of our hypotheses that contained both 'conditional' and 'unconditional' outcomes. We focused our mean effect index of trust on four community-wide measures that were administered to all respondents. We examine separately trust of ex-combatants and of migrants – questions that were asked of subsamples that knew these subgroups.

The same issues apply to the forgiveness hypothesis. The original index attempted to aggregate across three subgroups: those that had experienced hurt in the past; those who personally knew the perpetrator; and those whose persecutor still resided in the village. But the sample sizes of the final two subgroups, constituting affirmative responses from 12 randomly sampled respondents, are too small for any useful analysis. We therefore look separately at the first group, covering forgiveness by all those who had experienced hurt in the past.

Similar aggregation issues applied to our analysis of conflict resolution. Our original PAP proposed to examine the change in the proportion of conflicts that were resolved, as well as the proportion of conflicts that were resolved without a third party and the proportion that were resolved satisfactorily. Of course, the latter two variables are conditional on the former. In addition, a separate grouping proposed to look at the proportion resolved by chiefs and the fines levied by the chiefs; but fines apply to a much broader set of cases than just conflicts resolved by chiefs.

Moreover, one of our indicators – resolution without a third party – displayed zero variation in the baseline sample, which makes it infeasible to create a mean index using this measure. Given this limitation, we simply show each of the indicators in the table. Since there was a reported increase in the proportion of conflicts resolved by chiefs, we thought it would be informative to see whether there was any decrease in resolution in the other categories. And so, in addition, we examined resolution via mediation with friends and family. These results are discussed further below.

Finally, we had to make some changes to the social network questions. The way the questions were asked in the baseline survey proved problematic, because some answers displayed limited variation. In the wave-one baseline, the questions were prompted too strongly: each of the 12 respondents was asked to consider, in turn, each of the 11 other individuals being surveyed (for example: ‘Now consider John Koroma. Would you be willing to share a farm boundary with John Koroma?’). Consequently, most of the respondents were mentioned by everyone else; but this was not a meaningful measure of actual connectedness. By contrast, in the second-wave baseline, the questions were not prompted at all (for example: ‘Name all the people you would be willing to share a farm boundary with’). As a result, respondents were almost never mentioned by anyone else, which again did not serve as a meaningful measure.

Given these problems, in the endline surveys we undertook several additional changes. First, we dropped the farm-boundary question entirely, since this showed the least variation. In one part of the survey, the questions were unprompted (for example: ‘Say you have a problem. Think of all the people from this village who you would go to for advice and help. Who are these people?’). The unprompted questions were used to construct how many times the respondent lists someone else. In another part of the survey, the questions were lightly prompted, with the respondent being asked to consider jointly the 11 other respondents being surveyed (for example: ‘Of the people I named, name the people (if any) who you would go to for advice or help’). The lightly prompted questions were used to construct an indicator for how often the respondent was listed by someone else. Furthermore, since the social network questions were very time intensive to administer, in the prompted format we no longer asked ‘Who would you ask to go and collect money for you?’ As a result of these changes, we had to conduct cross-sectional analysis on a subset of the social network questions that were administered in the endline surveys.

Table 1 provides a summary of all the hypotheses.

Table 1: Summary of hypotheses in pre-analysis plan

No.	Hypothesis	Change?	Reason for change
1	Programme implemented according to objectives		
2	Forgiveness	Y	Original aggregated conditional outcome
3	Psychological well-being		
4	Attitudes towards ex-combatants		
5	War attitudes		
6	Attitudes towards gender		
7	Incidence of conflict and crime		
8	Conflict resolution	Y	Original aggregated conditional outcome
9	Entrenchment of traditional sources of power	Y	Original aggregated conditional outcome/no control group variation in one indicator
10	Trust	Y	Original aggregated conditional outcome
11	Social divisions	Y	Change to social networks question
12	Social networks	Y	Change to social networks question
13	Group membership		
14	Public goods provision		
15	Economic activity		
16	Economic outcomes		

Notes: The first two columns list the number and hypothesis, as listed in the PAP. The third column indicates if there was a change in how the hypothesis was examined relative to what was pre-specified. The fourth column provides a brief description of the reason for the change, each of which is detailed in the report.

6. Evaluation

6.1 Evaluation design

FT planned to expand its operations to new sections of Sierra Leone, starting in 2011. Given this, we used random assignment to evaluate the impact of the intervention. Importantly, our evaluation was completely independent of the intervention. We provided no input into the design of its pre-existing intervention. Also, there was no overlap between the FT staff who implemented the programme and our enumerators who collected the data.

As a first step in the evaluation, FT identified sections that were willing to participate in the programme in its five districts of operation. These sections were identified prior to the start of the community consultation process, or indeed any other part of the programme.

The evaluation then proceeded in waves, enabling FT to work within its capacity. Forty communities comprised the first wave of the evaluation and 60 the second wave. The programme was also implemented in a third wave. However, data collection to evaluate this wave was interrupted by the Ebola crisis in Sierra Leone. Our field staff had to be evacuated while we were in the midst of collecting behavioural measures.

Within each wave, we first surveyed the communities at baseline. We matched sections into pairs, stratified by district based on an `optimal greedy' algorithm (Greevy et al. 2004; Imai et al. 2009), using baseline data on exposure to violence, conflict incidence, economic activity and psychological health. We then randomly assigned one section in each pair to treatment. Random assignment was based on a public lottery. Our research team travelled to each district. Together with FT staff members, in front of the eligible section chiefs, they selected the treatment section from each pair.

We conducted household and village surveys in two villages within each section. One village was the section headquarters, where the ceremony was typically held, and the second a randomly chosen village within each section. Respondents within the villages were selected using two stages of an in-field random sampling process. First, we randomly sampled 12 households in each village. Then we randomly sampled one individual aged 18–65 in each household. A village-level survey was also conducted at baseline and endline. Due to mechanical error in the hand-held devices used for data collection, baseline data are missing for six villages, and endline data are missing for an additional six villages. If a village-level variable is one of just a few key indicators in an index, we construct the index by dropping those villages, rather than imputing missing values to them.

In both waves, we sought to resurvey the same respondents who were interviewed at baseline. We went to great lengths to minimise attrition. We conducted several follow-up visits and also tracked down respondents who had moved to neighbouring villages. We developed two measures of attrition. The first equals one if a baseline respondent did not receive an endline survey at all. This attrition rate is 7 per cent. The second applies only to the first round and equals one if a baseline respondent was not available for both endline surveys. This conservative attrition rate is 13 per cent. As is shown in Table 2,

neither attrition indicator is predicted by treatment, meaning that attrition was not differential in treatment communities relative to control communities.

Table 2: Attrition

	(1) Attrition-Round 1	(2) Attrition-Round 2	(3) Attrition	(4) Overall Attrition-Broad	(5) Overall Attrition-Narrow
Treatment	0.017	0.022	-0.007	0.010	-0.003
	(0.011)	(0.017)	(0.014)	(0.012)	(0.009)
Sample	Wave 1	Wave 1	Wave 2	Waves 1 and 2	Waves 1 and 2
Attrition rate	7.04%	10.71%	10.98%	13.22%	7.05%
Obs.	952	952	1,430	2,382	2,382
R-sqr.	0.021	0.049	0.036	0.043	0.070

Notes: Each column represents a separate regression of an attrition measure on treatment assignment. In the first (second) column, attrition equals one if respondent was not resurveyed in the first (second) endline round in Wave 1. Overall Attrition-Broad equals one if the respondent is missing from either the endline round in Wave 1 or the endline in Wave 2. Overall Attrition-Narrow equals one if the respondent is missing in both endline rounds in Wave 1 and the endline in Wave 2. Variables not shown include section-pair fixed effects. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

Two of our project associates oversaw the data collection. Data were collected electronically by enumerators on tablets and gathered by the project associate in the process of data collection. We were thus able to verify data quality while data collection was taking place. As an additional step of quality control and accountability, supervisors, who were part of the survey teams, went back to a random subset of villages to resurvey a subset of individuals. The enumerators were made aware of this prior to data collection.

The evaluation team secured ethical approval to conduct the evaluation from the institutional review boards of both New York University and the University of Oxford, as well as from IPA's review board

Our initial power calculations were based on the National Public Services datasets – one of the few datasets in Sierra Leone that also asks questions related to conflict and public goods provision. Given the wide range of innovative outcomes we wish to measure and the lack of previous data on these outcomes, we were conservative with our initial power calculations. In our sample of 100 communities, we were sufficiently powered to detect an impact of 0.14 standard deviations, if the correlation between baseline and endline was 0.4 and the intra-cluster correlation is 0.05. The size of our impact ranges between 0.1 and 0.2, and the intra-cluster correlation ranges between 0.05 and 0.15.

6.2 Timeline

The interventions in Wave 1 villages began with bonfire ceremonies that took place between late April and June 2011. In these sections, we were able to collect two rounds of endline data – the first round of the endline survey was administered approximately 9 months after the intervention, and the second round was administered 31 months

afterwards. In Wave 2 villages, bonfire ceremonies occurred in March through June 2012. We administered one endline round for these sections, approximately 19 months after the intervention. (There was a longer delay in the second wave, to ensure that our data collection did not coincide with elections.) We thus present our results using these three sets of endline surveys.

Table 3: Impact evaluation timeline

Year	Month	WAVE 1	WAVE 2
2011	Jan		
	Feb		
	March	**Baseline**	
	April	**Intervention**	
	May		
	June		
	July		
	Aug		
	Sept		**Baseline**
	Oct		
	Nov		
	Dec		
2012	Jan		
	Feb		
	March		**Intervention**
	April	**Endline-Round 1**	
	May		
	June		
	July		
	Aug		
	Sept		
	Oct		
	Nov		
	Dec		
2014	Jan		
	Feb	**Endline-Round 2**	**Endline**
	March		

6.3 Empirical strategy

Our main specifications combine all three sets of endline surveys – two rounds for the first wave and one round for the second wave. Our main results therefore look at the average impact across all time periods. Where possible, we also control for the baseline value of the dependent variable. We opt for this estimator (otherwise known as the analysis of covariance (ANCOVA) specification), as it accounts for the covariance between pre- and post-treatment outcomes (Frison and Pocock 1992; McKenzie 2012), and has more power than a difference-in-differences estimator.

We allow the baseline dependent variable to exert different effects across rounds and waves. The estimating equation can be represented as:

$$y_{risvspw} = \beta_0 + \beta_1 T_s + \rho_P + \beta_2 y_{0isvspw} + \delta_r + \delta_r y_{0isvspw} + \lambda_w y_{0isvspw} + \varepsilon_{risvspw}$$

where $y_{0isvspw}$ and $y_{risvspw}$ denote outcomes at baseline and endline round *r* respectively, for individual *i* in village *v*, section *s*, section-pair *p* and wave *w*. ρ_P denotes section-pair fixed effects, which account for section-level matching in the allocation of treatment (Bruhn and McKenzie 2009). T_s is assignment to treatment, and β_1 measures the treatment effect. δ_r is a round effect which equals one for the second-round endline. $\delta_r y_{0isvspw}$ is the interaction of this second-round indicator with the baseline outcome, and allows the baseline to exert different effects over time. $\lambda_w\}$ denotes a wave effect, which equals one for sections in the second wave. Since each wave includes different sections, these wave effects are subsumed by section-pair effects. $\lambda_w y_{0isvspw}$ denotes the interaction of the wave effect with the baseline outcome. This allows baseline variables to have different effects for the Wave 2 sections. This is particularly important, since we are only able to include the pared-down baseline outcomes collected in the second-wave baseline survey (discussed in more detail below). For our main results, there are no other controls beyond what is specified in the estimating equation.

For outcomes where we do not have any baseline values, we utilise a simple cross-sectional specification:

$$y_{risvspw} = \beta_0 + \beta_1 T_s + \rho_P + \delta_r + \varepsilon_{risvspw}$$

In all specifications, we cluster the standard errors at the section level, the unit of treatment allocation. There are one section in Wave 1 and two sections in Wave 2 that do not match the treatment assignment: these sections were assigned to control, and yet six of the respondents in one of the villages and eight of the respondents in the two others reported attending a bonfire ceremony. However, we utilise assignment to treatment in all of our specifications. Thus, ceremony participation among control respondents may lead to an understatement of the effect.

We use two types of mean indices, following Anderson (2008) and Kling and others (2007). Both indices standardise the indicators and sum across these measures. The Anderson (2008) approach weights the standardised outcomes by the inverse of the variance–covariance matrix. This places less weight on indicators that add no extra information, due to high variance or high correlation with other indicators. The approach of Kling and others (2007) accounts for missing values by imputing the mean of the control (treatment) group to missing values for the control (treatment) group. Given missing data in some of our key indicators, the loss of observations by aggregating across indicators without imputation is at times substantial. Thus, we use mean indices constructed using the approach of Kling and others (2007) for the main results, but also present all results using the approach of Anderson (2008) for robustness checks.

Our empirical strategy is thus the intent to treat, and identifies the total treatment effect that stems from both the direct and the indirect effects of the programme – both for those who participated and for those who did not participate in the reconciliation ceremony. For example, direct effects will arise from those who participated directly in the bonfire ceremony, as 45 per cent of our randomly sampled respondents reported doing. However, there may also be indirect effects arising from those who did not participate directly. For example, a household member may attend the bonfire ceremony and develop a more positive outlook on his or her community, and subsequently convince other household members to join community groups. These spillovers underscore why

the section is the stable unit of treatment, and why it is important to randomise this intervention at the community level.[4]

6.4 Data

We administered a rich set of survey questions. This section provides an overview of the measures used in our primary analysis: forgiveness, psychological well-being and social capital. The online appendix details measures used for additional results.

We constructed a **forgiveness** scale from the Rye Subscale of the Enright Forgiveness Inventory (Subkoviak et al. 1995), which consists of 12 questions on a four-point Likert scale. This scale captures three dimensions of forgiveness (Enright and The Human Development Study Group 1991): (1) affect – including feelings against the perpetrator, such as resentment, anger, compassion and love; (2) judgement – opinions and beliefs about the perpetrator; and (3) behaviour – actions or expressed desired actions against perpetrators, such as revenge or acts of kindness.[5]

These questions were administered to respondents who reported that they had been physically or emotionally hurt during the war. While all endline surveys included these 12 questions, the second-wave baseline included a subset of seven questions. This subset is used to form a pared-down index, which serves as a baseline control for second-wave observations.

For **psychological well-being**, we aggregated three indices for PTSD, anxiety and depression. The 11 questions we used to construct the PTSD index were drawn from the 17-item PTSD Symptom Scale (Foa et al. 1993; Foa and Tolin 2000), which assesses the presence and severity of PTSD symptoms according to the Fourth Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). These are the same questions as the PTSD Checklist, Civilian Version, as developed by the National Center for PTSD (Weathers et al. 1991). These questions focus on three categories of symptoms: re-experience, arousal and avoidance. The responses to these questions are aggregate on a four-point Likert scale.

The second-wave baseline collected a more pared-down set of indicators for the forgiveness, depression and anxiety indices (7, 5 and 6 indicators, respectively). The ANCOVA specification still allows us to control for the pared-down measures at baseline (with no introduction in bias, but a reduction in statistical power). Nonetheless, for the sake of inclusion, we also report analysis where the endline indices are restricted to the set of indicators also found in the baseline.

[4] They also present a challenge to instrumenting ceremony attendance with treatment assignment, since the spillovers point to a potential violation of the exclusion restriction using this approach.

[5] Respondents are asked if they strongly agree, agree, disagree or strongly disagree with questions such as: ‘Do you spend time thinking about ways to get revenge on the person who wronged you?’ To give this scale an intuitive meaning, we code agreement in the negative domain, and disagreement on the positive domain symmetrically around 0. We follow the literature in coding across responses in steps of one, for resultant values ranging from -1.5 to 1.5. Since linear regression is unaffected by affine transformations, coding the variable on other values such as 1 to 4 does not affect the results.

Previous validation exercises for research purposes have demonstrated that both the PTSD Symptom Scale and the PTSD Checklist show high internal consistency and high test–retest reliability, concurrent validity and high correlation with other diagnostic scales (Foa 2006; Ventureyra et al. 2002; Foa et al. 1997). See Norris and Hamblen (2004) for a detailed review.

Our 7 depression and 10 anxiety questions are drawn from the Zung Depression and Zung Anxiety indices (Zung 1971), respectively. These indices have been found to have high internal consistency (Knight et al. 1983). The responses to these questions are also aggregate on a four-point Likert scale.

In administering these psychometric indicators in Sierra Leone, we face the perennial question of whether scales created in the developed world are culturally relevant for such a context. To adapt these questions to the Sierra Leonean context, we piloted our surveys extensively and adjusted the wording of the scales in order to reflect the Krio language better. For example, one PTSD question is: 'Have you had recurrent or intrusive distressing thoughts or recollections about the assault?' We adapted this question so that it read: 'In the last month, did you sit and think of bad bad things that happened to you even though you don't want to think of it?'

Moreover, the psychological questions we administered also correspond closely to those administered in other recent studies in post-conflict countries within Sub-Saharan Africa. The Child Posttraumatic Stress Disorder Reaction Index (CPTSD-RI) (Steinberg et al. 2004; Hawkins and Radcliffe 2006), for example, includes the same 17-item scale that measures severity of PTSD symptoms according to DSM-IV. A validation study of CPTSD-RI sexual abuse survivors in Zambia shows high reliability and concurrent validity (Murray et al. 2011); and it has also been used in a sample of Ugandan and Congolese child soldiers (Bayer et al. 2007). As another example, the Survey of War-Affected Youth conducted in Uganda (Blattman and Annan 2010) also administered 9 out of our 11 PTSD questions and 6 out of the 7 anxiety questions. Furthermore, 15 of the 17 depression and anxiety questions correspond to the Johns Hopkins 25-item checklist for anxiety and depression (60). Studies in a sample of former child soldiers in Sierra Leone (Betancourt et al. 2011) and HIV-positive pregnant women in Tanzania (Kaaya et al. 2002) have demonstrated good psychometric properties of this scale, with a Cronbach's alpha of 0.88 and 0.93 in these respective study sites.

Moreover, all three of our psychological indices also show strong internal consistency. Our PTSD, anxiety and depression indices have a Cronbach's alpha of 0.936, 0.901 and 0.831, respectively. When looking at only the pared-down questions, the anxiety and depression indices have a Cronbach's alpha of 0.851 and 0.730, respectively. We find it encouraging that even the pared-down indices are strongly internally consistent (though this reduction in indicators only affects *baseline* measures in the second wave).

We also measured **attitudes towards ex-combatants and war participation**. We gauged beliefs on the culpability of ex-combatants by asking to what extent respondents agree with the statements: 'Those that did bad things in the past would do it again if they had the chance' and 'People who joined the RUF are not responsible for what they did, since they were drugged.' For war attitudes, we asked if the respondent would fight again, or believes that others would fight again.

We additionally measured **gender attitudes** by asking whether the respondent agreed with seven circumstances under which it is acceptable for 'a man to beat his wife', and whether a wife has a right to express her own opinions.

To gauge impacts on **social networks**, we asked the respondents to list those of the 11 other respondents whom they consider to be very good friends and would ask for advice and help.[6] We used this to construct how many times a respondent was named by someone else. We also got the respondents to list all the people in the village that they would ask to collect money for them and would ask for help. We were only able to conduct cross-sectional analyses with these questions, since they were asked differently in the baseline and endline surveys.

We also collected a battery of social capital questions developed by the World Bank.[7] We separately administered **questions on trust of ex-combatants and migrants**, and four questions on trust of community members (which comprise an index of generalised trust). We examined trust of ex-combatants and migrants separately, since these questions were only administered to individuals who knew someone from these groups. Migrants were of interest, because many former combatants migrated out of their communities after the war, which creates ambiguity as to whether a migrant is in fact a former combatant. Also, migrants are typically considered a marginalised group, and so an increase in trust of them could reflect more inclusive attitudes. The definition of trust was discussed in detail during the training, in order to ensure that the enumerators had a shared understanding of the meaning of the word. We also distinguished between trust towards villagers and trust towards outsiders, since it is plausible that an intervention that promotes stronger ties within the community could succeed at the expense of trust towards those from outside the community.

Our **group membership index** considers both membership and meeting attendance for organisations such as PTAs and religious groups. Our public goods measure includes monetary and labour contributions to public facilities and community groups; the number of community projects; contributing money to a family in need over the past three months; and participation in road brushing, a common form of road maintenance in Sierra Leone.

We tracked the incidence of **conflicts** related to loans, land, property and religion that the respondent had experienced over the past six months. We also recorded the method of **conflict resolution** and the degree of satisfaction with conflict resolution. We asked about fines levied as a form of punishment. In addition, we tracked the incidence of violent and non-violent crimes. Finally, we recorded inter-village disputes over this period, although this is a village-level variable and, as discussed in the paper, we are missing a considerable number of village surveys.

Our measure of **social tensions** comes from the World Bank[8] and includes an indicator of the extent to which divisions between non-marginalised and marginalised groups (migrants versus non-migrants, the young versus the old, and the poor versus the rich)

[6] The enumerator first named the other 11 villagers who were being surveyed, then asked whom the respondent would choose. The enumerators were trained to emphasise that the respondent should not simply name everyone.
[7] http://go.worldbank.org/BOA3AR43W0
[8] http://go.worldbank.org/BOA3AR43W0

escalate into violence; it also covers feelings of inclusion, as measured by the extent to which respondents feel they would benefit from community resources, such as donations, and the extent to which they feel their voice is heard.
We also tracked various types of **economic activity**. Our index includes measures of frequency and size of lending and borrowing; time spent working on other people's farms over the past three months;[9] the number of communal farms; and the number and use of traders in the community.

To measure **economic well-being**, we constructed an index of household assets and housing quality, using principal component analysis. This index includes 16 household goods and whether the roof is made of straw and the walls of mud. We also asked respondents to report their subjective assessments of their ability to meet basic household needs, such as school fees and healthcare, and their perceptions of whether their household situation has improved within the past year. The assets measure, along with the two subjective assessments, comprise an index of economic outcomes used in the analysis.

7. Results

7.1 Programme implementation

Table 4 reports treatment coefficients on metrics of programme implementation. It shows that the programme was well implemented: respondents in the treatment villages were significantly more likely to have heard of Fambul Tok and to have participated in one of its activities. The depth of village participation was large (not shown in table): over 80 per cent of survey respondents in the treatment communities had heard of FT and 72 per cent of respondents knew that a bonfire ceremony had taken place in their community. In 100 per cent of treatment communities, *at least* one person knew that a bonfire ceremony had taken place. This corresponds with what FT reported to us.[10] Participation in the bonfire ceremony was high: 45 per cent of respondents in the treatment communities reported having actually attended the ceremony.

Table 4: Programme implementation

VARIABLES	Coeff.	Std. Error	Obs.	R-sqr.
Heard of Fambul Tok	0.426***	(0.030)	3,003	0.296
Fambul Tok held bonfire	0.689***	(0.057)	3,008	0.576
Fambul Tok communal farm	0.190***	(0.036)	3,008	0.343
Fambul Tok Peace Tree	0.265***	(0.033)	3,008	0.273
Fambul Tok Peace Mothers group	0.406***	(0.046)	3,008	0.381

Notes: Each row represents a separate regression of the outcome shown in the first column on treatment assignment. Variables not shown include section-pair fixed effects and the second-round indicator. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

[9] In a poor agricultural society such as Sierra Leone, sharing of farm labour serves as an important form of economic exchange.
[10] We used two different methods of monitoring: we attended and observed a subset of the bonfire ceremonies; and during our endline household and village surveys we asked about participation in the Fambul Tok activities in the treatment communities.

7.2 Descriptive statistics, baseline balance and attrition

Table 5 presents descriptive statistics of key variables. Individual and village-level characteristics are from the baseline data.

The population we surveyed resided in highly impoverished conditions. Over 70 per cent had no formal education, and less than 8 per cent lived in a village with a market. In addition, these individuals experienced extensive violence during the civil war: 54% had a family member killed; 33% were beaten; 2% reported having been maimed; and 3% acknowledged that they had been raped. This numbers are likely to be an under-estimates given the sensitivity of these outcomes. (Table 8 also shows that there is a negative relationship between the violence exposure variables and our psychological measures at baseline, including affect towards former combatants, as captured by forgiveness, PTSD, anxiety and depression. These correlations suggest that the psychological measures capture meaningful variation.

Table 6 shows balance statistics on main outcome variables at baseline, as well as some key individual and village-level characteristics.[11] Most outcomes display balance, with the exception of trust. Statistically, we expect to observe imbalance in some indicators purely by chance. Moreover, the imbalance goes in different directions for different measures, which suggests that these are statistical aberrations: the index of generalised trust is higher, while trust of migrants is lower in treatment communities. Both are also only significant at the 10 per cent level.

In addition, in our main specification we control for the baseline dependent variable, and this accounts directly for the confounding effects of potential imbalance. Finally, for robustness, we also present specifications where we control for imbalanced baseline indicators in all outcome regressions.

Next, we examine effects on attrition. We develop two measures of attrition: the attrition rate of those who appeared in the baseline but were missing from either endline round in Wave 1 or the endline in Wave 2 is 13 per cent (315 out of 2,382 individuals); while the attrition rate for those missing from both endline rounds in Wave 1 or the endline in Wave 2 is 7 per cent (168 out of 2,382 individuals). Table 2 shows that neither of these attrition measures, nor the attrition measure of each endline round separately, is predicted by treatment.

[11] We present the pared-down measures where we have limited measures in the Wave 2 baseline.

Table 5: Descriptive statistics

VARIABLES	Obs.	Mean	Std. Dev.	Min.	Max.
Individual and village characteristics (baseline):					
Gender	2,212	0.549	0.498	0	1
No formal education	2,208	0.717	0.451	0	1
Occupation farmer	2,345	0.744	0.436	0	1
Market in village	2,075	0.085	0.279	0	1
Village size (number households)	2,135	185.727	331.354	9	2,811
Beaten	2,097	0.329	0.470	0	1
Maimed	2,099	0.020	0.138	0	1
Raped	2,092	0.030	0.170	0	1
Family member killed	2,157	0.535	0.499	0	1
Saw violence	1,749	0.440	0.496	0	1
Panel outcomes (baseline and endline):					
Forgive perpetrators	4,296	-0.217	3.901	-10.5	10.5
Ex-combatants would not fight	5,191	2.552	0.794	1	4
Trust of ex-combatants	3,016	1.823	0.981	1	4
Trust of migrants	4,484	3.074	0.775	1	4
Index of generalised trust	5,212	0.014	0.686	-1.987	1.932
Index of community group participation	5,218	0.011	0.429	-0.573	2.437
Attitude towards wife beati	5,185	10.468	1.930	4	12
Index of economic outcomes	5,222	-0.017	0.555	-1.545	6.350
Index of economic activity	5,222	0.002	0.481	-1.052	11.839
Index of group tensions	5,212	0.004	0.581	-2.778	1.470
Index of psychological well-being	5,205	-0.035	0.839	-5.506	1.907
Less PTSD	5,067	26.769	5.746	0	33
Less anxiety	5,141	13.356	3.929	0	21
Less depression	5,158	10.988	2.380	0	15
Cross-sectional outcomes (endline):					
Forgive perpetrators	2,434	2.502	5.408	-18	18
Index of war attitudes	3,000	-0.011	0.692	-1.675	2.526
If another war, people would not fight	3,000	0.770	0.421	0	1
People would not be a part of another rebellion	3,000	0.838	0.368	0	1
If another war, you would not fight	3,000	0.045	0.207	0	1
Index of network strength	3,008	0.047	0.817	-1.144	27.597
Number of people respondent would approach for advice/help	3,005	2.961	2.193	0.000	47.000
Number of people respondent would ask to collect money	3,005	3.214	5.223	0	244
Number of times respondent listed by others a friend	3,008	2.236	2.023	0	13
Number of times respondent listed by others for advice/help	3,008	3.419	2.887	0	16

Table 6: Baseline balance

VARIABLES	T-C	Std. Error	Obs.
Market in village	0.002	(0.033)	2,075
Village size	16.486	(30.803)	2,158
No formal education	-0.000	(0.015)	2,208
Forgive perpetrators	-0.107	(0.153)	1,862
Ex-combatants would not fight	0.011	(0.034)	2,191
Trust of ex-combatants	-0.022	(0.054)	1,546
Trust of migrants	-0.059*	(0.032)	1,962
Index of generalised trust	0.052*	(0.029)	2,211
Index of community group participation	-0.016	(0.020)	2,213
Index of public goods contributions	-0.034	(0.021)	2,214
Index of psychological well-being	0.023	(0.038)	2,202
Attitude towards wife beating	-0.350	(0.226)	912
Index of conflict and crime (village-level variable)	-0.071	(0.072)	190
Index of economic activity	-0.026	(0.024)	2,214
Index of economic outcomes	0.019	(0.031)	2,214

Notes: Each row represents a separate regression of the baseline variable shown in the first column on treatment assignment. All regressions include section-pair fixed effects. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

7.3 Forgiveness and views of ex-combatants

Table 7 assesses the relationship between reconciliation and forgiveness. The outcome in the first row measures whether those hurt during the war have forgiven the perpetrators, based on the Rye Subscale of the Enright Forgiveness Index. The second row is the equivalent measure for the subset of questions in both baseline waves.

These results show that the reconciliation process increased forgiveness substantially. The coefficient in the second row is 0.277. At endline, the control group mean measure was 0.951. A percentage calculation would imply a 29 per cent effect, but this may not be meaningful, since forgiveness is measured by aggregating questions on a Likert scale.

We instead benchmark the effect by considering how experience of violence during the war influenced an individual's baseline forgiveness. Of course, war exposure was widespread – over 80 per cent of the sample reported having been hurt in some way. To ensure that we benchmark against consequential war experiences, we examine how experience of more extreme forms of violence – getting raped, maimed or having a family member killed – influences baseline forgiveness, relative to other forms of hurt. For example, Table 8 shows that rape reduces the baseline affect towards ex-combatants by 1.21. Comparing this to 0.277 would suggest that the reconciliation process offset negative feelings towards perpetrators by 23 per cent. Having a family member killed reduces this measure by 0.920, which implies an offsetting effect of 30 per cent.

Table 7: Forgiveness, trust and attitudes towards war

VARIABLES	Control mean	Coeff.	Std. Error	Obs.	R-sqr.	Specification
Forgiveness						
Forgive perpetrators	2.264	0.571**	(0.227)	2,010	0.131	ANCOVA
Forgive perpetrators (based on questions in both baselines)	0.951	0.277*	(0.145)	2,085	0.121	ANCOVA
Attitudes related to the war						
Index of attitudes towards ex-combatants	-	-0.007	(0.029)	2,980	0.075	ANCOVA
Indicators:						
Those who did bad things during the war would do it again	2.582	0.018	(0.030)	2,966	0.060	ANCOVA
Rebels are not responsible for their actions	2.832	-0.025	(0.029)	2,966	0.089	ANCOVA
Index of war attitudes	-	-0.024	(0.030)	3,000	0.057	CS
Indicators:						
If another war, people would not fight	0.780	-0.023	(0.016)	3,000	0.099	CS
People would not be a part of another rebellion	0.853	-0.030**	(0.015)	3,000	0.070	CS
If another war, you would not fight	0.038	0.013*	(0.007)	3,000	0.040	CS
Trust						
How much do you trust ex-combatants?	1.875	0.177**	(0.079)	900	0.222	ANCOVA
How much you trust migrants to this community?	3.161	0.123***	(0.033)	2,203	0.172	ANCOVA
Index of generalised trust in community	-	0.006	(0.027)	2,996	0.135	ANCOVA
Indicators:						
People are honest and can be trusted	2.598	0.014	(0.026)	2,994	0.126	ANCOVA
People in village are honest and can be trusted	2.858	-0.010	(0.020)	2,976	0.167	ANCOVA
People in community would not betray fellow community members	2.550	0.003	(0.028)	2,976	0.059	ANCOVA
Money left out accidentally will still be there an hour later	0.365	0.010	(0.020)	2,956	0.141	ANCOVA

Notes: Each row represents a separate regression of the outcome shown in the first column on treatment assignment. CS denotes a cross-sectional specification. Variables not shown in all regressions include section-pair fixed effects and the second-round indicator. ANCOVA specifications also include the baseline outcome variable, and the interaction of this variable with both the second-round indicator and the second-wave indicator. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level. The control mean is the mean in the control group at endline.

Table 8: War exposure, baseline forgiveness and psychological health

VARIABLES	(1) Forgive perpetrators	(2) Less PTSD	(3) Less anxiety	(4) Less depression
Raped	-1.211**	-2.358***	-0.512	-0.529
	(0.544)	(0.849)	(0.575)	(0.462)
Observations	1,470	1,918	1,986	1,999
Maimed	-0.564	-2.471***	-1.193**	-0.613
	(0.803)	(0.928)	(0.494)	(0.536)
Observations	1,475	1,925	1,990	2,005
Family member killed	-0.920***	-1.140***	-0.402**	-0.330***
	(0.232)	(0.286)	(0.195)	(0.120)
Observations	1,500	1,972	2,039	2,056

Notes: Each cell represents a separate regression of the respondent's war exposure on the baseline measure of the variables Forgive perpetrators, Less PTSD, Less anxiety and Less depression. All regressions include section-pair fixed effects. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

These forgiveness effects are based on attitudinal questions, which raises concerns that the respondents may simply be saying what they believe the surveyors want to hear. There are three factors that mitigate this concern. First, our evaluation was independent and our survey team remained completely separate from the implementing NGO over the course of the study. Second, we asked these questions between 9 and 31 months after the bonfire ceremonies took place, not in its direct aftermath, and so talk of forgiveness was not fresh in the respondents' minds. Finally, our respondents were victimised in traumatic ways, experiencing events such as amputations. Given these experiences, it would not be without psychological cost for them to respond that they no longer felt hatred towards the perpetrator or that they wished the perpetrator well, if that did not reflect a shift in their underlying perspective.

Did these effects on forgiveness also shift individuals' attitudes towards the war? To gauge views on former fighters, we aggregate two questions. We first show the treatment effect on the mean index of these indicators. The coefficient on this regression captures the effect measured in standard deviation units. The rows underneath also show results from separate regressions of the component indicators. The results indicate that overall there were no significant shifts in attitudes around the culpability of former combatants. This suggests that people can grant forgiveness, even if they continue to feel that the combatants were responsible for their actions.

Next, we examine an index of attitudes towards future war participation. This includes three indicators of whether respondents believe that they, or other members of their community, would fight in a future rebellion. We utilise a cross-sectional specification, since these variables were not collected in the baseline of the second wave. The results indicate no significant impact on this outcome. This suggests that granting forgiveness for past violence does not necessarily shift an individual's beliefs regarding future violence.

At the bottom of Table 7 we also examine impacts on trust. The first two outcomes measure trust of two socially marginalised groups – ex-combatants and migrants.[12] The third measures generalised levels of trust in the community, indexing four variables around the perceived honesty and trustworthiness of community members. The results show a clear pattern. The treatment significantly increases trust of marginalised groups, without exerting a significant impact on general levels of trust. Trust of ex-combatants increased by 9 per cent, while trust of migrants increased by 4 per cent.

7.4 Social capital

Since reconciliation is aimed at forgiving perpetrators, it is reassuring to see that the process led to increases in trust towards this group. But since it does not produce shifts in general trust, this still leaves open the question as to whether the process influences people's willingness to engage with each other and to form ties. To gain leverage on this question, Table 9 examines impacts on the strength of social networks, using an index that aggregates four outcomes. These network measures were collected comparably only at the endline, and so we are restricted to using cross-sectional specifications. Even without baseline controls, we observe significant increases in the index of network strength. The largest effects stem from the second and third indicators. The means indicate that each respondent is listed, on average, two or three times as a friend or as someone to seek help from in control areas. The reconciliation treatment boosts each network measure by 11 per cent. In short, individuals rely more on each other and are more connected to each other in treatment communities.[13]

If the reconciliation process improves individuals' outlook on their community, it may have the capacity to alter their engagement with the community more broadly. We have two ways of gauging the community orientation of individuals' behaviour.

First, Table 10 examines their participation in community groups, based on both membership and meeting attendance. The treatment exerts a clear, positive effect on this aggregate index. The coefficients on the individual indicators suggest that the largest increases occurred in PTA and religious group participation. For example, PTA membership and meeting attendance were 25% and 45% higher in treatment communities, respectively, while religious group membership and meeting participation were 20% and 31% higher, respectively.

Youth group membership and women's group meeting attendance also increased. In fact, the coefficients are positive for almost all other groups, with the exception of secret societies. This effect is noteworthy: since secret societies have closed membership dominated by the elite (Murphy 1980), decreased participation in this group is consistent with substitution towards more broad-based community organisations. The advantage of the index is that it aggregates the individual effects and demonstrates an overall effect across these various groups.

[12] As discussed in section 6.4 Data, sometimes migrants are perceived to be former combatants.
[13] It is possible that, aside from its impact on healing, attending the ceremony may have generated friendships through an alternative social channel. However, we find no significant differential impact of ceremony attendance on the mean effect of social networks, which casts doubt on the importance of this account. (The coefficient on the interaction term of treatment with attendance is 0.005 and the standard error is 0.110).

Table 9: Social networks

VARIABLES	Control mean	Coeff.	Std. Error	Obs.	R-sqr.	Specification
Index of network strength	-	0.099***	(0.001)	3,008	0.061	CS
Indicators:						
Number of people respondent would approach for advice or help	2.894	0.148**	(0.033)	3,005	0.056	CS
Number of people respondent would ask to collect money for them	3.144	0.155	(0.279)	3,005	0.026	CS
Number of times respondent listed as good friend	2.123	0.232**	(0.013)	3,008	0.192	CS
Number of times respondent listed as someone to ask for advice or help	3.245	0.362***	(0.005)	3,008	0.199	CS

Notes: Each row represents a separate regression of the outcome shown in the first column on treatment assignment. CS denotes a cross-sectional specification. Variables not shown in all regressions include section-pair fixed effects and the second-round indicator. ANCOVA specifications also include the baseline outcome variable, and the interaction of this variable with both the second-round indicator and the second-wave indicator. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level. The control mean is the mean in the control group at endline.

Table 10: Reconciliation and participation in community groups

VARIABLES	Control Mean	Coeff.	Std. Error	Obs.	R-sqr.
Index of participation in community groups	-	0.058***	(0.017)	3,004	0.160
Index of participation in community groups – without women's membership/meetings	-	0.064***	(0.017)	3,004	0.162
Indicators:					
PTA membership	0.137	0.034**	(0.016)	2,732	0.223
Village development committee membership	0.091	0.013	(0.011)	2,737	0.141
Youth group membership		0.015*	(0.008)	2,738	0.144
Women's group membership	0.118	0.022	(0.014)	2,004	0.138
Secret society membership	0.358	-0.058***	(0.019)	2,770	0.338
Religious group membership	0.286	0.055***	(0.020)	2,729	0.179
PTA meeting attendance	0.082	0.037**	(0.015)	2,739	0.138
Village development committee meeting attendance	0.068	0.008	(0.010)	2,734	0.106
Youth group meeting attendance	0.066	0.007	(0.008)	2,739	0.090
Women's group meeting attendance	0.075	0.024*	(0.013)	2,004	0.095
Secret society meeting attendance	0.056	-0.005	(0.008)	2,766	0.057
Religious group meeting attendance	0.190	0.058***	(0.016)	2,714	0.103
Community meeting attendance	0.626	0.006	(0.013)	2,983	0.077

Notes: Each row represents a separate ANCOVA regression of the outcome shown in the first column on treatment assignment. Variables not shown include section-pair fixed effects, the second-round indicator, the baseline outcome variable and the interaction of the baseline outcome variable with both the second-round indicator and the second-wave indicator. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level. The control mean is the mean in the control group at endline.

The women's group effect may raise a concern that the effects on aggregate group participation are driven by membership of Peace Mothers groups in treatment areas. However, when we remove both women's group variables from the index, the coefficient remains significant and increases in magnitude, suggesting that this is not the case.

Second, Table 11 examines effects on individuals' contributions to public goods. This index aggregates a number of different measures, including: whether individuals contribute labour or money to building and maintaining public facilities (such as bridges, schools, wells and clinics); 'road brushing', which is a common form of road maintenance; monetary contributions to needy families; and labour and monetary contributions to PTAs, village development committees, and youth and women's groups.

There is again a significant increase in the mean effect index. Among the individual indicators, the effect is most precisely estimated for PTA contributions, where the implied effect is approximately 32 per cent. The effect is also marginally insignificant for contributing to public facilities broadly, with a p-value of 0.12 and an implied effect of 7 per cent. While the estimates for the other variables are not individually significant at the 10% level, the implied effect for contributing to women's groups is also substantial (approximately 20%), as is the effect of a decision to give to those in need (8%). However, we again verify that the mean effect is not automatically driven by the women's group effect. In fact, when we remove this indicator, the coefficient again increases in magnitude and precision.

These results on network strength, group participation and public goods contributions provide evidence that the reconciliation process led to improvements in social capital. Individuals were more likely to view one another as a source of support and to contribute charitably towards community needs. The impacts on behaviour also suggest that the forgiveness effects do not arise from respondents saying what they believe surveyors want to hear – rather, more positive affect towards former combatants is coupled with actual behavioural changes in treated areas.

Table 11: Reconciliation and contributions to public goods

VARIABLES	Control Mean	Coeff.	Std. Error	Obs.	R-sqr.
Index of public goods contributions	-	0.042*	(0.022)	3,008	0.171
Index of public goods contributions (without contributions to women's group)	-	0.046**	(0.023)	3,008	0.184
Indicators appearing in endline and baseline:					
Contributed to public facilities	0.397	0.029	(0.019)	2,911	0.078
Brushed roads	0.290	0.005	(0.014)	2,898	0.171
Number of community projects (village-level variable)	0.527	-0.049	(0.055)	2,901	0.308
Contributed to PTA	0.066	0.023*	(0.013)	2,732	0.105
Contributed to village development committee	0.062	0.001	(0.008)	2,737	0.119
Contributed to youth group	0.069	-0.002	(0.006)	2,738	0.081
Contributed to women's group	0.064	0.021**	(0.010)	2,004	0.076
Contributed money to someone in need	0.196	0.015	(0.016)	2,866	0.097

Notes: Each row represents a separate ANCOVA regression of the outcome shown in the first column on treatment assignment. Variables not shown include section-pair fixed effects, the second-round indicator, the baseline outcome variable and the interaction of the baseline outcome variable with both the second-round indicator and the second-wave indicator. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level. The control mean is the mean in the control group at endline.

7.5 Psychological health

In this section, we examine how reconciliation influences individual healing, by examining its impact on psychological well-being. Table 12 examines impacts on our index of psychological well-being, which includes measures of PTSD, anxiety and depression. The first row presents the index of complete indicators (with pared baseline controls for wave two). The second row presents the index with pared-down anxiety and depression measures at endline. Both indices show that respondents in treatment communities experienced a deterioration in these outcomes. The coefficients indicate that the indices fell by approximately 0.14 standard deviation units in treated areas. Regressions of the individual indicators suggest that this negative impact stems from a worsening of all three psychometric measures.[14]

The continuous PTSD measure can also be converted into a dichotomous measure of whether an individual suffers from clinical PTSD. We construct one such measure to discern the magnitude of the trauma effect.[15] As shown in Table 12, this dichotomous measure suggests that the prevalence of PTSD increased by 36%: 8% of the individuals in the control group displayed signs of PTSD and this fraction was 2.9% higher in the treated areas. These numbers firstly indicate that a sizable fraction of individuals (in the control areas) continued to suffer from PTSD nearly a decade after the end of the war. Secondly, they suggest a substantial worsening of this psychological outcome owing to the reconciliation treatment.

We can also benchmark the estimates against baseline violence exposure (Table 6). If we compare the treatment effects against the baseline effect of being maimed, the treatment is predicted to worsen PTSD by 28%, depression by 47% and anxiety by 37%. We can also benchmark these effects against baseline violence exposure (Table 8). If we compare against the baseline effect of being maimed, the treatment is predicted to worsen PTSD by 28%, depression by 47% and anxiety by 37%.

However, different types of violence exert different impacts on baseline psychological well-being, and these translate into variable implied effects. For example, being maimed corresponds to higher baseline PTSD relative to having a family member killed.

These sizable negative impacts on psychological well-being challenge the view that reconciliation promotes individual healing through catharsis or forgiveness. Rather, they are more consistent with the view that talking about the past brings up painful memories and can potentially retraumatise individuals. Overall, the results suggest that societal and individual healing do not have to move concurrently in response to reconciliation processes.

---

[14] Note that control group means of the continuous psychometric indicators are less instructive for gauging magnitudes in percentage terms, since they are aggregations on a Likert scale. Under these scales, changing the value assigned to responses will not alter the regression coefficients, but it will alter the value of the control group mean.

[15] Following guidelines from the Clinician-Administered PTSD Scale (Weathers et al. 2013), we categorise a respondent as suffering from PTSD if he or she shows at least one symptom of re-experience, one symptom of avoidance, and at least two symptoms of increased arousal. For example, consider the question: ‘In the last month, did you sit and think of bad bad things that happened to you even though you don’t want to think of it?’ Of the responses ‘never’, ‘yes small small’, ‘yes sometimes’ and ‘yes often’, we conservatively code a respondent as having a symptom if they report either ‘yes sometimes’ or ‘yes often’.

Table 12: Psychological well-being

VARIABLES	Control mean	Coeff.	Std. Error	Obs.	R-sqr.
Index of psychological well-being (all indicators)	-	-0.147***	(0.033)	2,982	0.115
Index of psychological well-being (indicators in both baselines)	-	-0.138***	(0.031)	2,982	0.115
Indicators (in both baselines):					
Less PTSD	28.819	-0.683***	(0.197)	2,776	0.119
Less anxiety	14.945	-0.441***	(0.117)	2,895	0.142
Less depression	11.677	-0.289***	(0.069)	2,913	0.092
PTSD symptoms present	0.080	0.029***	0.011	2,776	0.057

Notes: Each row represents a separate ANCOVA regression of the outcome shown in the first column on treatment assignment. Variables not shown include section-pair fixed effects, the second-round indicator, the baseline outcome variable and the interaction of the baseline outcome variable with both the second-round indicator and the second-wave indicator. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level. The control mean is the mean in the control group at endline.

7.6 Economic activity and outcomes

Table 13 shows that economic outcomes and activities did not increase systematically in response to treatment. There is an interesting pattern, whereby an objective assets indicator increased, while perceptions of economic well-being decreased; this could reflect a more negative outlook from greater depression in treatment areas. However, as a whole, we see no consistent effects on economic variables. In particular, within economic activities, there were no significant effects on farming-related activities, such as labour devoted to working other people's farms. This suggests a limited impact of the FT communal farms. Furthermore, when we control for baseline imbalance (Table 14), the significant negative impact on economic outcomes disappears.

Table 13: Economic activity and economic outcomes

VARIABLES	Control mean	Coeff.	Std. Error	Obs.	R-sqr.
Index of economic outcomes	-	-0.036*	(0.019)	3,008	0.161
Index of economic outcomes – control for baseline of individual economic indicators	-	-0.027	(0.020)	2,839	0.227
Indicators:					
Objective indicator of household assets	0.047	0.145***	(0.055)	2,836	0.416
Perception that household needs are met	10.079	-0.117	(0.072)	2,835	0.102
Perception of overall household economic situation compared to one year ago	2.882	-0.131***	(0.037)	2,831	0.088
Baseline balance on individual economic outcome indicators					
Indicator: Household assets index	-	-0.031	(0.065)	2,205	
Indicator: Perception that household needs met	-	0.383***	(0.144)	2,131	
Indicator: Perceived satisfaction with household economic situation	-	-0.107*	(0.054)	2,133	
Index of economic activity	-	0.034	(0.026)	3,008	0.182
Indicators:					
Frequency of borrowing and lending	2.17	0.043	(0.036)	3,008	0.461
Monetary value of borrowing and lending	5.17	0.302	(0.221)	2,915	0.104
Respondent belongs to an *osusu* (savings group)	0.396	-0.015	(0.018)	2,950	0.144
Number of traders (village-level indicator)	9.356	0.743	(1.513)	2,710	0.501
Respondent buys from trader	0.899	-0.011	(0.011)	2,956	0.076
Number of communal farms (village-level indicator)	0.558	0.096	(0.103)	2,820	0.359
Respondent belongs to a labour gang	0.333	0.002	(0.016)	2,738	0.164
Days spent working on others' farms	7.96	0.473	(0.618)	2,418	0.130

Notes: In the first and third panels, each row represents a separate ANCOVA regression of the outcome shown in the first column on treatment assignment. Variables not shown include section-pair fixed effects, the second-round indicator, the baseline outcome variable and the interaction of the baseline outcome variable with both the second-round indicator and the second-wave indicator. In the second panel, the baseline value of each indicator is regressed on treatment assignment. In all specifications, standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level. The control mean is the mean in the control group at endline.

Table 14: Controlling for baseline imbalance

VARIABLES	Coeff.	Std. Error	Obs.	R-sqr.	Specification
Forgive perpetrators	0.548**	(0.239)	1,919	0.143	ANCOVA
Trust ex-combatants	0.222***	(0.076)	860	0.240	ANCOVA
Trust migrants	0.110***	(0.034)	2,084	0.181	ANCOVA
Index of generalised trust in community	0.003	(0.024)	2,832	0.145	ANCOVA
Index of attitudes towards ex-combatants	-0.012	(0.030)	2,818	0.085	ANCOVA
Index of war attitudes	-0.025	(0.029)	2,831	0.062	CS
Index of network strength	0.130***	(0.035)	2,839	0.077	CS
Index of community group participation	0.060***	(0.018)	2,836	0.174	ANCOVA
Index of contributions to public goods	0.044*	(0.022)	2,839	0.180	ANCOVA
Index of psychological well-being	-0.142***	(0.033)	2,820	0.130	ANCOVA
Less PTSD	-0.664***	(0.196)	2,628	0.135	ANCOVA
Less anxiety	-0.395***	(0.118)	2,738	0.149	ANCOVA
Less depression	-0.279***	(0.069)	2,839	0.228	ANCOVA
Index of economic outcomes	-0.026	(0.020)	2,839	0.228	ANCOVA
Assets	0.149***	(0.054)	2,836	0.418	ANCOVA
Perception of household needs	-0.115	(0.071)	2,835	0.102	ANCOVA
Perception of economic situation	-0.130***	(0.037)	2,831	0.088	ANCOVA
Index of economic activity	0.037	(0.026)	2,839	0.192	ANCOVA
Index of social tensions	0.030	(0.022)	2,832	0.094	ANCOVA
Index of conflicts and crime	0.117	(0.083)	259	0.333	CS
Index of attitudes towards women	0.043	(0.026)	2,818	0.044	ANCOVA

Notes: Each row represents a separate regression of the outcome shown in the first column on treatment assignment. Indices are constructed using the Kling and others (2007) methodology. All regressions include section-pair fixed effects and the second-round indicator. ANCOVA specifications also include the baseline outcome variable, and the interaction of this variable with both the second-round indicator and the second-wave indicator. All specifications control for baseline measures of the trust index and the individual indicators comprising the economic outcomes index which showed imbalance. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

7.7 Persistence

A key issue is whether these effects persist over time. For example, do the estimates reflect just short-run effects on forgiveness? Also, are the effects on trauma sharpest in the first months after the ceremony, and do they subsequently recede?

We are able to explore short-run and long-run effects using Wave 1 data, where the endlines were administered 9 and 31 months after the ceremonies, respectively. Table 15 presents the results separately by these two rounds.

Since the first wave includes less than half the sections in the evaluation, this is a relatively underpowered sample, and some of the effects are not individually significant. Yet, the pattern in this table shows that most effects – both positive and negative – last over time.

First, the forgiveness effect persists: to the extent that respondents have been able to let go of their anger towards former combatants, these changes do not appear to recede over time. While trust towards former combatants is individually insignificant in both rounds, these coefficients are not significantly distinguishable from each other at the 5 per cent level, indicating that the effects do not diminish in a meaningful manner. Trust of

migrants also persists, and there are even short-run improvements in attitudes towards former combatants and the generalised trust measures, though these effects fall – and significantly so – over the longer horizon.

Notably, increased participation in community groups is also sustained for over two and a half years down the road. Public goods contributions and social networks are individually insignificant in both subsamples, but the coefficients display a similar pattern, increasing in magnitude during the second round. As such, the reconciliation effects do not reflect a short-run burst in community orientation and civic engagement that subsequently fade away.

At the same time, Table 15 also establishes that the negative psychological impacts persist over the course of the two rounds. These effects point to the potentially lasting damage associated with participating in the reconciliation process.

Table 15: Persistence of effects

Sample:	Round 1			Round 2		
VARIABLES	Coeff.	Std. Error	Obs.	Coeff.	Std. Error	Obs.
Forgive perpetrators	0.986***	(0.272)	550	1.231***	(0.361)	521
Trust ex-combatants	0.100	(0.073)	241	0.048	(0.198)	203
Trust migrants	0.140**	(0.053)	653	0.119*	(0.069)	564
Index of generalised trust in community	0.119**	(0.050)	878	-0.009	(0.038)	845
Index of attitudes towards ex-combatants	0.115**	(0.052)	875	-0.065	(0.055)	841
Index of war attitudes	0.015	(0.033)	828	-0.039	(0.063)	789
Index of network strength	0.015	(0.027)	885	0.119	(0.085)	850
Index of community group participation	0.038*	(0.022)	884	0.084**	(0.040)	847
Index of contributions to public goods	0.024	(0.033)	885	0.035	(0.046)	850
Index of psychological well-being	-0.166***	(0.052)	873	-0.170***	(0.058)	837
Index of economic outcomes	-0.014	(0.029)	885	-0.014	(0.023)	850
Assets	0.195**	(0.092)	879	0.182**	(0.083)	842
Perception of household needs	-0.052	(0.142)	806	-0.146*	(0.080)	780
Perception of economic situation	-0.113*	(0.057)	811	-0.067	(0.060)	784
Index of economic activity	-0.025	(0.041)	885	0.021	(0.035)	850
Index of conflict and crime	0.201	(0.139)	80	0.130	(0.122)	78
Index of social tensions	0.021	(0.026)	878	0.060	(0.039)	845
Index of attitudes towards women	0.068*	(0.038)	877	0.007	(0.045)	844

Notes: These results present separate estimates for the two endline rounds in Wave 1. Each row represents a separate regression of the outcome shown in the first column on treatment assignment. Variables not shown include section-pair fixed effects and the second-round indicator. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level. The control mean is the mean in the control group at endline.

7.8 Additional results: conflicts and conflict resolution, gender attitudes

Finally, we find no systematic reductions in crimes and conflicts, or improvements in their resolution, though we do observe greater resolution by chiefs and less by friends and family (Table 16). The overall conflict effects suggest that the reconciliation process

influences outcomes related to the war, but does not prevent the occurrence of other day-to-day disputes. These results reinforce the idea that war-related reconciliation is an important element of the intervention. If merely gathering the community was crucial in improving interpersonal dynamics, then we should observe other outcomes, such as day-to-day disputes falling or economic activities increasing.

Table 16: Societal conflicts

VARIABLES	Control mean	Coeff.	Std. Error	Obs.	R-sqr.	Specification
Social tensions						
Index of social tensions	-	0.028	(0.021)	2,996	0.085	ANCOVA
Indicators:						
Dominant groups do not benefit more from community resources	1.127	0.026	(0.040)	2,963	0.094	ANCOVA
Marginalised groups benefit from community resources	7.367	0.053	(0.056)	2,809	0.092	ANCOVA
Respondent feels included and respected in the community	0.95	0.000	(0.005)	2,960	0.050	ANCOVA
Perception that social divisions escalated into conflict	0.506	0.037	(0.037)	2,943	0.140	ANCOVA
Conflict and crime						
Index of conflict and crime	-	0.112	(0.072)	274	0.275	CS
Indicators:						
Number of conflicts	0.158	0.002	(0.019)	274	0.320	CS
Number of crimes	0.039	-0.005	(0.007)	274	0.226	CS
Number of violent crimes	0.004	0.003	(0.003)	273	0.172	CS
Number of inter-village conflicts	0.124	0.122***	(0.042)	274	0.295	CS
Conflict resolution						
Resolved	0.85	-0.057	(0.050)	172	0.330	ANCOVA
Satisfactorily resolved	0.753	-0.107	(0.067)	172	0.456	ANCOVA
Resolved without third party	0.218	-0.036	(0.035)	172	0.603	ANCOVA
Resolved with mediation from family/friends	0.428	-0.141**	(0.055)	172	0.547	ANCOVA
Resolved by chief	0.43	0.103*	(0.058)	172	0.326	ANCOVA
Fined by chief	0.09	-0.007	(0.009)	280	0.255	ANCOVA
Gender attitudes						
Index of attitude towards women	-	0.044*	(0.025)	2,982	0.036	ANCOVA
Indicators:						
Attitude towards wife beating	18.856	0.081	(0.115)	2,957	0.036	ANCOVA
Belief that a wife has a right to her own opinion	0.888	0.019**	(0.008)	2,957	0.055	ANCOVA

Notes: Each row represents a separate regression of the outcome shown in the first column on treatment assignment. CS denotes a cross-sectional specification. Variables not shown in all regressions include section-pair fixed effects and the second-round indicator. ANCOVA specifications also include the baseline outcome variable, and the interaction of this variable with both the second-round indicator and the second-wave indicator. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level. The control mean is the mean in the control group at endline.

We also see improvements in attitudes towards women (Table 16): respondents are more likely to believe that a wife has a right to express her own opinion and are less likely to believe that there are circumstances where it is acceptable for a husband to beat his wife. This is likely to be due to the influence of the Peace Mothers groups. However, we do not place a lot of weight on this result, because the effect is relatively small and statistical significance is not robust to adjusting the standard errors to account for multiple comparisons (more below).

7.9 Robustness checks

We also present additional robustness checks. We control for the presence of FT communal farms (Table 17), since only some treatment areas had a farm. Our core results are unaffected, suggesting that our results are not driven by the communal farms. The results suggest that, while FT has several components, our results are likely to be due to the reconciliation component, rather than to auxiliary impacts on economic activity from communal farms.

Finally, Table 18 shows that constructing the mean indices using the approach of Anderson (2008) and controlling for baseline imbalance does not meaningfully affect our results.

Table 17: Controlling for FT communal farm

VARIABLES	Coeff.	Std. Error	Obs.	R-sqr.	Specification
Forgive perpetrators	0.640**	(0.248)	2,010	0.132	ANCOVA
Trust ex-combatants	0.122	(0.080)	900	0.227	ANCOVA
Trust migrants	0.129***	(0.038)	2,203	0.172	ANCOVA
Index of generalised trust in community	0.038	(0.030)	2,996	0.139	ANCOVA
Index of attitudes towards ex-combatants	-0.001	(0.031)	2,980	0.075	ANCOVA
Index of war attitudes	-0.015	(0.034)	3,000	0.058	CS
Index of network strength	0.132***	(0.036)	3,008	0.064	CS
Index of community group participation	0.052***	(0.017)	3,004	0.161	ANCOVA
Index of contributions to public goods	0.044*	(0.024)	3,008	0.171	ANCOVA
Index of psychological well-being	-0.161***	(0.036)	2,982	0.116	ANCOVA
Less PTSD	-0.732***	(0.224)	2,776	0.115	ANCOVA
Less anxiety	-0.455***	(0.128)	2,895	0.139	ANCOVA
Less depression	-0.300***	(0.074)	2,913	0.090	ANCOVA
Index of economic outcomes	-0.042**	(0.020)	3,008	0.161	ANCOVA
Assets	0.105*	(0.055)	2,991	0.403	ANCOVA
Perception of household needs	-0.167**	(0.080)	2,857	0.083	ANCOVA
Perception of economic situation	-0.125***	(0.042)	2,860	0.081	ANCOVA
Index of economic activity	0.029	(0.029)	3,008	0.182	ANCOVA
Index of conflict and crime	0.112	(0.072)	274	0.275	CS
Index of social tensions	0.056**	(0.023)	2,996	0.089	ANCOVA
Index of attitudes towards women	0.041	(0.027)	2,982	0.035	ANCOVA

See Table 11 for notes.

Table 18: Impacts using indices as constructed by Anderson (2008)

VARIABLES	Coeff.	Std. Error	Obs.	R-sqr.	Specification
Index of attitudes towards ex-combatants	-0.004	(0.029)	2,960	0.073	ANCOVA
Index of war attitudes	-0.014	(0.028)	3,000	0.044	CS
Index of generalised trust in community	0.015	(0.029)	2,915	0.121	ANCOVA
Index of network strength	0.076**	(0.012)	3,005	0.112	CS
Index of participation in community groups	0.035**	(0.017)	1,930	0.159	ANCOVA
Index of public goods contributions	0.022	(0.025)	1,853	0.206	ANCOVA
Index of psychological well-being (all indicators)	-0.143***	(0.034)	2,635	0.121	ANCOVA
Index of psychological well-being (indicators in both baselines)	-0.133***	(0.031)	2,667	0.120	ANCOVA
Index of economic outcomes	-0.039*	(0.020)	2,831	0.134	ANCOVA
Index of economic activity	-0.018	(0.035)	1,861	0.239	ANCOVA
Index of social tensions	0.028	(0.021)	2,996	0.085	ANCOVA
Index of conflict and crime	-0.001	(0.064)	273	0.051	CS
Index of attitudes towards women	0.044*	(0.026)	2,920	0.037	ANCOVA

Notes: Each row represents a separate regression of the outcome shown in the first column on treatment assignment. CS denotes a cross-sectional specification. Variables not shown in all regressions include section-pair fixed effects and the second-round indicator. ANCOVA specifications also include the baseline outcome variable, and the interaction of this variable with both the second-round indicator and the second-wave indicator. Standard errors are clustered at the section level. *** is significant at the 1% level, ** is significant at the 5% level and * is significant at the 10% level.

7.10 Multiple comparison corrections

Our study examines impacts on multiple outcomes that are conceptually related to one another. Since multiple tests raise the risk of falsely rejecting true null hypotheses, here we account for multiple comparisons by controlling for the FDR using the method proposed by Benjamini and Hochberg (1995).

The Benjamini–Hochberg method requires selecting a threshold rate, FDR, and ranking n hypotheses in a family from the highest to the lowest *p-value*. The i^{th} hypothesis is then assigned the false discovery rate critical significance level of $d_i = FDR * (i/n)$. For example, if we select an FDR of 5 per cent and have 10 hypotheses, the critical level of significance for the least significant hypothesis (with the highest *p-value*) is 0.05, while the critical significance level for the most significant hypothesis (with the lowest *p-value*) is 0.05 / 10 = 0.005. Each hypothesis is then determined to be significant after accounting for the false discovery rate if the *p-value*$_i < d_i$.

One important factor in making multiple comparison corrections is how hypotheses are grouped into different families of outcomes, since this determines the degree of penalty applied to the p-values. We present two approaches. We apply the FDR controlling procedure to all the hypotheses as they were grouped together in the sections of our PAP. These groupings were thematically related. For example, forgiveness and psychological well-being were grouped together, since they are both psychology-related outcomes. Table 19 shows these results. We rank each hypothesis from least to most significant, and the first column shows these rankings. The final three columns show three different levels of the adjusted critical level of significance, based on different threshold FDRs that we are willing to accept.

The estimates indicate that none of our main results change with the adjustments. For example, the most precisely estimated effect under social capital was trust of migrants (with a p-value of 0.0003, which is significant at the 1 per cent level, in unadjusted terms). The maximal penalty is applied to this outcome, yielding an adjusted critical significance level of 0.001 for an FDR of 1 per cent (which is a tougher criterion to meet than the standard 0.01 for unadjusted 1 per cent significance). Yet 0.0003 < 0.001, and so the effect remains significant at the 1 per cent level, even with this adjustment. Analogously, our index of contributions was only significant at the 10 per cent level (in unadjusted terms) with a p-value of 0.055. As the table shows, it continues to remain significant under an FDR of 10 per cent (after facing an adjusted critical significance level of 0.056 instead of the standard 0.10 for unadjusted 10 per cent significance). This table shows that effects on social capital remain in place even after we include general trust, as well as measured trust towards all various subgroups.

Table 19 further shows that none of the other additional results that were statistically significant with unadjusted p-values continue to be statistically significant after we adjust for multiple comparisons. For example, the impacts on economic outcomes (p = 0.058) and attitudes towards women (p = 0.074) appeared to be significant at the 10 per cent level when considered in unadjusted terms, but do not remain so after we account for an FDR of at least 10 per cent. This does not necessarily mean that there are no true effects, but simply we cannot reject the null hypothesis of no impact after accounting for multiple comparisons.

Table 19: Adjusting for multiple comparisons with pre-analysis plan groupings

Comparison *i*		Variable	Coeff.	p-value	Adjusted critical significance level (d_i)		
					FDR = 0.01	FDR = 0.05	FDR = 0.1
Social capital							
	9	Index of generalised trust in community	0.006	0.816	0.01	0.050	0.100
	8	Trust of former Sierra Leonean Army	0.019	0.783	0.009	0.044	0.089
	7	Trust of former CDF	0.029	0.638	0.008	0.039	0.078
	6	Index of social divisions	0.028	0.199	0.007	0.033	0.067
	5	Index of contributions to public goods	0.042+	0.055	0.006	0.028	0.056
	4	Trust of rebel ex-combatants	0.177+	0.027	0.004	0.022	0.044
	3	Index of network strength	0.112+++	0.002	0.003	0.017	0.033
	2	Index of community group participation	0.058+++	0.001	0.002	0.011	0.022
	1	Trust migrants	0.123+++	0.0003	0.001	0.006	0.011
Forgiveness and psychological well-being							
	2	Forgive perpetrators	0.571++	0.0134	0.01	0.050	0.100
	1	Index of psychological well-being	-0.147+++	0.00002	0.005	0.025	0.050
Attitude and beliefs							
	3	Index of attitudes towards ex-combatants	-0.007	0.821	0.01	0.050	0.100
	2	Index of war attitudes	-0.024	0.422	0.007	0.033	0.067
	1	Index of attitude towards women	0.044	0.074	0.003	0.017	0.033
Conflict and conflict resolution							
	7	Fined by chief	-0.009	0.342	0.01	0.050	0.100
	6	Resolved without third party	-0.036	0.308	0.009	0.043	0.086
	5	Resolved	-0.057	0.260	0.007	0.036	0.071
	4	Index of conflict and crime	0.112	0.122	0.006	0.029	0.057
	3	Satisfactorily resolved	-0.107	0.117	0.004	0.021	0.043
	2	Resolved by chief	0.103	0.082	0.003	0.014	0.029
	1	Resolved with mediation from family/friends	-0.141	0.013	0.001	0.007	0.014
Economic activity and welfare							
	2	Index of economic activity	0.034	0.187	0.01	0.050	0.100
	1	Index of economic outcomes	-0.036	0.058	0.005	0.025	0.050

Notes: This table adjusts for multiple comparisons through an FDR controlling procedure, based on the groups in the PAP. Each hypothesis is ranked in order of lowest p-value to highest p-value. Column 1 shows this ranking. Column 4 reports the unadjusted p-value. The final three columns show the critical adjusted levels of significance for different false discovery rate thresholds. +++ denotes that the effect is significant with an FDR of 1%, ++ is significant with an FDR of 5%, + is significant with an FDR of 1%.

While the results above are presented for groupings under the PAP, one could argue that a different grouping would be relevant for conceptualising families. Most notably, we posit a close conceptual relationship between forgiveness and social capital outcomes in describing the potential mechanism under our section on 'Healing under Reconciliation'. Under this conceptualisation, social networks and associated norms such as trust, as well as greater community participation and contribution, may result as a consequence of forgiveness. Under this theory of change, forgiveness and these social capital outcomes arguably belong in the same family of outcomes.

Thus, for robustness, we also apply the FDR controlling approach to this alternative grouping in Table 20. Panel A shows that our results are again unaffected. Finally, in Panel B, we further incorporate psychological well-being into the grouping, though arguably it is a separate category of outcome. We again find that our results remain unaffected. Thus, adjusting for multiple comparisons under various families does not influence our findings.

Table 20: Adjusting for multiple comparisons

Comparison	Variable	Coeff.	p-value	Adjusted critical level of significance (d_i)		
				FDR = 0.01	FDR = 0.05	FDR = 0.1
Panel A. Forgiveness and social capital						
9	Index of ge	0.006	0.816	0.010	0.050	0.100
8	Trust of for	0.019	0.783	0.009	0.044	0.089
7	Trust of for	0.029	0.638	0.008	0.039	0.078
6	Index of co	0.042+	0.055	0.007	0.033	0.067
5	Trust rebel	0.177++	0.027	0.006	0.028	0.056
4	Forgive per	0.571++	0.0134	0.004	0.022	0.044
3	Index of ne	0.112+++	0.002	0.003	0.017	0.033
2	Index of co	0.058+++	0.001	0.002	0.011	0.022
1	Trust migra	0.123+++	0.0003	0.001	0.006	0.011
Panel B. Forgiveness, psychological well-being and social capital						
10	Index of ge	0.006	0.816	0.010	0.050	0.100
9	Trust of for	0.019	0.783	0.009	0.045	0.090
8	Trust of for	0.029	0.638	0.008	0.040	0.080
7	Index of co	0.042+	0.055	0.007	0.035	0.070
6	Trust rebel	0.177++	0.027	0.006	0.030	0.060
5	Forgive per	0.571++	0.0134	0.005	0.025	0.050
4	Index of ne	0.112+++	0.002	0.004	0.020	0.040
3	Index of co	0.058+++	0.001	0.003	0.015	0.030
2	Trust migra	0.123+++	0.0003	0.002	0.010	0.020
1	Index of psy	-0.147+++	0.00002	0.001	0.005	0.010

Notes: This table adjusts for multiple comparisons through an FDR controlling procedure, applied to the conceptually related hypotheses around forgiveness and social capital (Panel A) as well as psychological well-being (Panel B). Each hypothesis is ranked in order of lowest p-value to highest p-value. Column 1 shows this ranking. Column 4 reports the unadjusted p-value. The final three columns show the critical adjusted levels of significance for different false discovery rate thresholds. +++ denotes that the effect is significant with an FDR of 1%, ++ is significant with an FDR of 5%, + is significant with an FDR of 1%.

7.1 Heterogeneous impacts

In line with our PAP, we also look at subgroup effects, testing if the impact of the programme is larger/smaller based on gender or exposure to violence. We find no differential effects based on gender (Table 21), and limited impact based on exposure to violence (Table 22) or being a former combatant (Table 23). Experiencing more violence may generate a greater need for reconciliation, or may make it more difficult to reconcile. The null violence interactions are consistent with this theoretical ambiguity and suggest that neither effect dominates the other.

Table 21: Impacts by gender

	T		T x Female				
					Obs.	**R-sqr.**	**Specification**
VARIABLES	Coeff.	Std. Error	Coeff.	Std. Error			
Forgive perpetrators	0.249	(0.341)	0.534	(0.568)	2,009	0.138	ANCOVA
Trust ex-combatants	0.165	(0.109)	0.024	(0.144)	900	0.222	ANCOVA
Trust migrants	0.132***	(0.045)	-0.021	(0.061)	2,203	0.174	ANCOVA
Index of generalised trust in community	-0.013	(0.035)	0.032	(0.048)	2,995	0.136	ANCOVA
Index of attitudes towards ex-combatants	-0.054	(0.043)	0.083	(0.055)	2,979	0.077	ANCOVA
Index of war attitudes	-0.024	(0.044)	-0.001	(0.051)	2,999	0.058	CS
Index of network strength	0.137**	(0.060)	-0.059	(0.076)	3,004	0.075	CS
Index of community group participation	0.069***	(0.025)	-0.024	(0.031)	3,003	0.164	ANCOVA
Index of contributions to public goods	0.035	(0.029)	-0.001	(0.033)	3,004	0.196	ANCOVA
Index of psychological well-being	-0.133***	(0.042)	-0.034	(0.066)	2,981	0.123	ANCOVA
Less PTSD	-0.526**	(0.260)	-0.328	(0.394)	2,775	0.124	ANCOVA
Less anxiety	-0.534***	(0.174)	0.127	(0.266)	2,894	0.149	ANCOVA
Less depression	-0.270***	(0.097)	-0.066	(0.156)	2,912	0.103	ANCOVA
Perception of household needs	-0.222*	(0.112)	0.153	(0.152)	2,856	0.086	ANCOVA
Perception of economic situation	-0.083	(0.060)	-0.092	(0.093)	2,860	0.084	ANCOVA
Index of economic activity (individual level)	0.029	(0.035)	-0.032	(0.046)	2,099	0.157	ANCOVA
Frequency of borrowing and lending	0.093	(0.064)	-0.102	(0.084)	3,004	0.465	ANCOVA
Monetary value of borrowing and lending	0.770**	(0.338)	-0.860**	(0.430)	2,911	0.106	ANCOVA
Respondent belongs to an *osusu* (savings group)	-0.018	(0.028)	0.010	(0.042)	2,949	0.151	ANCOVA
Respondent buys from trader	-0.003	(0.018)	-0.016	(0.023)	2,955	0.078	ANCOVA
Respondent belongs to a labour gang	-0.004	(0.026)	0.002	(0.035)	2,738	0.178	ANCOVA
Days spent working on others' farms	0.830	(1.106)	-1.124	(1.385)	2,414	0.145	ANCOVA
Index of social tensions	0.020	(0.029)	0.009	(0.046)	2,995	0.088	ANCOVA
Index of attitudes towards women	0.034	(0.036)	0.005	(0.058)	2,981	0.054	ANCOVA

Table 22: Impacts by exposure to violence

	T		T x Violence-exposed				
VARIABLES	Coeff.	Std. Error	Coeff.	Std. Error	Obs.	R-sqr.	Specification
Forgive perpetrators	0.552	(0.424)	-0.105	(0.587)	1,945	0.136	ANCOVA
Trust ex-combatants	0.229*	(0.116)	-0.045	(0.136)	873	0.226	ANCOVA
Trust migrants	0.172***	(0.049)	-0.084	(0.066)	2,108	0.175	ANCOVA
Index of generalised trust in community	-0.016	(0.036)	0.003	(0.046)	2,861	0.144	ANCOVA
Index of attitudes towards ex-combatants	-0.009	(0.040)	-0.011	(0.053)	2,848	0.075	ANCOVA
Index of war attitudes	-0.039	(0.039)	0.013	(0.049)	2,861	0.060	CS
Index of network strength	0.093**	(0.041)	0.040	(0.069)	2,868	0.068	CS
Index of community group participation	0.072***	(0.026)	-0.021	(0.033)	2,865	0.163	ANCOVA
Index of contributions to public goods	0.037	(0.028)	0.011	(0.032)	2,868	0.177	ANCOVA
Index of psychological well-being	-0.160***	(0.052)	0.011	(0.064)	2,852	0.121	ANCOVA
Less PTSD	-0.871***	(0.309)	0.298	(0.391)	2,662	0.123	ANCOVA
Less anxiety	-0.476**	(0.213)	0.003	(0.268)	2,778	0.144	ANCOVA
Less depression	-0.270**	(0.127)	-0.044	(0.162)	2,788	0.094	ANCOVA
Perception of household needs	-0.062	(0.128)	-0.138	(0.157)	2,732	0.085	ANCOVA
Perception of economic situation	-0.232***	(0.063)	0.154**	(0.076)	2,736	0.088	ANCOVA
Index of economic activity (individual level)	-0.006	(0.033)	0.016	(0.046)	2,868	0.187	ANCOVA
Index of social tensions	0.014	(0.032)	0.017	(0.042)	2,861	0.084	ANCOVA
Index of attitudes towards women	0.019	(0.039)	0.027	(0.053)	2,847	0.039	ANCOVA

Table 23: Impacts by ex-combatants

	T		T x Ex-combatant				
					Obs.	R-sqr.	Specification
VARIABLES	Coeff.	Std. Error	Coeff.	Std. Error			
Forgive perpetrators	0.462*	(0.234)	0.828	(0.854)	1,930	0.138	ANCOVA
Trust ex-combatants	0.209**	(0.084)	-0.057	(0.209)	868	0.227	ANCOVA
Trust migrants	0.128***	(0.035)	-0.009	(0.123)	2,081	0.178	ANCOVA
Index of generalised trust in community	-0.014	(0.025)	-0.006	(0.094)	2,819	0.141	ANCOVA
Index of attitudes towards ex-combatants	-0.015	(0.030)	0.001	(0.123)	2,806	0.076	ANCOVA
Index of war attitudes	-0.029	(0.028)	-0.092	(0.126)	2,819	0.062	CS
Index of network strength	0.100***	(0.036)	0.525	(0.573)	2,826	0.073	CS
Index of community group participation	0.060***	(0.019)	0.007	(0.095)	2,823	0.164	ANCOVA
Index of contributions to public goods	0.043*	(0.023)	0.064	(0.082)	2,826	0.176	ANCOVA
Index of psychological well-being	-0.154***	(0.037)	0.136	(0.157)	2,810	0.125	ANCOVA
Less PTSD	-0.703***	(0.223)	0.355	(0.914)	2,626	0.126	ANCOVA
Less anxiety	-0.495***	(0.127)	0.997	(0.604)	2,736	0.145	ANCOVA
Less depression	-0.301***	(0.076)	0.317	(0.348)	2,747	0.095	ANCOVA
Perception of household needs	-0.112	(0.075)	-0.150	(0.313)	2,700	0.085	ANCOVA
Perception of economic situation	-0.134***	(0.041)	0.026	(0.213)	2,704	0.086	ANCOVA
Index of economic activity (individual level)	0.004	(0.023)	0.064	(0.099)	2,826	0.187	ANCOVA
Index of social tensions	0.024	(0.023)	-0.064	(0.097)	2,819	0.082	ANCOVA
Index of attitudes towards women	0.040	(0.028)	0.093	(0.105)	2,805	0.041	ANCOVA

8. Conclusion and recommendations

Countries emerging from internal conflict face the challenge of rebuilding social capital and renewing their societies. Yet, we have a limited understanding of how to ignite this process. In particular, we know little about the workings of reconciliation efforts, which are commonly pursued as an avenue for healing societies and individuals.

We study a community-level reconciliation programme in Sierra Leone, which has been designed and implemented by a local NGO. In evaluating the programme, we find negative and positive consequences associated with this approach. On the one hand, the reconciliation process increased forgiveness and also improved social capital: social networks were stronger and people contributed more to their communities in treatment villages.

These are important impacts, since vast resources are spent on trying to improve social capital outcomes in post-conflict contexts. For example, a well-implemented community-driven development or reconstruction (CDD/R) programme in Sierra Leone, which seeks to unify communities by encouraging individuals to work together under participatory aid programmes, found no impact on local public goods provision (Casey et al. 2012). (CDD/R programmes are distinct from reconciliation efforts, but share the goal of improving social capital.)

At the same time, these benefits came at a substantial cost: the reconciliation process also exerted negative psychological impacts, leading to greater trauma, anxiety and depression in the treated areas. These effects persisted for nearly three years after the reconciliation process was complete. Our results suggest that societal and individual healing do not need to move in tandem as a response to reconciliation. Rather, one form of healing can even come at the expense of the other.

These results are in contrast to studies that have shown that forgiveness can improve psychological health, yet they are consistent with the psychological literature on the risk of retraumatisation due to a one-off exposure to a past traumatic event (Joseph and Gray 2008; van Emmerik et al. 2002; Brounéus 2010).

Our findings highlight the long shadow of war along two dimensions. The programme we study was implemented nearly 10 years after the end of the civil war. The positive effects on forgiveness and social capital suggest that the need for reconciliation persists long after the violence ends. At the same time, the negative psychological impacts indicate that truth telling opened up psychological wounds, pointing to the potency of war memories. An important question remaining is whether these effects would have differed, had reconciliation occurred in the direct aftermath of the conflict. For example, the psychological impacts may have been incrementally smaller if trauma levels were already high, owing to more recent memories of the war.

These results have implications not only for reconciliation efforts, but for a wide range of transitional justice programmes that require survivors to talk about their experience during the war. These include international courts that try perpetrators, truth commissions that uncover facts surrounding past atrocities, reconciliation programmes that encourage repentance and forgiveness, and hybrid local courts (such as the *gacaca* courts in Rwanda) that combine retribution with reconciliation. If survivors are worse off

by testifying at these forums, then interventions that promote transitional justice need to be reformed.

8.1 Recommendations

These dual results – good for social capital, bad for psychological health – imply that policymakers involved in reconciliation processes and truth commissions need to be aware of their potential psychological effects, and need to take this cost into account in their decision to conduct these programmes.

It is important to note, though, that our results in no way suggest that reconciliation programmes should be stopped. The large impacts on social capital are striking, especially given the lack of success of far more resource-intensive CDD/R programmes in improving social capital. And they serve other important, long-term functions for society that are difficult to quantify. For example, recorded testimonials guard against selective retelling of history for political purposes at a later stage; and a perception that justice was done can limit festering resentment.

Moreover, over the longer run, our results suggest that reconciliation programmes should be redesigned so that they minimise these psychological costs, while retaining the societal benefits. Although our research cannot tell us *how*, it is clear that programmes should be better integrated with psychological evidence on how to deal with past trauma: it is important that the survivor should relive the experience, but it must be in an environment where he or she feels safe, comfortable and in control of the memory, until the fear, sense of threat and resultant anxiety associated with the event subside (Brounéus 2008).

Combining these processes with ongoing counselling and/or group therapy may therefore mitigate the distress experienced by programme participants. If properly designed, group therapy is effective in treating PTSD, since it creates a safe and trusting environment in which to discuss the trauma, provides important social support and also allows for social learning (for example, sharing of coping strategies) (Foa et al. 2009). In fact, facilitated group discussions have the potential to promote both individual and societal healing. A programme in Rwanda that provided psycho-educational lectures in both large-group and small-group settings led to both a reduction in trauma and a positive orientation towards others (Staub et al. 2005).

Future work should systematically examine the impact of a combined effort along these lines.

Table 24: Robustness: Trust of ex-combatants and migrants

VARIABLES	Coeff.	Std. Error	Obs.	R-sqr.	Specification
Do you know any ex-combatants?	-0.034	(0.024)	2,970	0.186	ANCOVA
Do you know any migrants?	-0.012	(0.013)	3,008	0.116	ANCOVA
Do you know any ex-combatants?	-0.024	(0.024)	3,000	0.179	CS
Do you know any migrants?	-0.011	(0.013)	3,008	0.109	CS
How much do you trust ex-combatants?	0.145**	(0.066)	1,470	0.177	CS
How much do you trust migrants to this community?	0.083***	(0.032)	2,522	0.167	CS

Notes: Each row represents a separate regression of the outcome shown in the first column on treatment.

Bibliography

Al-Mabuk, RH, Enright, RD and Cardis, PA, 1995. Forgiveness education with parentally love-deprived late adolescents. *Journal of Moral Education*, 24(4), pp.427–44.

Anderson, ML, 2008. Multiple inference and gender differences in the effects of early intervention: a reevaluation of the Abecedarian, Perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484), pp.1481–95.

Arvay, MJ, 2001. Secondary traumatic stress among trauma counsellors: what does the research say? *International Journal for the Advancement of Counselling*, 23(4), pp.283–93.

Asmal, K, Asmal, L and Roberts, RS, 1994. *Reconciliation through truth: a reckoning of apartheid's criminal governance*. Cambridge: Cambridge University Press.

Backer, DA, 2004. The human face of justice: victims' responses to South Africa's Truth and Reconciliation Commission process. PhD dissertation, University of Michigan.

Backer, DA, 2004. Victims' responses to truth commissions: evidence from South Africa. In: M. Ndulo, ed. 2007. *Security, reconstruction, and reconciliation: when the wars end.* London: University College London, pp.165–96.

Baskin, TW and Enright, RD, 2004. Intervention studies on forgiveness: a meta-analysis. *Journal of Counseling and Development*, 82(1), pp.79–90.

Bayer, Christophe Pierre, Klasen, Fionna and Adam, Hubertus, 2007. Association of trauma and PTSD symptoms with openness to reconciliation and feelings of revenge among former Ugandan and Congolese child soldiers. *Jama*, 298(5), pp.555–59.

Beath, A, Christia, F and Enikolopov, R, 2013. Empowering women through development aid: evidence from a field experiment in Afghanistan. *American Political Science Review*, 107(3), pp.540–57.

Benjamini, Yoav and Hochberg, Yosef, 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), pp. 289–300.

Betancourt, TS, 2010. A longitudinal study of psychosocial adjustment and community reintegration among former child soldiers in Sierra Leone. *International Psychiatry: Bulletin of the Board of International Affairs of the Royal College of Psychiatrists*, 7(3), p.60.

Betancourt, TS, Borisova, II, de la Soudière, M and Williamson, J, 2011. Sierra Leone's child soldiers: war exposures and mental health problems by gender. *Journal of Adolescent Health*, 49(1), pp.21–28.

Betancourt, TS, Borisova, II, Williams, TP, Brennan, RT, Whitfield, TH, de la Soudière, M, Williamson, J and Gilman, SE, 2010. Sierra Leone's former child soldiers: a follow-up study of psychosocial adjustment and community reintegration. *Child Development*, 81(4), pp.1077–95.

Biggar, N, 2003. *Burying the past: making peace and doing justice after civil conflict*. Washington, DC: Georgetown University Press.

Bisson, J and Andrew, M, 2009. Psychological treatment of post-traumatic stress disorder (PTSD) (Review). *The Cochrane Library*, 1.

Blattman, C and Annan, J, 2010. The consequences of child soldiering. *The Review of Economics and Statistics*, 92(4), pp.882–98.

Blattman, C, Hartman, AC and Blair, RA, 2014. How to promote order and property rights under weak rule of law? An experiment in changing dispute resolution behavior through community education. *American Political Science Review*, 108, pp.100–20.

Bliese, PD, Wright, KM, Adler, AB, Cabrera, O, Castrol, CA and Hoge, CW, 2008. Validating the primary care posttraumatic stress disorder screen and the posttraumatic stress disorder checklist with soldiers returning from combat. *Journal of Consulting and Clinical Psychology*, 76(2), p.272.

Brahm, E, 2005. Patterns of truth: examining truth commission impact in cross-national context. Paper presented at the annual meeting of the International Studies Association, Honolulu, HI.

Brounéus, K, 2008. Truth-telling as talking cure? Insecurity and retraumatization in the Rwandan gacaca courts. *Security Dialogue*, 39(1), pp.55–76.

Brounéus, K, 2010. The trauma of truth telling: effects of witnessing in the Rwandan gacaca courts on psychological health. *Journal of Conflict Resolution*, 54(3), pp.408–37.

Bruhn, M and McKenzie, D, 2009. In pursuit of balance: randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1(4), pp.200–32.

Casey, K, Glennerster, R and Miguel, E, 2012. Reshaping institutions: evidence on aid impacts using a preanalysis plan, *The Quarterly Journal of Economics*, 127(4), pp.1755–812.

Cobban, H, 2002. The legacies of collective violence. *Boston Review*, April/May.

Conibere, R, Asher, J, Cibelli, K, Dudukovich, J, Kaplan, R and Ball, P, 2004. *Statistical appendix to the report of the Truth and Reconciliation Commission of Sierra Leone*. Palo Alto, CA: Benetech Human Rights Data Analysis Group.

Coyle, CT and Enright, RD, 1997. Forgiveness intervention with postabortion men. *Journal of Consulting and Clinical Psychology*, 65(6), pp.1042–46.

Cunningham, M, 2003. Impact of trauma work on social work clinicians: empirical findings. *Social Work*, 48(4), pp.451–59.

Curran-Everett, Douglas and Benos, Dale J, 2004. Guidelines for reporting statistics in journals published by the American Physiological Society. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 287(2), pp.L259–L261.

Curran-Everett, Douglas, 2000. Multiple comparisons: philosophies and illustrations. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 279(1), pp.R1–R8.

Derogatis, LR, Lipman, RS, Rickels, K, Uhlenhuth, EH and Covi, L, 1974. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. *Behavioral Science*, 19(1), pp.1–15.

Enright, RD and Fitzgibbons, RP, 2000. *Helping clients forgive: an empirical guide for resolving anger and restoring hope.* Washington, DC: American Psychological Association.

Enright, RD, Knutson Enright, JA, Holter, A, Baskin, T and Knutson, C, 2007. Waging peace through forgiveness in Belfast, Northern Ireland II: educational programs for mental health improvement of children. *Journal of Research in Education*, 17, pp.63–78.

Fearon, JD, Humphreys, M and Weinstein, J, 2009. Can development aid contribute to social cohesion after civil war? Evidence from a field experiment in post-conflict Liberia. *American Economic Review*, 99(2), pp.287–91.

Fearon, JD, Humphreys, M and Weinstein, J, 2015. How does development assistance affect collective action capacity? Results from a field experiment in post-conflict Liberia. *American Political Science Review*, 109(3), pp.450–69.

Foa, EB, 2006. Psychosocial therapy for posttraumatic stress disorder. *Journal of Clinical Psychiatry*, 67, pp.40–45.

Foa, EB and Tolin, DF, 2000. Comparison of the PTSD symptom scale–interview version and the clinician-administered PTSD scale. *Journal of Traumatic Stress*, 13(2), pp.181–91.

Foa, EB, Cashman, L, Jaycox, L and Perry, K, 1997. The validation of a self-report measure of posttraumatic stress disorder: the Posttraumatic Diagnostic Scale. *Psychological Assessment*, 9(4), p.445.

Foa, EB, Riggs, D, Dancu, C and Rothbaum, B, 1993. Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *Journal of Traumatic Stress,* 6(4), pp.459–74.

Freedman, SR and Enright, RD, 1996. Forgiveness as an intervention goal with incest survivors. *Journal of Consulting and Clinical Psychology*, 64(5), pp.983–92.

Frison, L and Pocock, SJ, 1992. Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Statistics in Medicine*, 11(13), pp.1685–704.

Gibson, JL, 2004. Does truth lead to reconciliation? Testing the causal assumptions of the South African truth and reconciliation process. *American Journal of Political Science*, 48(1), pp.201–17.

Hamber, R, 2003. Healing. In: D Bloomfield, T Barnes and L Huyse, eds. *Reconciliation after violent conflict: a handbook.* Stockholm: International Institute for Democracy and Electoral Assistance.

Hawkins, SS and Radcliffe, J, 2006. Current measures of PTSD for children and adolescents. *Journal of Pediatric Psychology*, 31(4), pp.420–30.

Hayner, PB, 2000. *Unspeakable truths: confronting state terror and atrocity*. New York: Routledge.

Humphreys, M, Sanchez de la Sierra, R and van der Windt, P, 2013. Fishing, commitment, and communication: a proposal for comprehensive nonbinding research registration. *Political Analysis*, 21(1), pp.1–20.

Joseph, JS and Gray, M, 2008. Exposure therapy for posttraumatic stress disorder. *The Journal of Behavior Analysis of Offender and Victim Treatment and Prevention*, 1(4), pp.69–79.

Kaaya, SF, Fawzi, MC, Mbwambo, JK, Lee, B, Msamanga, GI and Fawzi W, 2002. Validity of the Hopkins Symptom Checklist-25 amongst HIV-positive pregnant women in Tanzania. *Acta Psychiatrica Scandinavica*, 106(1), pp.9–19.

Kaldor, M and Vincent, J, 2006. *Evaluation of UNDP assistance to conflict-affected countries: case study Sierra Leone*. United Nations Development Programme Evaluation Office.

Kaminer, D, Stein, DJ, Mbanga, I and Zungu-Dirwayi, N, 2001. The Truth and Reconciliation Commission in South Africa: relation to psychiatric status and forgiveness among survivors of human rights abuses. *British Journal of Psychiatry*, 178, pp.373–77.

Kling, JR and Liebman, JB, 2004. *Experimental analysis of neighborhood effects on youth*. Cambridge, MA: John F. Kennedy School of Government, Harvard University.

Kling, JR, Liebman, JB and Katz, LF, 2007. Experimental analysis of neighborhood effects. *Econometrica*, 75(1), pp.83–119.

Knight, RG, Waal-Manning, HJ and Spears, GF, 1983. Some norms and reliability data for the State–Trait Anxiety Inventory and the Zung Self-Rating Depression scale. *British Journal of Clinical Psychology*, 22(4), pp.245–49.

Kritz, NJ, 1999. War crime trials: who should conduct them? In: B Cooper, ed. *War crimes: the legacy of Nuremberg.* New York: TV Books.

Lederach, JP, 1999. *The journey towards reconciliation*. Pennsylvania: Herald Press.

Lin, WF, Mack, D, Enright, RD, Krahn, D and Baskin, TW, 2004. Effects of forgiveness therapy on anger, mood, and vulnerability to substance use among inpatient substance-dependent clients. *Journal of Consulting and Clinical Psychology*, 72(6), pp.1114–21.

Mansuri, G and Rao, V, 2013. *Localizing development: does participation work?* Washington, DC: The World Bank.

McCann, IL and Pearlman, LA, 1990. Vicarious traumatization: a framework for understanding the psychological effects of working with victims. *Journal of Traumatic Stress*, 3(1), pp.131–49.

McKenzie, D, 2012. Beyond baseline and follow-up: the case for more T in experiments. *Journal of Development Economics*, 99(2), pp.210–21.

Mendeloff, D, 2004. Truth-seeking, truth-telling, and postconflict peacebuilding: curb the enthusiasm? *International Studies Review*, 6(3), pp.355–80.

Mendeloff, D, 2009. Trauma and vengeance: assessing the psychological and emotional effects of post-conflict justice. *Human Rights Quarterly*, 31(3), pp.592–623.

Murphy, WP, 1980. Secret knowledge as property and power in Kpelle society: elders versus youth. *Africa: Journal of the International African Institute*, 50(2), pp.193–207.

Murray, LK, Bass, J, Chomba, E, Imasiku, M, Thea, D, Semrau, K, Cohen, JA, Lam, C and Bolton, P, 2011. Validation of the UCLA Child Post traumatic stress disorder-reaction index in Zambia. *International Journal of Mental Health Systems*, 5(1).

Norris, Fran H and Hamblen, Jessica L, 2004. Standardized self-report measures of civilian trauma and PTSD. In: JP Wilson and TM Keane, eds. *Assessing psychological trauma and PTSD*, 2nd ed. New York: Guilford Press, pp. 63–102.

Paluck, EL, 2009. Reducing intergroup prejudice and conflict using the media: a field experiment in Rwanda. *Journal of Personality and Social Psychology*, 96(3), pp.574–87.

Pearlman, LA and Saakvitne, KW, 1995. *Trauma and the therapist: countertransference and vicarious traumatization in psychotherapy with incest survivors*. New York: WW Norton & Co.

Pettersson, T and Wallensteen, P, 2015. Armed conflicts, 1946–2014. *Journal of Peace Research*, 52(4), pp.536–50.

Putnam, RD, 1995. Bowling alone: America's declining social capital. *Journal of Democracy*, 6(1), pp.65–78.

Quinn, J, 2010. *The politics of acknowledgement: truth commissions in Uganda and Haiti*. Vancouver: UBC Press.

Reed, GL and Enright, RD, 2006. The effects of forgiveness therapy on depression, anxiety, and posttraumatic stress for women after spousal emotional abuse. *Journal of Consulting and Clinical Psychology*, 74(5), pp.920–29.

Rose, S, Bisson, J, Churchill, R and Wessely, S, 2002. Psychological debriefing for preventing post traumatic stress disorder (PTSD). *Cochrane Database of Systematic Reviews*, 2(2).

Rye, MS, 1998. Evaluation of a secular and a religiously integrated forgiveness group therapy program for college students who have been wronged by a romantic partner. Unpublished doctoral dissertation, Bowling Green State University, Bowling Green, OH.

Rye, MS, Pargament, KI, Pan, W, Yingling, DW, Shogren, KA and Ito, M, 2005. Can group interventions facilitate forgiveness of an ex-spouse? A randomized clinical trial. *Journal of Consulting and Clinical Psychology*, 73(5), pp.880–92.

Shaw, R, 2005. *Rethinking truth and reconciliation commissions: lessons from Sierra Leone*. United States Institute of Peace Special Report No. 130. Washington, DC: USIP Press.

Smith, LA, Gambette, C and Longley, T, 2004. *Conflict mapping in Sierra Leone: violations of international humanitarian law from 1991 to 2002: executive summary*. Brussels: No Peace Without Justice.

Staub, E, Pearlman, LA, Gubin, A and Hagengimana, A, 2005. Healing, reconciliation, forgiving and the prevention of violence after genocide or mass killing: an intervention and its experimental evaluation in Rwanda. *Journal of Social and Clinical Psychology*, 24(3), pp.297–334.

Steinberg, AM, Brymer, MJ, Decker, KB and Pynoos RS, 2004. The University of California at Los Angeles Post-traumatic Stress Disorder Reaction Index. *Current Psychiatry Reports*, 6(2), pp.96–100.

Subkoviak, MJ, Enright, RD, Wu, C-R, Gassin, EA, Freedman, S, Olson, LM and Sarinopoulos, I, 1995. Measuring interpersonal forgiveness in late adolescence and middle adulthood. *Journal of Adolescence*, 18(6), pp.641–55.

Sullivan, GM and Artino, A, 2013. Analyzing and interpreting data from Likert-type scales. *Journal of Graduate Medical Education*, 5(4), pp.541–42.

Truth and Reconciliation Commission, 1998. *Truth and Reconciliation Commission of South Africa Report.* Cape Town: South African Government.

van Emmerik, AA, Kamphuis, JH, Hulsbosch, AM and Emmelkamp, PM, 2002. Single session debriefing after psychological trauma: a meta- analysis, *The Lancet*, 360(9335), pp.766–71.

Ventureyra, VA, Yao, SN, Cottraux, J, Note, I and De May-Guillard, C, 2002. The validation of the Posttrauamtic Stress Disorder Checklist Scale in posttraumatic stress disorder and nonclinical subjects. *Psychotherapy and Psychosomatics*, 71(1), pp.47–53.

Weathers, FW, Huska, JA and Keane, TM, 1991. PCL-M for DSM-IV. Boston, MA: National Center for PTSD—Behavioral Science Division.

Zung, WW, 1971. A rating instrument for anxiety disorders. *Psychosomatics: Journal of Consultation Liaison Psychiatry*, 12(6), pp.371–79.

Zung, WW, Richards, CB and Short, MJ, 1965. Self-rating depression scale in an outpatient clinic: further validation of the SDS. *Archives of General Psychiatry*, 13(6), pp.508–15.

Other publications in the 3ie Impact Evaluation Report Series

The following reports are available from http://www.3ieimpact.org/evidence-hub/impact-evaluation-repository

Impact evaluation of the Menabe and Melaky development programme in Madagascar, 3ie Impact Evaluation Report 74. Ring, H, Morey, M, Kavanagh, E, Kamto, K, McCarthy, N, Brubaker, J and Rakotondrafara, C, 2018.

Impact evaluation of the Smallholder Dairy Commercialization Programme in Kenya, 3ie Impact Evaluation Report 73. Bonilla, J, McCarthy, N, Mugatha, S, Rai, N, Coombes, A and Brubaker, J, 2018.

Impact and adoption of risk-reducing drought-tolerant rice in India, 3ie Impact Evaluation Report 72. Yamano, T, Dar, MH, Panda, A, Gupta, I, Malabayabas, ML and Kelly, E, 2018.

Poverty and empowerment impacts of the Bihar Rural Livelihoods Project in India, 3ie Impact Evaluation Report 71. Hoffmann, V, Rao, V, Datta, U, Sanyal, P, Surendra, V and Majumdar, S 2018.

How should Tanzania use its natural gas? Citizens' views from a nationwide Deliberative Poll, 3ie Impact Evaluation Report 70. Birdsall, N, Fishkin, J, Haqqi, F, Kinyondo, A, Moyo, M, Richmond, J and Sandefur, J, 2018.

Impact evaluation of the conditional cash transfer program for secondary school attendance in Macedonia, 3ie Impact Evaluation Report 69. Armand, A and Carneiro, P, 2018.

Age at marriage, women's education, and mother and child outcomes in Bangladesh, 3ie Impact Evaluation Report 68. Field, E, Glennerster, R, Nazneen, S, Pimkina, S, Sen, I and Buchmann, N, 2018.

Evaluating agricultural information dissemination in western Kenya, *3ie Impact Evaluation Report 67.* Fabregas, R, Kremer, M, Robinson, J and Schilbach, F, 2017.

General equilibrium impact assessment of the Productive Safety Net Program in Ethiopia, 3ie Impact Evaluation Report 66. Filipski, M, Taylor, JE, Abegaz, GA, Ferede, T, Taffesse, AS and Diao, X, 2017.

Impact of the Uddeepan programme on child health and nutrition in India, 3ie Impact Evaluation Report 65. Kochar, A, Sharma, A and Sharma, A, 2017.

Evaluating oral HIV self-testing to increase HIV testing uptake among truck drivers in Kenya, 3ie Impact Evaluation Report 64. Kelvin, EA, Mwai, E, Romo, ML, George, G, Govender, K, Mantell, JE, Strauss, M, Nyaga, EN and Odhiambo, JO, 2017.

Integration of EPI and paediatric HIV services for improved ART initiation in Zimbabwe, *3ie Impact Evaluation Report 63.* Prescott, M, Boeke, C, Gotora, T, Mafaune, HW, Motsi, W, Graves, J, Mangwiro, A and McCarthy, E, 2017.

Increasing male partner HIV testing using self-test kits in Kenya, 3ie Impact Evaluation Report 62. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Evaluating the impact of community health worker integration into prevention of mother-to-child transmission of HIV services in Tanzania, 3ie Impact Evaluation Report 61. Nance, N, McCoy, S, Ngilangwa, D, Masanja, J, Njau, P and Noronha, R, 2017.

Using HIV self-testing to promote male partner and couples testing in Kenya, 3ie Impact Evaluation Report 60. Thirumurthy, H, Omanga, E, Obonyo, B, Masters, S and Agot, K, 2017.

Increasing male partner HIV self-testing at antenatal care clinics in Kenya, 3ie Impact Evaluation Report 59. Gichangi, A, Korte, JE, Wambua, J, Vrana, C and Stevens, D, 2017.

Impact of free availability of public childcare on labour supply and child development in Brazil, *3ie Impact Evaluation Report 58*. Attanasio, O, Paes de Barros, R, Carneiro, P, Evans, D, Lima, L, Olinto, P and Schady, N, 2017.

Estimating the effects of a low-cost early stimulation and parenting education programme in Mexico, 3ie Impact Evaluation Report 57. Cardenas, S, Evans, D and Holland, P, 2017.

The Better Obstetrics in Rural Nigeria study: an impact evaluation of the Nigerian Midwives Service Scheme, 3ie Impact Evaluation Report 56. Okeke, E, Glick, P, Abubakar, IS, Chari, AV, Pitchforth, E, Exley, J, Bashir, U, Setodji, C, Gu, K and Onwujekwe, O, 2017.

*The Productive Safety Net Programme in Ethiopia: impacts on children's schooling, labour and nutritional statu*s, *3ie Impact Evaluation Report 55.* Berhane, G, Hoddinott, J, Kumar, N and Margolies, A, 2016.

The impact of youth skills training on the financial behaviour, employability and educational choice in Morocco, 3ie Impact Evaluation Report 54. Bausch, J, Dyer, P, Gardiner, D, Kluve, J and Mizrokhi, E, 2016.

Using advertisements to create demand for voluntary medical male circumcision in South Africa, 3ie Impact Evaluation Report 53. Frade, S, Friedman, W, Rech, D and Wilson, N, 2016.

The use of peer referral incentives to increase demand for voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 52. Zanolini, A, Bolton, C, Lyabola, LL, Phiri, G, Samona, A, Kaonga, A and Harsha Thirumurthy, H, 2016.

Using smartphone raffles to increase demand for voluntary medical male circumcision in Tanzania, 3ie Impact Evaluation Report 51. Mahler, H and Bazant, E, 2016.

Voluntary medical male circumcision uptake through soccer in Zimbabwe, 3ie Impact Evaluation Report 50. DeCelles, J, Kaufman, Z, Bhauti, K, Hershow, R, Weiss, H, Chaibva, C, Moyo, N, Braunschweig, E, Mantula, F, Hatzold, K and Ross, D, 2016.

Measuring the impact of SMS-based interventions on uptake of voluntary medical male circumcision in Zambia, 3ie Impact Evaluation Report 49. Leiby, K, Connor, A, Tsague, L, Sapele, C, Koanga, A, Kakaire, J and Wang, P, 2016.

Assessing the impact of delivering messages through intimate partners to create demand for voluntary medical male circumcision in Uganda, 3ie Impact Evaluation Report 48. Semeere, AS, Bbaale, DS, Castelnuovo, B, Kiragga, A, Kigozi, J, Muganzi, A, Kambugu, A and Coutinho, AG, 2016.

Optimising the use of economic interventions to increase demand for voluntary medical male circumcision in Kenya, 3ie Impact Evaluation Report 47. Thirumurthy, H, Omanga, E, Rao, SO, Murray, K, Masters, S and Agot, K, 2016.

The impact of earned and windfall cash transfers on livelihoods and conservation in Sierra Leone, 3ie Impact Evaluation Report 46. Bulte, E, Conteh, B, Kontoleon, A, List, J, Mokuwa, E, Richards, P, Turley, T and Voors, M, 2016.

Property tax experiment in Pakistan: Incentivising tax collection and improving performance, 3ie Impact Evaluation Report 45. Khan, A, Khwaja, A and Olken, B, 2016.

Impact of mobile message reminders on tuberculosis treatment outcomes in Pakistan, 3ie Impact Evaluation Report 44. Mohammed, S, Glennerster, R and Khan, A, 2016.

Making networks work for policy: Evidence from agricultural technology adoption in Malawi, 3ie Impact Evaluation Report 43. Beaman, L, BenYishay, A, Fatch, P, Magruder, J and Mobarak, AM, 2016.

Estimating the impact and cost-effectiveness of expanding access to secondary education in Ghana, 3ie Impact Evaluation Report 42. Dupas, P, Duflo, E and Kremer, M, 2016.

Evaluating the effectiveness of computers as tutors in China, 3ie Impact Evaluation Report 41. Mo, D, Bai, Y, Boswell, M and Rozelle, S, 2016.

Micro entrepreneurship support programme in Chile, 3ie Impact Evaluation Report 40. Martínez, CA, Puentes, EE and Ruiz-Tagle, JV, 2016.

Thirty-five years later: evaluating the impacts of a child health and family planning programme in Bangladesh, 3ie Impact Evaluation Report 39. Barham, T, Kuhn, R, Menken, J and Razzaque, A, 2016.

Effectiveness of a rural sanitation programme on diarrhoea, soil-transmitted helminth infection and malnutrition in India, 3ie Impact Evaluation Report 38. Clasen, T, Boisson, S, Routray, P, Torondel, B, Bell, M, Cumming, O, Ensink, J, Freeman, M and Jenkins, M, 2016.

Evaluating the impact of vocational education vouchers on out-of-school youth in Kenya, 3ie Impact Evaluation Report 37. Hicks, JH, Kremer, M, Mbiti, I and Miguel, E, 2016.

Removing barriers to higher education in Chile: evaluation of peer effects and scholarships for test preparation, 3ie Impact Evaluation Report 36. Banerjee, A, Duflo E and Gallego, F, 2016.

Sustainability of impact: dimensions of decline and persistence in adopting a biofortified crop in Uganda, 3ie Impact Evaluation Report 35. McNiven, S, Gilligan, DO and Hotz, C 2016.

A triple win? The impact of Tanzania's Joint Forest Management programme on livelihoods, governance and forests, *3ie Impact Evaluation Report 34.* Persha, L and Meshack, C, 2016.

The effect of conditional transfers on intimate partner violence: evidence from Northern Ecuador, 3ie Impact Evaluation Report 33. Hidrobo, M, Peterman, A and Heise, L, 2016.

The effect of transfers and preschool on children's cognitive development in Uganda, 3ie Impact Evaluation Report 32. Gillian, DO and Roy, S, 2016.

Can egovernance reduce capture of public programmes? Experimental evidence from India's employment guarantee, 3ie Impact Evaluation Report 31. Banerjee, A, Duflo, E, Imbert, C, Mathew, S and Pande, R, 2015.

Improving maternal and child health in India: evaluating demand and supply strategies, 3ie Impact Evaluation Report 30. Mohanan, M, Miller, G, Forgia, GL, Shekhar, S and Singh, K, 2016.

Smallholder access to weather securities in India: demand and impact on production decisions, 3ie Impact Evaluation Report 28. Ceballos, F, Manuel, I, Robles, M and Butler, A, 2015.

What happens once the intervention ends? The medium-term impacts of a cash transfer programme in Malawi, 3ie Impact Evaluation Report 27. Baird, S, Chirwa, E, McIntosh, C and Özler, B, 2015.

Validation of hearing screening procedures in Ecuadorian schools, 3ie Impact Evaluation Report 26. Muñoz, K, White, K, Callow-Heusser, C and Ortiz, E, 2015.

Assessing the impact of farmer field schools on fertilizer use in China, 3ie Impact Evaluation Report 25. Burger, N, Fu, M, Gu, K, Jia, X, Kumar, KB and Mingliang, G, 2015.

The SASA! study: a cluster randomised trial to assess the impact of a violence and HIV prevention programme in Kampala, Uganda, *3ie Impact Evaluation Report 24.* Watts, C, Devries, K, Kiss, L, Abramsky, T, Kyegombe, N and Michau, L, 2014.

Enhancing food production and food security through improved inputs: an evaluation of Tanzania's National Agricultural Input Voucher Scheme with a focus on gender impacts, 3ie Impact Evaluation Report 23. Gine, X, Patel, S, Cuellar-Martinez, C, McCoy, S and Lauren, R, 2015.

A wide angle view of learning: evaluation of the CCE and LEP programmes in Haryana, 3ie Impact Evaluation Report 22. Duflo, E, Berry, J, Mukerji, S and Shotland, M, 2015.

Shelter from the storm: upgrading housing infrastructure in Latin American slums, *3ie Impact Evaluation Report 21.* Galiani, S, Gertler, P, Cooper, R, Martinez, S, Ross, A and Undurraga, R, 2015.

Environmental and socioeconomic impacts of Mexico's payments for ecosystem services programme, *3ie Impact Evaluation Report 20.* Alix-Garcia, J, Aronson, G, Radeloff, V, Ramirez-Reyes, C, Shapiro, E, Sims, K and Yañez-Pagans, P, 2015.

A randomised evaluation of the effects of an agricultural insurance programme on rural households' behaviour: evidence from China, 3ie Impact Evaluation Report 19. Cai, J, de Janvry, A and Sadoulet, E, 2014.

Impact of malaria control and enhanced literacy instruction on educational outcomes among school children in Kenya: a multi-sectoral, prospective, randomised evaluation, 3ie Impact Evaluation Report 18. Brooker, S and Halliday, K, 2015.

Assessing long-term impacts of conditional cash transfers on children and young adults in rural Nicaragua, *3ie Impact Evaluation Report 17.* Barham, T, Macours, K, Maluccio, JA, Regalia, F, Aguilera, V and Moncada, ME, 2014.

The impact of mother literacy and participation programmes on child learning: evidence from a randomised evaluation in India, 3ie Impact Evaluation Report 16. Banerji, R, Berry, J and Shortland, M, 2014.

A youth wage subsidy experiment for South Africa, 3ie Impact Evaluation Report 15. Levinsohn, J, Rankin, N, Roberts, G and Schöer, V, 2014.

Providing collateral and improving product market access for smallholder farmers: a randomised evaluation of inventory credit in Sierra Leone, 3ie Impact Evaluation Report 14. Casaburi, L, Glennerster, R, Suri, T and Kamara, S, 2014.

Scaling up male circumcision service provision: results from a randomised evaluation in Malawi, 3ie Impact Evaluation Report 13. Thornton, R, Chinkhumba, J, Godlonton, S and Pierotti, R, 2014.

Targeting the poor: evidence from a field experiment in Indonesia, 3ie Impact Evaluation Report 12. Atlas, V, Banerjee, A, Hanna, R, Olken, B, Wai-poi, M and Purnamasari, R, 2014.

An impact evaluation of information disclosure on elected representatives' performance: evidence from rural and urban India, 3ie Impact Evaluation Report 11. Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mahapatra, B, 2014.

Truth-telling by third-party audits and the response of polluting firms: Experimental evidence from India, 3ie Impact Evaluation Report 10. Duflo, E, Greenstone, M, Pande, R and Ryan, N, 2013.

No margin, no mission? Evaluating the role of incentives in the distribution of public goods in Zambia, *3ie Impact Evaluation Report 9.* Ashraf, N, Bandiera, O and Jack, K, 2013.

Paying for performance in China's battle against anaemia, *3ie Impact Evaluation Report 8.* Zhang, L, Rozelle, S and Shi, Y, 2013.

Social and economic impacts of Tuungane: final report on the effects of a community-driven reconstruction programme in the Democratic Republic of Congo, *3ie Impact Evaluation Report 7*. Humphreys, M, Sanchez de la Sierra, R and van der Windt, P, 2013.

The impact of daycare on maternal labour supply and child development in Mexico, *3ie Impact Evaluation Report 6*. Angeles, G, Gadsden, P, Galiani, S, Gertler, P, Herrera, A, Kariger, P and Seira, E, 2014.

Impact evaluation of the non-contributory social pension programme 70 y más in Mexico, *3ie Impact Evaluation Report 5*. Rodríguez, A, Espinoza, B, Tamayo, K, Pereda, P, Góngora, V, Tagliaferro, G and Solís, M, 2014.

Does marginal cost pricing of electricity affect groundwater pumping behaviour of farmers? Evidence from India, 3ie Impact Evaluation Report 4. Meenakshi, JV, Banerji, A, Mukherji, A and Gupta, A, 2013.

The GoBifo project evaluation report: Assessing the impacts of community-driven development in Sierra Leone, *3ie Impact Evaluation Report 3*. Casey, K, Glennerster, R and Miguel, E, 2013.

A rapid assessment randomised-controlled trial of improved cookstoves in rural Ghana, *3ie Impact Evaluation Report 2.* Burwen, J and Levine, DI, 2012.

The promise of preschool in Africa: A randomised impact evaluation of early childhood development in rural Mozambique, *3ie Impact Evaluation Report 1.* Martinez, S, Naudeau, S and Pereira, V, 2012.

This evaluation examines the impact of a reconciliation programme implemented in post-conflict Sierra Leone. Forums were set up with a group of villages where victims detailed the atrocities faced in war and perpetrators confessed to war crimes and sought forgiveness. The study finds that the programme led to greater forgiveness of those who perpetrated violence during the civil war. However, the reconciliation process also affected the psychological well-being of the participants, the effects of which persisted nearly three years after the intervention.

Impact Evaluation Series

International Initiative for Impact Evaluation
202-203, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

www.3ieimpact.org