

## RESEARCH ARTICLE

## PSYCHOLOGY

# Reconciling after civil conflict increases social capital but decreases individual well-being

Jacobus Cilliers,<sup>1</sup> Oeindrila Dube,<sup>2\*</sup> Bilal Siddiqi<sup>3</sup>

Civil wars divide nations along social, economic, and political cleavages, often pitting one neighbor against another. To restore social cohesion, many countries undertake truth and reconciliation efforts. We examined the consequences of one such effort in Sierra Leone, designed and implemented by a Sierra Leonean nongovernmental organization called Fambul Tok. As a part of this effort, community-level forums are set up in which victims detail war atrocities, and perpetrators confess to war crimes. We used random assignment to study its impact across 200 villages, drawing on data from 2383 individuals. We found that reconciliation had both positive and negative consequences. It led to greater forgiveness of perpetrators and strengthened social capital: Social networks were larger, and people contributed more to public goods in treated villages. However, these benefits came at a substantial cost: The reconciliation treatment also worsened psychological health, increasing depression, anxiety, and posttraumatic stress disorder in these same villages. For a subset of villages, we measured outcomes both 9 months and 31 months after the intervention. These results show that the effects, both positive and negative, persisted into the longer time horizon. Our findings suggest that policy-makers need to restructure reconciliation processes in ways that reduce their negative psychological costs while retaining their positive societal benefits.

Most wars today are civil wars (1), which divide countries along ethnic, economic, and political cleavages. For example, the Hutus targeted the Tutsis during Rwanda's genocide (in 1994), and illicit diamonds sustained Sierra Leone's civil war (over 1991–2002), pitting one neighbor against another. Because conflicts like this sever social ties among individuals, their prevalence has spurred efforts to promote social cohesion and improve social capital as a part of postconflict recovery (2–7).

Truth and reconciliation processes are a common approach used around the world to promote this type of rebuilding (8). These processes are founded on the idea that airing wartime grievances is the key to restoring social ties. As such, they bring war victims face-to-face with perpetrators through forums in which victims describe war atrocities and perpetrators confess to war crimes without facing prosecution. Proponents of this approach claim that reconciliation processes are highly effective—not just in rebuilding social capital and promoting societal healing but also in providing psychological relief to participants, aiding individual healing (9–15). Yet, we have little

knowledge of whether and how reconciliation processes help communities heal from conflict.

We have some evidence from past work that attitudes toward other groups can improve in the aftermath of nationwide Truth and Reconciliation Commissions (TRCs) (16) and with exposure to trauma counseling (17). Also, other types of interventions targeted toward individuals have been shown to reduce prejudice (18) and improve day-to-day dispute resolution (19). But what happens when we induce targeted, person-to-person forgiveness throughout a community? We lack rigorous evidence on how community-wide reconciliation influences either individual or societal healing (20, 21).

Our study seeks to address this gap in the literature. We conducted a randomized control trial of a reconciliation process in Sierra Leone that was designed and implemented by a Sierra Leonean nongovernmental organization (NGO) called Fambul Tok. Fambul Tok's intervention has several features common to truth and reconciliation processes around the world: It initiates forums in which victims describe the violence they experienced and perpetrators seek forgiveness for their crimes. Also, no one receives monetary compensation or is punished for participating. However, Fambul Tok's approach is distinct from nationwide truth and reconciliation because it conducts community-level reconciliation, holding forums at the level of the section, which on average includes 10 villages. We used random assignment to evaluate the impact of its work across 100 sections of Sierra Leone. Our evaluation was independent, and

we provided no input into the design of its program.

## War and reconciliation in Sierra Leone

More than 50,000 people were killed during Sierra Leone's civil war. Thousands more were raped and had limbs amputated, and 2.6 million people—more than half the population of ~4 million people (22)—were displaced as a part of the Revolutionary United Front (RUF) rebel group's campaign of terror against the population.

Much of the violence was neighbor-on-neighbor and took place among members of the same village. Child soldiers were frequently recruited by the RUF. Sometimes, they willingly rose up against local authorities in their village, and at other times, they were forced to commit atrocities against fellow villagers. The other armed actors in the conflict included the Sierra Leonean Army (SLA) and local militias called the Civil Defense Forces (CDF), which emerged in response to widespread civilian abuses and came to be revered for protecting the population against the rebels. Although all armed actors inflicted civilian casualties, the vast majority of the atrocities were committed by the RUF (23, 24).

After the conflict, the Sierra Leonean government set up a national TRC, but it only had the capacity to cover a small fraction of the war atrocities. Also, many rural Sierra Leoneans were unable to access the district capitals where the forums were held. As a result, a large part of the population was left out of the national reconciliation process.

Fambul Tok ("Family Talk" in Krio) was founded to address this gap in 2007, when it began initiating community-level reconciliation forums. As a part of its program, committees composed of community members were trained in trauma healing and mediation and conducted outreach to encourage victims and perpetrators to participate in the truth-telling process. This culminated in a 2-day bonfire ceremony in which victims described their experiences and perpetrators asked for forgiveness. The ceremonies were relatively cheap, costing between \$150 and \$200 in total, for all participants. They also incorporated traditional rituals to promote community healing. After the ceremony, Fambul Tok set up a symbolic Peace Tree in each village and, in some areas, communal farms to further sustain community healing. It additionally helped establish a Peace Mothers' group to discuss gender-targeted atrocities perpetrated during the war. As such, this intervention could have some impacts other than reconciliation—for example, on economic activity. Where we discuss alternative accounts, we lay out why effects on psychological health and social capital are likely due to reconciliation rather than these other impacts. (In supplementary text S3, we also discuss how local-level reconciliation processes such as the one implemented by Fambul Tok compare with national-level reconciliation processes.)

## Healing through reconciliation

Reconciliation processes such as this one could theoretically have both positive and negative psychological consequences. On the one hand,

<sup>1</sup>McCourt School of Public Policy, Georgetown University, 3700 O Street, Washington, DC 20057, USA. <sup>2</sup>Department of Politics, New York University, 19 West 4th Street, New York, NY 10012, USA. <sup>3</sup>Development Research Group, World Bank, 1818 H Street NW, Washington, DC 20433, USA.

\*Corresponding author. Email: odube@nyu.edu

they may improve psychological health if sharing war accounts has a cathartic effect (10, 11, 22) or leads to forgiveness, which has been shown to improve trauma, anxiety, and depression (25)—particularly when induced in the context of forgiveness therapies (26–31).

On the other hand, they may also prove traumatic because they evoke painful war memories without allowing for gradual habituation or desensitization (32, 33). In this regard, reconciliation processes are similar to single-session debriefing (34), which seeks to counsel patients by exposing them briefly and intensively to traumatic events but has been shown to have limited therapeutic value (34, 35). In contrast, gradual exposure therapy has been shown to be more effective for mitigating posttraumatic stress disorder (PTSD) (32, 36). Negative psychological effects associated with reconciliation need not be concentrated among those who were directly victimized; for example, other community members may experience vicarious traumatization (37–40) as they hear about new atrocities committed during the war.

In fact, studies of those who have testified in national TRCs suggest that this participation produces mixed emotional responses (41–43), may not improve psychological health (44), or may even correlate with worse psychological outcomes (34, 45). It is difficult to infer causal effects by comparing those who testified with those who did not because those who chose to testify may have experienced greater violence exposure or had a different psychological makeup. We use a randomized design to mitigate this type of endogeneity concern and better identify the effect of reconciliation on psychological outcomes, including trauma, anxiety, and depression. We also examined whether the effects vary systematically based on the degree to which individuals experienced war violence.

Reconciliation processes may also affect societal healing through their effects on social capital,

which is conceptualized as social networks, and norms such as trust and reciprocity that arise from these network ties (46). Social capital effects could also arise as a consequence of forgiveness. For example, individuals may stop avoiding places and activities associated with perpetrators and form social ties with them after forgiving them for past actions. They could also arise as a consequence of acknowledgment (47): People may be more willing to contribute to communities that have recognized that they were victimized, or that have recognized that they perpetrated crimes without punishing them for these past actions. To determine impacts on social capital, we examined outcomes such as social networks, participation in community groups, and contributions to public goods.

**Evaluation design**

In 2011, when Fambul Tok was poised to expand into new sections in its five districts of operation (Kailahun, Kono, Bombali, Moyamba, and Koinadugu), we used random assignment to assign some sections to the Fambul Tok treatment group and other sections to serve as a part of the control group. Geographically, sections are units that lie within districts, whereas villages are even smaller units that lie within sections.

In the supplementary materials (figs. S1 to S3), we show that these five districts are similar to other districts in Sierra Leone along key dimensions such as exposure to war violence and other socioeconomic characteristics. These similarities suggest that the findings of the study are also likely to hold for other areas of Sierra Leone, which helps boost the external validity of the study.

The evaluation occurred in waves so as to allow Fambul Tok to work within its capacity. The first wave included 40 sections, and the second wave included 60 sections. Data collection for a third wave was interrupted by the Ebola crisis in Sierra Leone in 2014. Our field staff had to be evacuated while we were collecting behavioral

measures. These 100 sections are also similar in key characteristics to other sections within the districts of study (table S1), which further bolsters potential generalizability to other areas of the country.

Within each section, we sampled two villages: One was the section headquarters, where the reconciliation ceremony was typically held, and the second was randomly chosen among what was on average nine remaining villages. Within each village, we interviewed a random sample of 10 to 12 adults, for a total of 2383 respondents across 200 villages. Almost all of our key outcome variables are individual-level responses from household surveys.

In wave one, we conducted endline surveys both 9 months and 31 months after the ceremonies took place, enabling us to determine both short-run and long-run effects. In wave two, endline surveys were conducted once, ~18 to 19 months after the ceremonies. The evaluation timeline, which spanned the 2011–2014 period, is shown in fig. S4.

For all endline rounds, we sought to resurvey the same respondents interviewed at baseline. We went to great lengths to minimize attrition, with repeat visits and by tracking respondents who had moved to neighboring villages.

The attrition rate of those who appeared in baseline but are missing from either endline round in wave one or the endline in wave two is 13% (315 out of 2382 individuals), and the attrition rate for those missing from both endline rounds in wave one or the endline in wave two is 7% (168 of 2382 individuals). As shown in table S2, neither of these attrition measures—nor the attrition measure of each endline round separately—is predicted by treatment (supplementary text S1).

We also used four village-level variables from a village survey. Because of a mechanical error in the hand-held devices used for data collection, this village-level survey is missing for five villages

**Table 1. The impact of reconciliation on forgiveness and trust.** Each row represents a separate regression of the outcome shown in the first column on treatment assignment. All specifications include section pair fixed effects and the second-round indicator, the baseline outcome variable, and its interaction with both the second-round indicator and the second-wave indicator. SEs 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.

Variables	Control mean	Coefficient	SE	Observations	R <sup>2</sup>
<i>Forgiveness</i>					
Forgive perpetrators	2.264	0.571**	(0.227)	2010	0.131
Forgive perpetrators (based on questions in both baselines)	0.951	0.277*	(0.145)	2085	0.121
<i>Trust</i>					
How much do you trust rebel excombatants?	1.901	0.177**	(0.079)	900	0.222
Indicator: Trust rebel excombatants somewhat or completely	0.328	0.073**	(0.036)	900	0.197
How much do you trust migrants to this community?	3.161	0.123***	(0.033)	2203	0.172
Indicator: Trust migrants somewhat or completely	0.861	0.058***	(0.012)	2203	0.094
Index of generalized trust in community	0	0.006	(0.027)	2996	0.135
<i>Indicators</i>					
People are honest and can be trusted	2.598	0.014	(0.026)	2994	0.126
People in village are honest and can be trusted	2.858	-0.010	(0.020)	2976	0.167
People in community would not betray fellow community members	2.550	0.003	(0.028)	2976	0.059
Money left out accidentally will still be there an hour later	0.365	0.010	(0.020)	2956	0.141

in baseline and a separate six villages in endline. As shown in table S3, whether a village is missing in either baseline or endline is also uncorrelated with treatment. The number of villages and individuals in our sample is shown in table S4, disaggregated by wave and round of data collection. (In supplementary text S1, we discuss additional robustness checks to confirm that missing village-level indicators do not affect our results.)

**Empirical strategy**

We used our baseline survey data to match sections into pairs stratified by district and randomly assigned one section in each pair into treatment and the other into control. The balance on key covariates is reported in table S5, and more details on matching and balance statistics are provided in supplementary text S3. To examine treatment effects, our main specification pools together endline surveys from both waves and rounds of the evaluation. For most outcomes, we have baseline data, which enables us to control for the baseline value of the dependent variable (48, 49). This approach reduces noise and increases power and has been commonly used in recent experimental studies in the social sciences [for example, (50)].

We estimate regressions that can be represented as

$$y_{rivspw} = \beta_0 + \beta_1 T_s + \rho_p + \beta_2 y_{Oivspw} + \delta_r + \delta_r y_{Oivspw} + \lambda_w y_{Oivspw} + \epsilon_{rivspw} \quad (1)$$

where  $y_{Oivspw}$  and  $y_{rivspw}$  denote outcomes at baseline and endline round  $r$ , respectively, for individual  $i$  in village  $v$ , section  $s$ , section-pair  $p$ , and wave  $w$ .  $\rho_p$  denotes section-pair fixed effects, which account for section-level matching in the allocation of treatment (51).  $\delta_r$  is a round effect that equals 1 for the second-round endline. The interaction term,  $\delta_r y_{Oivspw}$ , allows the baseline to exert different effects over time.  $\lambda_w$  is a wave effect that equals 1 for sections in the second wave. Because each wave includes different sections, wave effects are subsumed by section-pair effects.  $\lambda_w y_{Oivspw}$  allows baseline variables to have different effects for the wave-two sections. This control is particularly important because we are only able to include pared-down baseline outcomes collected in the second-wave baseline survey (a point discussed further in the Data section). Last,  $T_s$  is

assignment to treatment, and  $\beta_1$  measures the treatment effect.

If we did not have baseline data for an outcome, we estimated cross-sectional specifications of the form

$$y_{rivspw} = \beta_0 + \beta_1 T_s + \rho_p + \delta_r + \epsilon_{rivspw} \quad (2)$$

We clustered the standard errors (SEs) at the section level, which is the unit of treatment allocation. This accounts for the potential correlation of errors across individuals within a section (and implicitly, within a village, because a section is larger than a village).

There are three sections in which some responses do not match treatment assignment; these sections were assigned to control, and yet six of the respondents in one village and eight respondents in the other two reported attending a bonfire ceremony. However, we used assignment to treatment in estimating all of our specifications. Thus, ceremony participation among control respondents may lead to an understatement of the effect.

Many of our outcomes are mean effect indices that first standardize and then sum various indicators used to measure similar concepts. We used the methodology of (52), which imputes missing values before aggregation. The indicators are standardized by subtracting control group means and dividing by control group standard deviations, so that the control group means for the indices are zero by construction. In supplementary text S3, we provide greater detail on this method, and in table S6, we show robustness to an alternate method that does not first impute missing values (53).

To avoid fishing for significant effects (4, 6), we registered a Pre-Analysis Plan (PAP) in the Evidence in Governance and Politics (EGAP) repository before analysis of any endline data from either wave one or wave two. The PAP outlines the indicators comprising each index and all the hypotheses to be tested. A copy can be found at <http://egap.org/registration/622>. All hypotheses specified in our PAP are listed in table S7. We present results for six of the hypotheses in Tables 1 to 6 and 10 others in the supplementary materials. In supplementary text S2, we discuss the PAP in more detail and also the few circumstances under which we deviated from the prespecified group-

ing, owing to issues aggregating conditional and unconditional outcomes or to changes in how the social network data were collected over rounds.

In addition, we show in tables S8 and S9 that adjusting for multiple comparisons by controlling for rates of false discovery (54–56) does not affect any of our main results (supplementary text S3).

**Data**

In terms of our dependent variables, we used the Rye Forgiveness Scale to construct an index of forgiveness toward former perpetrators (57). This is a sum of 12 questions [and a subset of the 60 questions in the Enright Forgiveness Inventory (58)], answered on a four-point Likert scale, which were administered to those who reported being physically or emotionally hurt during the war. These questions are designed to measure affect as well as cognitive and behavioral responses toward former perpetrators.

The questions in this index are listed in table S10. Whereas all three endline surveys and the first-wave baseline included these 12 questions, the second-wave baseline included a subset of seven questions, which serve as a pared-down baseline control for second-wave observations. However, both indices show high internal consistency: Cronbach’s  $\alpha$  is 0.865 for the full forgiveness index and 0.824 for the pared-down forgiveness index.

To measure trust, we aggregated four questions on perceived trust and honesty of community members into an index of generalized trust. We also asked separate questions on degree of trust toward former RUF rebel combatants (to whom we refer as “rebel ex-combatants,” for brevity) as well as migrants, many of whom are former combatants who left their villages after the war. We also measured trust of former members of the SLA and the CDF (supplementary text S2). These trust questions are based on a 4-point Likert scale (with responses “trust completely,” “trust somewhat,” “distrust somewhat,” and “distrust completely”). In order to aid the interpretation of our results, we also constructed a binary variable indicating whether the respondent trusts the relevant subgroup or not.

To gauge impacts on social networks, we asked respondents to identify people from the 9 to 11 other respondents in that village whom they

Downloaded from <http://science.sciencemag.org/> on May 12, 2016

**Table 2. Reconciliation and social networks.** Each row represents a separate regression of the outcome shown in the first column on treatment assignment. All specifications are cross-sectional because we do not have baseline measures of these dependent variables. All regressions also include section pair fixed effects and the second round indicator. SEs 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.

Variables	Control mean	Coefficient	SE	Observations	R <sup>2</sup>
Index of network strength	0	0.099***	(0.028)	3008	0.061
Indicators					
Number of people respondent would approach for advice / help	2.894	0.148**	(0.069)	3005	0.056
Number of people respondent would ask to collect money for them	3.144	0.155	(0.142)	3005	0.026
Number of times respondent listed as good friend	2.123	0.232**	(0.091)	3008	0.192
Number of times respondent listed as someone to ask for advice / help	3.245	0.362***	(0.126)	3008	0.199

would consider a good friend and would ask for advice and help. We used this to construct a measure of how many times a respondent was named by someone else. We also asked the respondent to list all the people in the village they would ask to collect money for them and ask for help. We standardized and summed these four measures into a mean effect index. We were only able to conduct cross-sectional analyses with these questions because they were asked differently in the baseline and endline surveys (supplementary text S2).

We constructed a community group participation index based on whether respondents were members of organizations such as Parent Teacher Associations (PTAs) and religious groups and whether they attended group meetings. We also constructed an index of public goods contributions based on whether individuals contributed money or labor to community groups or to building public facilities (including bridges, schools, wells, and health clinics), gave money to a family in need, or participated in road-brushing (a common form of road maintenance), as well as the number of community projects in their village.

Turning to psychological health, we measured PTSD using 11 questions from the PTSD Symptom Scale that assesses the presence and severity of PTSD symptoms according to the 4th Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). This scale has been validated for research purposes (59, 60) and shown to have good psychometric properties, including high internal consistency and test-retest reliability (59). We also drew 7 depression and 10 anxiety questions from the Zung Depression and Zung Anxiety indices (61, 62). The second-wave baseline included a subset of seven and five questions on anxiety and depression, respectively, which again form pared-down baseline-dependent variable

controls. The indices for PTSD, anxiety, and depression are sums of questions answered on a four-point Likert scale (all the questions are listed in table S11). We further aggregated these three indices into a mean effect index of psychological health. We inverted the indicators so that a reduction in the index indicates worse psychological health.

The psychometric scales from which we drew our questions have typically been assessed in developed country contexts, which raise questions around whether they are culturally relevant and valid for a developing country such as Sierra Leone. We piloted our survey instruments extensively and adapted the wording of the psychological measures to the Sierra Leonean context so that they better reflect the informality of Krio language. Furthermore, our scales correspond closely to scales used in other recent studies set in postconflict parts of sub-Saharan Africa, where they have demonstrated good psychometric properties. For example, 15 of our 17 questions on anxiety and depression are also a part of the Johns Hopkins 25-Item Checklist for Anxiety and Depression (63). An adapted version of this scale shows strong internal consistency among adults who were formerly child soldiers in Sierra Leone (64–66). Although our PTSD scale has not been applied in Sierra Leone, it uses the same questions as the Child Posttraumatic Stress Disorder Reaction Index (CPTSD-RI) (67, 68), which has been tested on a population of Ugandan and Congolese child soldiers (69). Moreover, the psychological wellbeing questions we used also exhibit high internal consistency in our sample, with a Cronbach's  $\alpha$  ranging between 0.831 and 0.936 (supplementary text S1).

We also converted the continuous PTSD measure into a dichotomous indicator of whether an individual suffers from clinical PTSD or severe

trauma, following guidelines from the Clinician-Administered PTSD Scale (70). This is coded as 1 if the individual showed at least one symptom of reexperience, one symptom of avoidance, and at least two symptoms of increased arousal. We did not prespecify that we would look at this indicator in our PAP but do so to better gauge the magnitude of the effects on PTSD.

In terms of our sample, because the 10 to 12 respondents were randomly chosen in each village, some may have been victims during the war and others perpetrators. Our main results examine average impacts on all respondents. However, we also collected data on the ways in which respondents were exposed to violence to determine whether the treatment effect varies based on victimization. In our PAP, we defined a violence-exposed individual as one who was beaten, raped, maimed, abducted, or saw violence during the war. We discuss alternate measures in supplementary text S5. We also define someone as an ex-combatant based on a self-reported measure and whether they indicated that they were abducted and forced to carry a gun after getting abducted. There is likely to be extensive underreporting in both measures, which means the ex-combatant variable is likely measured with noise.

Descriptive statistics of key variables are presented in table S12. The surveyed respondents reside in impoverished conditions. More than 70% have no formal education, and less than 8% live in a village with a market. They also experienced extensive war violence: 54% had a family member killed, 33% were beaten, 2% report being maimed, and 3% report that they were raped. These latter numbers are also likely to be underestimates, given the sensitivity of these measures.

Respondents in treatment sections were very familiar with Fambul Tok's reconciliation program

**Table 3. Reconciliation and participation in community groups.** 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, 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. SEs 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.

Variables	Control mean	Coefficient	SE	Observations	R <sup>2</sup>
Index of participation in community groups	0	0.058***	(0.017)	3004	0.160
Index of participation in community groups, without women's membership or meetings	0	0.064***	(0.017)	3004	0.162
Indicators					
PTA membership	0.137	0.034**	(0.016)	2732	0.223
Village development committee membership	0.091	0.013	(0.011)	2737	0.141
Youth group membership	0.101	0.015*	(0.008)	2738	0.144
Women's group membership	0.118	0.022	(0.014)	2004	0.138
Secret society membership	0.358	-0.058***	(0.019)	2770	0.338
Religious group membership	0.286	0.055***	(0.020)	2729	0.179
PTA meeting attendance	0.082	0.037**	(0.015)	2739	0.138
Village development committee meeting attendance	0.068	0.008	(0.010)	2734	0.106
Youth group meeting attendance	0.066	0.007	(0.008)	2739	0.090
Women's group meeting attendance	0.075	0.024*	(0.013)	2004	0.095
Secret society meeting attendance	0.056	-0.005	(0.008)	2766	0.057
Religious group meeting attendance	0.190	0.058***	(0.016)	2714	0.103
Community meeting attendance	0.626	0.006	(0.013)	2983	0.077

(table S13), indicating that the intervention was well implemented.

## Results

Our findings on reconciliation and the forgiveness of former perpetrators are presented in Table 1, top. The forgiveness index (of 12 questions) is 0.571 higher in treatment areas ( $SE = 0.227$ ,  $P = 0.013$ ), over the control group mean of 2.264. The pared down forgiveness index (of seven questions) is 0.277 higher ( $SE = 0.145$ ,  $P = 0.059$ ) than that of its control group mean of 0.951.

Because the forgiveness indices are summed on a Likert scale, the coefficients cannot be interpreted in percent terms by comparing them with control group means. Under these scales, changing the value assigned to responses will not alter the regression coefficients but will alter the control group mean, yielding a different implied percent effect. Moreover, no one valuation is necessarily more appropriate than another because units have no inherent meaning in Likert scales (supplementary text S1) (77).

To gauge whether the effects on forgiveness are large, we instead benchmark the treatment effect against how exposure to specific forms of war violence affected feelings toward perpetrators, as reflected in the forgiveness index at baseline. For example, having a family member killed lowered baseline forgiveness by 0.920 ( $SE = 0.232$ ,  $P < 0.001$ ) (table S14). Thus, the reconciliation program can be said to offset this effect and increase forgiveness by 30% ( $0.277/0.920 = 0.301$ ). This approach is speculative because we cannot observe the causal effect of violence exposure on forgiveness, and so we are benchmarking our treatment effect against a correlation. As such, the interpretation of magnitudes in this manner should be taken as suggestive.

These forgiveness effects are based on survey responses, which raise potential concerns that respondents may say what they believe surveyors

want to hear. But there are four factors that mitigate the concern that the results are driven by social desirability bias. First, our surveyors are completely independent of the implementing NGO, so they would not be associated with messages of reconciliation. Second, we asked these questions 9 to 31 months after the reconciliation ceremonies take place, so talk of forgiveness is not fresh on respondents' minds. Third, respondents are not simply asked whether they have forgiven the perpetrator, but rather asked a series of questions designed to gauge their feeling and behavior toward excombatants (such as avoidance), which are arguably less subject to this type of bias. Last, our respondents experienced traumatic forms of victimization, such as amputations and the killing of family members, so it is not psychologically costless for them to say that they no longer feel anger toward their perpetrator, unless this reflects an underlying change in their perspective. However, to further bolster this interpretation, we also discuss whether these forgiveness effects go hand-in-hand with changes in the community orientation of individuals' behavior.

Next, we examine impacts on trust. As shown in Table 1, bottom, the reconciliation treatment increases trust toward both rebel excombatants and migrants. Looking at the binary indicator, respondents are on average 7.3 percentage points ( $SE = 0.036$ ,  $P = 0.046$ ), or 22.2%, more likely to trust a rebel excombatant and 5.8 percentage points ( $SE = 0.012$ ,  $P < 0.001$ ), or 6.7%, more likely to trust a migrant. Higher trust of migrants suggests greater inclusion of this marginalized group, whose members are sometimes difficult to distinguish from excombatants. In contrast, there is no discernible impact on trust toward former members of the SLA or CDF (table S8). This indicates that the reconciliation process led to changes in trust toward those who perpetrated atrocities during the war, namely former members of the RUF. Although all of these trust ques-

tions are administered to subsets of individuals who know members of each of these groups, our specifications restrict the sample to those who knew group members at both baseline and endline because they include controls for the baseline-dependent variable. In table S15, we further verify that these results are not driven by compositional changes in who knows members of these groups.

Because reconciliation is aimed at forgiving former war perpetrators, it is reassuring to see that the process did increase trust toward former rebel combatants. Yet, at the same time, there is no significant impact on the index of trust toward community members generally (Table 1, bottom). Moreover, the reconciliation process also did not alter individuals' beliefs that former combatants and other community members would fight again in the future (table S16). Both null effects raise questions as to whether the treatment altered individuals' interactions with other community members.

To further investigate this question, we examine impacts on social networks (Table 2). The coefficient on the mean effect index implies that the index of network strength is 0.099 standard deviation (SD) units larger in treatment sections than control sections ( $SE = 0.028$ ,  $P = 0.001$ ). Because the index is an aggregation of various indicators, effect sizes have a more intuitive meaning if we look at the individual indicators constituting the index.

For example, the number of individuals whom respondents would ask for advice or help increases by 0.148 above the control group mean of 2.894 ( $SE = 0.069$ ,  $P = 0.033$ ), implying a 5% increase. The tendency to be listed as a good friend and as someone to ask for advice or help both also increase by 11% ( $SE = 0.091$ ,  $P = 0.013$  and  $SE = 0.126$ ,  $P = 0.005$ , respectively). As we discuss in supplementary text S5, we see no significant differential impact of ceremony attendance on the

**Table 4. Reconciliation and contributions to public goods.** Each row represents a separate regression of the outcome shown in the first column on treatment assignment. All specifications include section pair fixed effects and the second-round indicator. All specifications also include the baseline outcome variable and its interaction with both the second-round indicator and the second-wave indicator, except for "Contributed money to someone in need." Because we do not have the second-wave baseline-dependent variable for this outcome, we instead control for the other baseline measures of public goods contributions and their interaction with the second-round indicator and the second-wave indicator in the regression of this outcome. \*\*\* 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.

Variables	Control mean	Coefficient	SE	Observations	R <sup>2</sup>
Index of public goods contributions	0	0.042*	(0.022)	3008	0.171
Index of public goods contributions (without contributions to women's group)	0	0.046**	(0.023)	3008	0.184
Index of public goods contributions (indicators in both baselines)	0	0.046**	(0.022)	3008	0.171
Indicators					
Contributed to public facilities	0.397	0.029	(0.019)	2911	0.078
Brushed roads	0.290	0.005	(0.014)	2898	0.171
Number of community projects (village-level variable)	0.539	-0.060	(0.057)	2841	0.356
Contributed to PTA	0.066	0.023*	(0.013)	2732	0.105
Contributed to village development committee	0.062	0.002	(0.009)	2737	0.119
Contributed to youth group	0.069	-0.002	(0.006)	2738	0.081
Contributed to women's group	0.064	0.021**	(0.010)	2004	0.076
Contributed money to someone in need	0.178	0.010	(0.019)	2039	0.100

social networks index, which suggests that this effect does not arise as a mere consequence of social interactions generated at the ceremony.

We next examine whether the reconciliation process altered the community orientation of individuals' behavior. Estimates on participation in community groups are presented in Table 3. The mean effect index is significantly higher in treatment sections by 0.058 SD units (SE = 0.017, P = 0.001). This overall effect reflects two different types of impacts among individual indicators. Membership and meeting attendance for almost all of the individual community groups increase, with effect sizes ranging from 11% above the control group mean of 0.10 for youth group membership (SE = 0.008, P = 0.066) to 45% above the control group mean of 0.08 for PTA meeting attendance (SE = 0.015, P = 0.017). In contrast, membership and meeting attendance decrease for secret societies by 16% (SE = 0.019, P = 0.004) and 9% (SE = 0.008, P = 0.516), respectively. These effects are interesting because these groups have a closed membership dominated by the elite (72). Thus, these decreases are consistent with substitution toward more broad-based community organizations.

Because the Peace Mothers' Groups are initiated as a part of the intervention, we verify that women's groups do not mechanically drive this result: Dropping women's group membership and attendance from the index does not meaningfully affect the estimate.

The effects on contributions to public goods are gauged in Table 4. The mean effect index is 0.042 SD units larger in treatment villages (SE = 0.022, P = 0.055). Dropping the women's group indicators and the indicator for giving to someone

in need (for which we do not have second-wave baseline data) does not substantively alter the estimate. Also, dropping the indicator of the number of community projects in a village from the index does not meaningfully affect the estimate (table S17), suggesting that imputation of missing village-level data does not drive this result.

The coefficient on public goods contributions is the smallest of our significant effects, among the mean effect indices. However, looking within the index again shows that effects on underlying indicators vary in magnitude. The effects are most precisely estimated and largest for contributions to PTAs and women's groups, where implied increases are 32% (SE = 0.013, P = 0.097) and 20% (SE = 0.01, P = 0.045), albeit from relatively small control group means of 6.6 and 6.4%, respectively. The implied effect for contributing to public facilities is 7% (SE = 0.019, P = 0.126), but from a relatively large control group mean of 40%.

These effects on networks, participation, and contribution also support our interpretation that the forgiveness effects are not driven through socially desirable responses regarding anger toward perpetrators because they are coupled with changes in the community orientation of individuals' actions. They also indicate that the reconciliation process boosted social capital as individuals formed more friendships and contributed more to their communities, although these changes were not accompanied by increases in general trust, which increased specifically for migrants and former rebel combatants.

Next, we turn from societal healing to individual healing. The effects on psychological well-being are presented in Table 5. The first row presents the index of complete indicators (with

pared baseline controls for wave two). The second row presents the index with just the subset of indicators appearing in the wave two baseline. Both versions show that psychological health was significantly lower in the treatment villages, by 0.147 and 0.138 SD units, respectively (SE = 0.033, P < 0.001 and SE = 0.031, P < 0.001). This overall negative impact stems from a worsening of all three psychological measures.

The dichotomous indicator of clinical PTSD indicates that severe trauma was 36% higher in treatment sections, above the control group mean of 8% (SE = 0.011, P = 0.006). The control group means of the continuous psychometric indicators are again not useful for gauging magnitudes in percent terms because they are also aggregations on a Likert scale. If we instead take the alternate approach of comparing the treatment effect against baseline effects of being maimed (table S14), the treatment is predicted to worsen PTSD by 28%, depression by 47%, and anxiety by 37%. Thus, both the percent effects with the dichotomous PTSD indicator and the more speculative approach of benchmarking against violence exposure are consistent with one another and imply substantial effects.

We found that all of these effects, both positive and negative, are also robust to alternate specifications, as discussed in supplementary texts S3 and S6 and presented in tables S18 and S19.

The negative impacts on psychological well-being suggest that confronting the past through reconciliation processes may be deeply distressing. But are these effects concentrated among victims, specifically? This is important for gauging distributional consequences of the program.

To examine whether the psychological impacts are larger for those who were victimized during

**Table 5. Reconciliation and psychological well-being.** The top portion of the table examines the average treatment effect. Each row represents a separate regression of the outcome shown in the first column on treatment assignment. The control mean is the mean in the control group at endline. The bottom portion examines how the treatment effect varies according to individuals' exposure to violence. All specifications include section pair fixed effects and the second-round indicator, the baseline outcome variable, and its interaction with both the second-round indicator and the second-wave indicator. SEs 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.

Variables	Control mean	Coefficient	SE	Observations	R <sup>2</sup>	
<i>Average effect</i>						
Index of psychological well-being (all indicators)	0	-0.147***	(0.033)	2982	0.115	
Index of psychological well-being (indicators in both baselines)	0	-0.138***	(0.031)	2982	0.115	
<i>Indicators (in both baselines)</i>						
Less PTSD	28.819	-0.683***	(0.197)	2776	0.119	
Less anxiety	14.945	-0.441***	(0.117)	2895	0.142	
Less depression	11.677	-0.289***	(0.069)	2913	0.092	
Clinical PTSD symptoms present	0.080	0.029***	0.011	2776	0.057	
<i>Effect by violence exposure (saw violence, was raped, maimed, beaten, or abducted)</i>						
		<b>T</b>		<b>T × violence-exposed</b>		
		<b>Coefficient</b>	<b>SE</b>	<b>Coefficient</b>	<b>SE</b>	
Index of psychological well-being	-0.160***	(0.052)	0.011	(0.064)	2852	0.121
Index of psychological well-being (indicators in both baselines)	-0.147***	(0.052)	0.005	(0.064)	2852	0.121
Less PTSD	-0.871***	(0.309)	0.298	(0.391)	2662	0.123
Less anxiety	-0.476**	(0.213)	0.003	(0.268)	2778	0.144
Less depression	-0.270**	(0.127)	-0.044	(0.162)	2788	0.094
Clinical PTSD symptoms present	0.038**	(0.018)	-0.010	(0.022)	2662	0.058

the war, we interact treatment with our pre-specified measure of violence exposure (Table 5). We have limited power to identify these heterogeneous treatment effects, so we considered the magnitudes of the coefficients instead of focusing solely on statistical significance. However, the coefficients on the interaction terms are not just imprecisely estimated but also differ in sign across indicators. (The coefficient is negative for depression, indicating worse effects for victims, but positive for PTSD and anxiety.)

These results are consistent with the idea that even nonvictims may experience a worsening of psychological health from going through a reconciliation process. For example, other community members may experience vicarious traumatization from hearing about atrocities done to others (36–39).

Another way of gauging the distributional consequence of the reconciliation treatment is to see whether the impact on social capital is smaller (or larger) for victims. These interaction effects are examined in table S20 with two measures of violence exposure. The coefficients on the interaction term of treatment and victimization are mixed in sign, small in size, and imprecisely estimated for outcomes such as social networks, public goods contributions, and community group participation. Thus, victims do not appear to partake systematically less in social capital improvements.

We examine in table S21 whether effects vary for excombatants. Here, some of the interaction terms are quite large in magnitude; for example, the effect on the psychological well-being index implies that the negative effect on those who are not excombatants is nearly offset for excombatants. These effects are imprecisely estimated in part because the excombatant variable is likely to be underreported and measured with noise. Thus, it is difficult to draw definitive conclusions on the basis of these heterogeneous effects, and future work should probe this further.

A key issue is whether these effects persist over time. We present in Table 6 short-run and long-run effects using the two rounds of wave-one data.

Because wave one includes fewer than half the sections in the evaluation, this is a relatively underpowered sample, and some of the effects are individually insignificant. Yet, the broad pattern implied by the coefficients indicates that both the positive and negative effects are sustained.

First, the impacts on all three psychological measures persist up to 31 months. This suggests that the war memories invoked by the reconciliation process are powerful and do not fade quickly.

At the same time, the effects on forgiveness and social capital outcomes also persist. Although the effect on trust of former rebel combatants is individually insignificant in both rounds, the coefficients are not significantly distinguishable from each other at the 5% level, indicating that they do not recede over time. Trust of migrants also persists, and there are even short-run improvements in generalized trust measures, although these effects fall, and significantly so, over the longer horizon. The coefficients on public goods contributions and social networks, if anything, increase in magnitude, suggesting that the effects do not recede. The effect on community group participation is also individually significant in both rounds. As such, reconciliation appears to boost the community orientation of individual behavior in a manner that does not subsequently fade away.

We interpret the results above as indicating that the reconciliation process itself affects both individual and societal healing. We also consider and present evidence against two alternative accounts, drawing on data for additional outcomes.

The first alternative account posits that the reconciliation ceremony may be relatively unimportant, whereas other components of the intervention—such as the Peace Mother's Group, Communal Farms, or Peace Tree—actually drive the estimated effects. We think that this is unlikely because treatment effects on forgiveness, social capital, and psychological health are not statistically distinguishable for men and women (table S22), nor if we include a control for communal farms (table S23). The effect on economic outcomes is even negative in sign (table S24), further suggesting

limited impacts of communal farms. The inefficient capturing effects on the resolution of day-to-day disputes, which was the focus of the Peace Tree, is also negative and imprecise (table S25). Moreover, it is difficult to see how the negative effects on psychological well-being could emerge as a response to these other components. Together, these results suggest that the reconciliation component of the intervention is an important driver of the estimated effects.

A second alternative account posits that the reconciliation component may be driving the psychological effects, but the social capital outcomes arise from simply getting community members together in a gathering. We think that this is unlikely because it has proven incredibly difficult to move social capital outcomes in Sierra Leone. For example, a large-scale Community-Driven Restoration (CDR) program was implemented in Sierra Leone in 2008 in one of the same districts as in our study. This program spent \$100 per household and fostered ongoing gatherings of the community in village-wide meetings in order to promote inclusive governance and collective action. A randomized evaluation found it successfully delivered economic benefits but had no effects on social capital outcomes such as community group participation, as measured with indicators similar to those used in our study (4). Given that social capital outcomes did not move in response to a well-implemented and well-resourced intervention, it is hard to see how a 2-day gathering initiated by Fambul Tok could deliver persistent effects on similar outcomes for up to nearly 3 years after the intervention, unless it entailed a deeper transformation of person-to-person interactions.

Second, if simply getting people together improved person-to-person interactions, then we should also have observed reductions in societal tensions and the incidence of other day-to-day disputes. But again, little support for this idea is provided in table S25. For example, the coefficient on the number of conflicts is only 0.002 (relative to a mean of 0.16) (SE = 0.019,  $P = 0.894$ ). Rather, we observe improvements in

**Table 6. Persistence of effects.** These results present separate estimates for the two endline rounds in wave one. 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. The final column indicates whether the specification also includes the baseline outcome variable, and its interaction with both the second-round indicator and the second-wave indicator. SEs 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.

Variables	Round 1			Round 2			Baseline-dependent variable controls?
	Coefficient	SE	Observations	Coefficient	SE	Observations	
Forgive perpetrators	0.986***	(0.272)	550	1.231***	(0.361)	521	Y
Trust rebel excombatants	0.100	(0.073)	241	0.048	(0.198)	203	Y
Trust migrants	0.140**	(0.053)	653	0.119*	(0.069)	564	Y
Index of generalized trust in community	0.119**	(0.050)	878	-0.009	(0.038)	845	Y
Index of network strength	0.015	(0.027)	885	0.119	(0.085)	850	N
Index of community group participation	0.038*	(0.022)	884	0.084**	(0.040)	847	Y
Index of contributions to public goods	0.024	(0.033)	885	0.035	(0.046)	850	Y
Index of psychological well-being	-0.166***	(0.052)	873	-0.170***	(0.058)	837	Y

outcomes that are specific to the war, such as forgiveness of war perpetrators and trust of former rebel combatants. This reiterates the idea that talking about the war is important in giving rise to the observed effects. In the supplementary materials, we examine these additional outcomes further by gauging their long-run impacts (table S19) and discussing them in greater detail.

## Conclusion

Our findings highlight the long shadow of war along two dimensions. The reconciliation forums we analyzed were held nearly a decade after the end of Sierra Leone's civil war. Yet, 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 these war memories when they are evoked suddenly (32–34).

These psychological effects do not preclude the possibility that individuals who forgave in response to reconciliation gained a psychological benefit—but they do suggest that these gains were offset by other negative impacts, such as the difficulty of coping with negative memories. In that regard, they corroborate the idea that forgiving is not the same as forgetting (73). They also suggest that forgiveness stemming from an intense, one-time event that evokes negative memories may differ in its psychological impact relative to forgiveness stemming from ongoing therapy (74).

Overall, our results indicate that the gains in societal healing associated with reconciliation came at a substantial cost in individual psychological healing. As such, they imply that policymakers need to find ways of holding reconciliation processes that reduce these psychological costs, while retaining the societal benefits. For example, it is possible that the negative psychological impacts may be smaller or even reversed if reconciliation efforts are held in the direct aftermath of the war, when trauma symptoms are high and people have yet to move on in their own way (75). A second possibility lies in combining reconciliation with other types of complementary interventions. For example, coupling these programs with sustained counseling—as used by forgiveness therapies (26–31), exposure therapy (32, 36), or trauma healing interventions (17)—may help mitigate the detrimental impacts. Given the global prevalence of conflict and postconflict reconciliation, future research should explore alternate designs for efforts aimed at unifying societies in the aftermath of war.

## REFERENCES AND NOTES

- P. Tettersson, P. Wallensteen, *J. Peace Res.* **52**, 536–550 (2015).
- J. Fearon, M. Humphreys, J. Weinstein, *Am. Econ. Rev.* **99**, 287–291 (2009).
- J. Fearon, M. Humphreys, J. Weinstein, *Am. Polit. Sci. Rev.* **109**, 450–469 (2015).
- K. Casey, R. Glennerster, E. Miguel, *Q. J. Econ.* **127**, 1755–1812 (2012).
- A. Beath, F. Christia, R. Nikolopov, *Am. Polit. Sci. Rev.* **107**, 540–557 (2013).
- M. Humphreys, R. Sanchez de la Sierra, P. van der Windt, *Polit. Anal.* **21**, 1–20 (2013).
- G. Mansuri, V. Rao, *Localizing Development: Does Participation Work?* (The World Bank, 2013).
- E. Brahm, "Patterns of truth: Examining truth commission impact in cross-national context," paper presented at the annual meeting of the International Studies Association, Honolulu, HI (2005).
- N. Kritz, *War Crimes: The Legacy of Nuremberg*, B. Cooper, Ed. (TV Books, 1999).
- H. Cobban, *Boston Review* **27**, 2 (2003).
- B. Hamber, *Reconciliation After Violent Conflict: A Handbook* (International Institute for Democracy and Electoral Assistance, 2003).
- N. Biggar, *Burying the Past: Making Peace and Doing Justice After Civil Conflict* (Georgetown Univ. Press, 2003).
- J. Lederarch, *The Journey Towards Reconciliation* (Herald Press, 1999).
- K. Asmal, L. Asmal, R. Roberts, *Reconciliation Through Truth: A Reckoning of Apartheid's Criminal Governance* (Cambridge Univ. Press, 1994).
- Truth and Reconciliation Commission, *Truth and Reconciliation Commission of South Africa Report* (South African Government, 1998).
- J. Gibson, *Am. J. Pol. Sci.* **48**, 201–217 (2004).
- E. L. Staub, A. Pearlman, A. Gubin, A. Hagegimama, *J. Soc. Clin. Psychol.* **24**, 297–334 (2005).
- E. L. Paluck, *J. Pers. Soc. Psychol.* **96**, 574–587 (2009).
- C. Blattman, A. C. Hartman, R. A. Blair, *Am. Polit. Sci. Rev.* **108**, 100–120 (2014).
- D. Mendeloff, *Int. Stud. Rev.* **6**, 355–380 (2004).
- D. Mendeloff, *Hum. Rights Q.* **31**, 592–623 (2009).
- M. Kaldor, J. Vincent, *Evaluation of UNDP assistance to conflict-affected countries: Case study Sierra Leone* (United Nations Development Programme Evaluation Office, 2006).
- R. Conibere et al., *Statistical Appendix to the Report of the Truth and Reconciliation Commission of Sierra Leone* (Benetech Human Rights Data Analysis Group, 2004), pp. 1–38.
- L. A. Smith, C. Gambette, T. Longley, *Conflict Mapping in Sierra Leone: Violations of International Humanitarian Law from 1991 to 2002: Executive Summary* (No Peace Without Justice, 2004).
- R. D. Enright, R. Fitzgibbons, *Helping Clients Forgive: An Empirical Guide for Resolving Anger and Restoring Hope* (American Psychological Association, 2000).
- R. H. Al-Mabuk, R. D. Enright, P. A. Cardis, *J. Moral Educ.* **24**, 427–444 (1995).
- C. T. Coyle, R. D. Enright, *J. Consult. Clin. Psychol.* **65**, 1042–1046 (1997).
- S. R. Freedman, R. D. Enright, *J. Consult. Clin. Psychol.* **64**, 983–992 (1996).
- W. F. Lin, D. Mack, R. D. Enright, D. Krahn, T. W. Baskin, *J. Consult. Clin. Psychol.* **72**, 1114–1121 (2004).
- M. S. Rye et al., *J. Consult. Clin. Psychol.* **73**, 880–892 (2005).
- G. L. Reed, R. D. Enright, *J. Consult. Clin. Psychol.* **74**, 920–929 (2006).
- J. Joseph, M. Gray, *J. Behav. Anal. Offend. Victim Treat. Prevent.* **1**, 69–79 (2008).
- A. A. van Emmerik, J. H. Kamphuis, A. M. Hulsbosch, P. M. G. Emmelkamp, *Lancet* **360**, 766–771 (2002).
- K. Brounéus, *J. Conflict Resolut.* **54**, 408–437 (2010).
- S. Rose, J. Bisson, R. Churchill, S. Wessely, *Cochrane Database Syst. Rev.* **2**, CD000560 (2002).
- J. Bisson, M. Andrew, *Cochrane Libr.* **2009**, 1 (2009).
- I. L. McCann, L. Pearlman, *J. Trauma. Stress* **3**, 131–149 (1990).
- L. A. Pearlman, W. Karen, Saakvitne, *Trauma and the Therapist: Countertransference and Vicarious Traumatization in Psychotherapy with Incest Survivors* (Norton, 1995).
- M. J. Arvey, *Int. J. Adv. Couns.* **23**, 283–293 (2001).
- M. Cunningham, *Soc. Work* **48**, 451–459 (2003).
- R. Shaw, "Rethinking Truth and Reconciliation Commissions: Lessons from Sierra Leone," United States Institute of Peace Special Report 130 (USIP Press, 2005).
- D. Backer, "The human face of justice: Victims' responses to South Africa's truth and reconciliation commission process," thesis, University of Michigan (2004).
- D. Backer, *Security, Reconstruction, and Reconciliation: When the Wars End*, M. Ndulo, Ed. (University College London, 2007) p. 165–196.
- D. Kaminer, D. J. Stein, I. Mbangi, N. Zungu-Dirwayi, *Br. J. Psychiatry* **178**, 373–377 (2001).
- K. Brounéus, *Secur. Dialogue* **39**, 55–76 (2008).
- R. D. Putnam, *J. Democracy* **6**, 65–78 (1995).
- J. Quinn, *The Politics of Acknowledgement: Truth Commissions in Uganda and Haiti* (UBC Press, 2010).
- L. Frison, S. J. Pocock, *Stat. Med.* **11**, 1685–1704 (1992).
- D. McKenzie, *J. Dev. Econ.* **99**, 210–221 (2012).
- A. Banerjee et al., *Science* **348**, 1260799 (2015).
- M. Bruhn, D. McKenzie, *Am. Econ. J. Appl. Econ.* **1**, 200–232 (2009).
- J. R. Kling, J. B. Liebman, L. F. Katz, *Econometrica* **75**, 83–119 (2007).
- M. L. Anderson, *J. Am. Stat. Assoc.* **103**, 1481–1495 (2008).
- Y. Benjamini, Y. Hochberg, *J. R. Stat. Soc.*, **B 57**, 289–300 (1995).
- D. Curran-Everett, D. J. Benos, *Am. J. Physiol. Lung Cell. Mol. Physiol.* **287**, L259–L261 (2004).
- D. Curran-Everett, *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **279**, R1–R8 (2000).
- M. S. Rye, "Evaluation of a secular and a religiously integrated forgiveness group therapy program for college students who have been wronged by a romantic partner," thesis, Bowling Green State University, Bowling Green, OH (1998).
- M. J. Subkoviak et al., *J. Adolesc.* **18**, 641–655 (1995).
- E. Foa, D. Riggs, C. Dancu, B. Rothbaum, *J. Trauma. Stress* **6**, 459–473 (1993).
- E. B. Foa, D. F. Tolin, *J. Trauma. Stress* **13**, 181–191 (2000).
- W. W. Zung, C. B. Richards, M. J. Short, *Arch. Gen. Psychiatry* **13**, 508–515 (1965).
- W. W. Zung, *Psychosomatics* **12**, 371–379 (1971).
- L. R. Derogatis, R. S. Lipman, K. Rickels, E. H. Uhlenhuth, L. Covi, *Behav. Sci.* **19**, 1–15 (1974).
- S. Theresa, *Integr. Psychiatry* **7**, 60–62 (2010).
- T. S. Betancourt et al., *Child Dev.* **81**, 1077–1095 (2010).
- T. S. Betancourt, I. I. Borisova, M. de la Soudière, J. Williamson, *J. Adolesc. Health* **49**, 21–28 (2011).
- A. M. Steinberg, M. J. Brymer, K. B. Decker, R. S. Pynoos, *Curr. Psychiatry Rep.* **6**, 96–100 (2004).
- S. S. Hawkins, J. Radcliffe, *J. Pediatr. Psychol.* **31**, 420–430 (2006).
- C. P. Bayer, F. Klases, H. Adam, *JAMA* **298**, 555–559 (2007).
- F. W. Weathers, J. A. Huska, T. M. Keane, *PCL-M for DSM-IV* (National Center for PTSD—Behavioral Science Division, 1991).
- G. M. Sullivan, A. R. Artino Jr., *J. Grad. Med. Educ.* **5**, 541–542 (2013).
- W. P. Murphy, *Africa J. Int. Afr. Inst.* **50**, 193–207 (1980).
- T. W. Baskin, R. D. Enright, *J. Couns. Dev.* **82**, 79–90 (2004).
- R. D. Enright, A. Holter, T. Baskin, C. Knutson, *J. Res. Ed.* **17**, 63–78 (2007).
- E. B. Foa, *J. Clin. Psychiatry* **67** (suppl. 2), 40–45 (2006).

## ACKNOWLEDGMENTS

We thank J. Caulker from Fambul Tok and L. Hoffman from Catalyst for Peaker for enabling us to study Fambul Tok's initiative. We are grateful to E. Dixon, Q.-A. Do, M. Gilligan, S. Mullainathan, and D. Stasavage for providing detailed comments and suggestions on the manuscript and to T. Betancourt for providing valuable input on the psychometric survey measures used in the study. We additionally thank A. Mansaray, A. Ahmed, J. Creighton, and N. Hasham (in chronological order of involvement) for outstanding research assistance in the field. This project was funded by 3ie (grant code OW2.253) and the J-PAL Governance Initiative (subaward 5710003563) and received extensive implementation support from Innovations for Poverty Action. We thank them, without implicating them, for making this project possible. This research also received institutional review board (IRB) approval from Oxford University (ref. SSD/CURECI/11-028), New York University (HS# 11-8528; IRB#: 14-9936), and Innovations for Poverty Action (Protocol 534.11February002). All errors and omissions are our own. Upon publication, the replication data and code from this study will be made publicly available through the Interuniversity Consortium for Political and Social Research (ICPSR) data depository. The Pre-Analysis Plan, which is already posted publicly in the EGAP depository, will also be made available as a part of the ICPSR replication package.

## SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6287/787/suppl/DC1  
Supplementary Texts S1 to S6

Figs. S1 to S4

Tables S1 to S26

References (76–98)

30 November 2015; accepted 29 March 2016  
10.1126/science.aad9682





**Reconciling after civil conflict increases social capital but decreases individual well-being**

Jacobus Cilliers, Oeindrila Dube and Bilal Siddiqi (May 12, 2016)  
*Science* **352** (6287), 787-794. [doi: 10.1126/science.aad9682]

Editor's Summary

**The psychological cost of reconciliation**

During civil wars, individuals and communities who were previously good neighbors can end up fighting each other. One approach to reknit these sundered social ties is to bring perpetrators and victims together in truth and reconciliation forums. Cilliers *et al.* found that these forums have helped to reestablish social bonds in Sierra Leone, but that they have also imposed a cost on the victims' mental health (see the Perspective by Casey and Glennerster).

*Science*, this issue p. 787; see also p. 766

---

This copy is for your personal, non-commercial use only.

---

- Article Tools** Visit the online version of this article to access the personalization and article tools:  
<http://science.sciencemag.org/content/352/6287/787>
- Permissions** Obtain information about reproducing this article:  
<http://www.sciencemag.org/about/permissions.dtl>

*Science* (print ISSN 0036-8075; online ISSN 1095-9203) is published weekly, except the last week in December, by the American Association for the Advancement of Science, 1200 New York Avenue NW, Washington, DC 20005. Copyright 2016 by the American Association for the Advancement of Science; all rights reserved. The title *Science* is a registered trademark of AAAS.

# Supplementary Materials to:

## Reconciling after civil conflict increases social capital but reduces individual well-being

### Text 1: Data and Sampling

#### Construction of Mean Effect Indices

We use two methods for aggregating our outcomes into mean effect indices. For our main results, we use the methodology of Kling Liebman Katz (hereafter, KLK) (52). Constructing these indices entails orienting each outcome so higher values always indicate “better” outcomes; standardizing them by subtracting mean and dividing by the standard deviation from the control group; imputing missing values in the treatment (control) group with the treatment (control) mean; and then adding the standardized indicators together into an index that gives equal weight to each indicator. This approach has the advantage that it helps us account for missing data. Effects on the index can be interpreted as the average of results for separate measures, scaled to standard deviation units (52).

We also construct indices using a second methodology of Anderson (53), which does not impute missing values, and which weights the standardized 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 KLK methodology has the advantage that it accounts for missing values. If there are many outcomes in an index, and there is some non-response in each outcome, aggregating across outcomes without imputation can produce a large loss of observations in the aggregate index. On the other hand, when an individual is missing data on an outcome, simply averaging across non-missing outcomes would implicitly place greater weight on those non-missing outcomes (76). Thus our main results use the KLK methodology and we present robustness using the Anderson methodology. As detailed below, our results are similar under the two approaches.

#### Description of Data on Additional Outcomes

Here, we provide an overview of the additional variables collected in our survey that we use for results presented in the supplement.

We tracked various types of economic activity. Our index includes measures of the frequency and size of lending and borrowing; time spent working on other people's farms over the past 3 months; 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 (PCA). This index includes 16 household goods and whether the roof is made out of straw and the walls out of mud. We also asked respondents to report their subjective assessments of their ability to meet basic household needs such as school fees and health care, and their perceptions on 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.

In the household survey, we tracked the incidence of conflicts related to loans, land, property and religion that the respondent had experienced over the past 6 months. We similarly tracked the incidence of violent and non-violent crimes. These measures are aggregated across households in a village to create the number of events at the village level. We also recorded inter-village disputes over the past 6 months from a village-level survey. These four measures are used to form an index of conflict and crime. A limitation here is that the village survey is missing for 5 villages at baseline (and 6 at endline); and the inter-village disputes variable is further missing for a 6<sup>th</sup> village at

baseline. Thus we examine effects on this index cross-sectionally to avoid losing observations from the missing baseline data.

We also recorded whether conflicts were resolved, satisfactorily resolved, and how they were resolved -- i.e., without a third party, with mediation by friends and family, and by the chief. We summed across individuals in each village and scaled by the number of conflicts to attain proportions of these variables at the village level. Thus these measures are restricted to villages that had some positive level of conflict. In addition, we asked about fines levied as a form of punishment. These variables are used to gauge impacts on conflict resolution and the entrenchment of traditional authorities in conflict resolution.

Our measure of social divisions comes from the World Bank (See: <http://go.worldbank.org/BOA3AR43W0>). It includes an indicator of the extent to which divisions between non-marginalized and marginalized groups (migrants vs. non-migrants, the young vs. old, and the poor vs. rich) escalate into violence; as well as 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. The indicator of whether divisions between non-marginalized and marginalized groups is missing for the second wave baseline so we use pared down baseline dependent variables for these outcomes.

We additionally measured gender attitudes. We asked respondents whether a wife has a right to express her own opinions. We also asked respondents if they felt it is acceptable for "a man to beat his wife" under seven different circumstances, coding their responses ("Always" "Sometimes" and "Never.") on a 3-point Likert scale. The wave two baseline asked a limited set of questions related to attitudes toward wife beating (covering four out of seven circumstance questions); it also didn't administer the question about whether a wife has a right to form her own opinion, so we control for pared down baseline dependent variable controls when examining this specification.

## **Likert-Scale Responses**

Our forgiveness and psychological wellbeing measures are sums of questions answered on Likert scales. For the forgiveness questions, 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?" (Table S10 lists all 12 questions). To give the responses an intuitive meaning, we code agreement in the negative domain, and disagreement on the positive domain symmetrically around 0. Since there are four responses, and we use the convention of coding across responses in steps of one, the resultant values range from -1.5 to 1.5. Since linear regression is invariant to affine transformations, coding the variable with other values in steps of one, such as 1 to 4, does not affect the estimated regression coefficients. However, the latter coding does yield a different control group mean. So, if the regression coefficients under the two coding approaches were compared against their respective control group means, they would imply different percent effects. This highlights one reason why estimates for outcomes coded on a Likert scale should not be benchmarked against control group means and interpreted in percent terms.

## **Psychological Wellbeing Indicators**

The 11 questions we use to construct the PTSD index are drawn from the 17-item PTSD Symptom Scale (59-60), that assesses presence and severity PTSD Symptoms according to the 4<sup>th</sup> Diagnostic and Statistical Manual of Mental Disorders (DSM-IV). These are the same questions as the PTSD Checklist, Civilian Version (PCV-C) as developed by at the National Center for PTSD (70). These questions focus on three categories of symptoms: re-experience, arousal, and avoidance. The responses to these questions are aggregate on a 4-point Likert scale.

Previous validation exercises for research purposes have demonstrated that both the PTSD Symptom Scale and PTSD Checklist show high internal consistency and high test-retest reliability, concurrent validity, and high correlation with other diagnostic scales (59-60, 77-79). See (80) for a detailed review.

Our seven depression and 10 anxiety questions are drawn from the Zung Depression and Zung Anxiety indices (61-62), respectively. These indices have been found to have high internal consistency (81). The responses

to these questions are also aggregate on a 4-point Likert scale. Table S11 provides a complete list of all the psychological wellbeing questions in our study.

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 contexts. To adapt these questions to the Sierra Leonean context we piloted our surveys extensively and adjusted the wording of the scales to better reflect Krio language. For example, one PTSD question is: "Have you had recurrent or intrusive distressing thoughts or recollections about the assault". We adapted this question so it reads: "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) (67-68), 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 (82); and it has also been used in a sample of Ugandan and Congolese child soldiers (69). As another example, the Survey of War-Affected Youth (SWAY) conducted in Uganda (83) also administered nine out of our eleven PTSD questions and six out of the seven anxiety questions.

Furthermore, 15 out 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 (66) and HIV-positive pregnant women Tanzania (84) 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 only looking at 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).

## Missing Village-Level Survey Data

Due to mechanical error in data collection, our village-level survey is missing for five villages in baseline, and a separate group of six villages in endline. This missing data is unlikely to have substantial effects on our estimates for two reasons. First, in Table S3, we verify that whether the village is missing in baseline, or endline, or in either is not predicted by treatment. Second, fortunately, our analysis only uses four variables from the village-level surveys, as most of the indicators come from the household-survey.

The four variables are used in the construction of three mean effect indices. Since the indices constructed using the KLK methodology imputes missing values, here we discuss the steps we took to verify that missingness in these four indicators do not skew our estimates

Two of the indices are comprised of many individual level indicators from the household survey along with one village level variable. For example, the public goods index (Table 4) is comprised of seven individual level indicators along with the village-level community projects variable; and the index of economic activity (Table S24) has six individual level indicators along with number of traders and communal farms in the village. In these cases, we take the approach of constructing alternate indices dropping the village level indicators for these two outcomes, and present these as additional robustness checks.

One of the indices, on conflicts and crime, is composed of three other indicators that come from the household survey (aggregated to the village level), along with the inter-village disputes variable that comes from the village-level survey. We examine effects on this index cross-sectionally, without controls for the baseline dependent variable, to avoid losing 6 village-level observations missing in the baseline (5 from the missing village survey, and 6<sup>th</sup> which is missing only the inter-village disputes variable).

However, the inter-village disputes variable is also missing for 6 villages at endline. Since the conflict index is analyzed at the village level, and imputing missing values to 6 villages constitutes a relatively large share of

the village-level observations, we take the conservative step of dropping these 6 villages missing inter-village disputes in constructing the entire index. As an additional check, we also examine effects instead by dropping the inter-village conflicts indicator from the conflict index (as we did with the other indices that include village level variables). We find that our results look similar under either approach (Table S17). These additional steps help ensure that the missing data for the four village level variables do not affect our results.

## **Attrition and Sample size**

We measure attrition as those who were surveyed in the baseline but were not re-surveyed at endline. For wave 1, the attrition rate for respondents who appeared in the baseline but are missing in the round 1 endline is 7%, and for those who are missing in the round 2 endline is 10%. The attrition rate for the wave 2 endline is 11%. Pooled across the two waves and three endline rounds, the attrition rate of those who appeared in baseline but are missing from either endline round in wave 1 or the endline in wave 2 is 13%; and the attrition rate for those missing from both endline rounds in wave 1 and the endline in wave 2 is 7%. Table S2 shows that neither of these attrition measures, nor the attrition measure of each endline separately, are predicted by treatment.

Table S4 shows a breakdown of individuals and villages surveyed by wave and round. It shows that out of 952 households that we surveyed in baseline wave one, we were able to re-survey 885 (850) households in the first (second) round of endline data collection. Note that a total of 794 households were surveyed in both rounds, and 941 households were surveyed in either round. Out of the 1430 households we surveyed in wave 2, we were able to re-survey 1273 in the endline. Turning to the village survey, the table shows that village-level data is missing for five villages in wave one baseline, four villages in the wave two endline, and two villages in the second round of the wave one endline.

Since the village-level survey was separate from the household-level survey, the missing village level surveys do not induce aggregate attrition of individuals out of the sample. In other words, we still have individual level indicators from the household survey for respondents in villages with missing village-level surveys.

## **Text 2: Pre-Analysis Plan**

Our Pre-analysis Plan (PAP) was registered at the EGAP Repository December 12, 2012. It can be found here: <http://egap.org/registration/622>. As in (4), we finalized the plan while the endline data was 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.

Owing to differences in data collection across the two waves, we developed a modified pre-analysis plan for the subsequent waves. But in pooling data across the first and second waves, we followed the aggregation specified in the plan for the first wave, which was developed before we had access to any endline data.

## **Changes to Pre-Analysis Plan**

Table S7 shows the 16 hypotheses that were specified in our PAP. We examine all of these hypotheses, presenting 6 in the main tables and 10 in the supplemental tables. However, we had to change the way we analyzed some of the indicators from the pre-specified approach. The table marks which hypotheses were changed. Most of these changes were motivated by issues of attempting to aggregate conditional and unconditional outcomes, while others were motivated by changes in how we asked social network questions over time. We discuss each of these issues and detail the associated changes to the hypotheses.

### ***Aggregating Conditional and Unconditional Outcomes***

As in (4) and (6), some of our original hypotheses combined 'conditional' outcomes that relate to a sub-sample of respondents, with 'unconditional' outcomes, that relate to the full sample. However, this could create a

bias in the index using the approach of (52) due to sample selection, and may induce false rejection of the null in the (51) index, if the two sample sizes differ substantially. For example, in measuring trust, all respondents answered how much they trust people in general, but only respondents who personally knew former rebel combatants answered how much they trusted these individuals, which was less than half the sample.

The approach of (52) would create a composite trust index restricted to this latter half. But the selected sample of individuals who know ex-combatants could have different levels of generalized trust, and, their trust levels may also respond differently to treatment. So, this could produce a biased treatment effect that is not representative of the full sample. The approach of (51) 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:

- 1) **Trust (H10)**. Our original trust index proposed to group together four indicators of trust in community members and villagers generally, which were administered to all respondents, along with indicators of trust of specific groups including rebel ex-combatants and migrants, as well as former members of the Sierra Leonean Army (SLA) and the Civil Defense Forces (CDF), all of which were administered only to the individuals who reported knowing members of these groups. Thus, we instead created an index of general trust using the four questions on community-wide trust that were administered to all respondents. We also examined trust in the sub-groups separately, focusing on the RUF rebels, who committed the vast majority of atrocities during the war, as well as migrants, some of whom are former rebel combatants who moved after the war, and who, as a group, are typically socially marginalized within the villages where they reside. For completeness, we also separately examined trust in the CDF and trust in the SLA (in Table S9), though the CDF in particular are unlikely to be the focus of reconciliation processes since they were revered for their role in defending the civilian population against the rebels. One disadvantage of having to conduct multiple tests for each of these sub-groups is that it increases the risk of falsely rejecting true null hypotheses. However, Tables S8 and S9 shows that our results are not meaningfully affected by accounting for these multiple comparisons.
- 2) **Forgiveness (H2)**. These same aggregation of conditional and unconditional issues apply to this hypothesis. The PAP proposed to aggregate across three sub-groups: those who reported being hurt during the war, those who reported personally knowing the perpetrator, and those whose perpetrator still resides in the village. The sample sizes of the final two subgroups are too small for any useful analysis. We therefore separately examine forgiveness of all people in the first group, those who reported being hurt during the war.
- 3) **Conflict Resolution (H8) and Traditional Authority Entrenchment (H9)**. Similar issues apply to our analysis of conflict resolution (H8), which we present as auxiliary results in Supplementary Table S25. Our original PAP proposed to aggregate the proportion of conflicts resolved, as well as the conflicts resolved satisfactorily, as well as the conflicts resolved without a third party. Conceptually, the latter two are conditional on the former, since only conflicts that were resolved could have been resolved via a particular method. Moreover, resolution without a third party displays no variation in the second wave baseline control group (as there were no such types of resolution). This makes it infeasible to create a mean index with this variable since the index scales by the control group standard deviation.

In addition, a separate grouping around Traditional Authority Entrenchment (H9) proposed to look at the proportion of conflicts resolved by chiefs as well as fines levied by the chiefs, but fines apply to a much broader set of cases than conflicts resolved by chiefs, which are only conditional on conflicts occurring in the village and those conflicts getting resolved. Given these issues, we present all five of the indicators individually, without aggregating them.

Since we observed that there was an increase in the fraction of conflicts resolved by chiefs, we thought it would be informative to assess if there were decreases in conflict resolution within other categories. So we

additionally examined resolution via mediation with friends and family, along with the other pre-specified indicators (in Table S25). However, this was not a pre-specified outcome.

### *Changes to Social Network Survey Questions*

We also made some changes to the social network questions. The way the questions were asked in the baseline survey was problematic, so the answers displayed very limited variation. As a result, in the endline, we entirely dropped one question that showed least variation: "who would you share a farm boundary with?". We also changed the way we asked the remaining questions so as to improve the variation in responses. In one part of the survey the question was unprompted: The respondent was asked to list the total number of people he or she would: (i) approach for help/advice; and (ii) ask to collect money for them. In another part of a survey the questions were prompted: the respondent was asked to choose, from the 11 others surveyed in the village, whom they: (i) consider a close friend; and (ii) would approach for advice or help. The prompted questions were used to calculate how often a respondent was listed by someone else. Since asking social network questions twice was very time intensive, we no longer asked the prompted question for "who would you ask to collect money?".

This implies two additional changes:

- 4) **Social Divisions (H11).** This index does not include the altered version of the social network question which would only be available cross-sectionally, since all other indicators were available in baseline and endline.
- 5) **Social Networks (H12).** We are only able to examine the social network measures cross-sectionally, from the endline data. We also dropped two of the indicators that were based on prompted questions: the numbers of times that people listed the respondent as someone (i) they would ask to go collect money; (ii) they would share a farm boundary with. All remaining questions are available cross-sectionally.

Finally, we created three dichotomous indicators of variables that are analyzed on a Likert scale (trust of former rebel combatants, trust of migrants, and our measure of PTSD), in order to gauge the magnitude of effects more meaningfully. We did not pre-specify we would do this in the PAP, but state this explicitly in the main paper when we analyzing these outcomes.

## **Text 3: Additional Details on Evaluation Design**

### **Section Matching and Baseline Balance**

We conducted baseline surveys in each study section, and matched them into pairs, stratified by district, using baseline data on exposure to violence, conflict incidence within villages, an index of household assets, economic activity and psychological health. We then randomly assigned one section in each pair to treatment and the other to control.

Table S5 shows balance statistics on main outcome variables at baseline. (To ensure comparability, we present the pared down versions for outcomes where we had limited measures in the wave two baseline.) Most outcomes display balance, with the exception of the trust variables. Statistically, we expect to observe imbalance in some indicators purely by chance. Moreover, the imbalance goes in different directions for the two trust measures, which suggests that these are statistical aberrations. The index of generalized trust is higher while trust of migrants is lower in treatment communities, and these differences are both only significant at the 10% level. We uncover imbalance in one other context. While the overall economic outcomes index is balanced, individual economic indicators within the index also show imbalance, as shown in the second panel of Table S24

Our main specifications already control for the baseline dependent variable, thus accounting for potential confounds stemming from imbalance in that dependent variable. For example, the regression of the generalized trust index will account for imbalance in baseline trust. But if baseline imbalance in trust also affects the estimated

treatment effect on community group participation, or other outcomes, this may be a reason to control for baseline trust in additional regressions.

In Table S19, we present estimates for all outcomes, but controlling for the baseline index of generalized trust and the individual economic indicators that show imbalance in the middle panel of Table S24. Note that we cannot control for trust of migrants (without losing 30% of the sample) since this question is only administered to those who report knowing a migrant. The results in Table S19 verify that controlling for all other imbalanced indicators do not influence our results.

## **Total Treatment Effect**

Our empirical strategy identifies the total treatment effect which stems from both the direct and indirect effects of the reconciliation process. For example, direct effects will arise from those who participated in the bonfire ceremony, as 40% of our respondents reported doing in treatment sections. However, indirect effects may also arise. For example, an individual who attended the bonfire ceremony may develop a more positive outlook on their community, join community groups, and convince friends and family to do the same. Spillovers of this type highlight why the section is the stable unit of treatment, and why it is important to randomize this intervention at this aggregate level, rather than the individual level. They also present a challenge to instrumenting ceremony attendance with treatment assignment, since the spillovers would serve as a potential violation of the exclusion restriction.

## **Text 4: External Validity and Other Related Literature**

### **Local and National Truth and Reconciliation Processes Around the Globe**

Truth and reconciliation processes come in different forms. One common approach is a national Truth and Reconciliation Commission (TRC). To date, 31 post-conflict countries have instituted TRCs. Examples include: Chad, 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. These commissions often have the power to grant blank amnesty to those who confess to crimes, though fears of prosecution may still remain within the population. For example, the Sierra Leonean TRC granted amnesty to all but the highest level commanders and perpetrators of violence, but people still feared they would be prosecuted, which limited their participation in the process (41). (See (85) for more details on the nature and functions of various commissions).

Local-level reconciliation processes such as the one implemented by Fambul Tok shares some key elements with national-level reconciliation processes, such as truth-telling and bringing victims and perpetrators together in the same setting. But these also differ in notable ways. For example, a reconciliation process run by a non-governmental organization is likely to reduce fear of prosecution, and thereby facilitate more broad-based participation. Also, local level processes are more likely to bring together victims and perpetrators from the same villages, facilitating collective acknowledgement of local war-time atrocities. To the extent that this acknowledgement is critical for transforming individuals' views toward their community, and facilitating their willingness to contribute to their communities (47), local reconciliation efforts may spur larger effects on social capital outcomes.

It is worth noting that besides Sierra Leone, several other countries have also conducted local level reconciliation processes. Examples include the Teso Peace and Human Rights Development Initiative in Kenya, the Acholi Religious Leaders Peace Initiative in Uganda, and the South African Center for the Study of Violence and Reconciliation. The National Platform for Peace and Reconciliation (NPPR) has also initiated community-based reconciliation in select states of South Sudan. Given these other local reconciliation efforts, and important commonalities of local and national reconciliation, our study examines a question that holds relevance to a number of other national contexts.

### **Fambul Tok Areas vs. Non Fambul Tok Areas in Sierra Leone**



Within Sierra Leone, our study takes place within the five districts where Fambul Tok operates, so it is also important to consider whether these districts are similar to the rest of the country. We draw on nationally representative data from the 2007 National Public Services Survey to examine this issue. This year was chosen because the survey contains questions on exposure to violence and trust, and because it marks the start of Fambul Tok operations.

As shown in Figures S1 to S3, the average levels of war violence in these districts is comparable to the war violence experienced in the rest of Sierra Leone, including or excluding Western Area (which hosts the capital city of Freetown). In fact, the Fambul Tok districts display considerable diversity in war exposure, including some of the districts most exposed to violence (such as Kono and Kailahun in the east) and some of the districts least exposed to violence (such as Bombali). These figures also show that basic socioeconomic characteristics such as employment in farming, formal education, and household access to toilets are similar across Fambul Tok districts and non Fambul Tok districts, as is trust in other community members, a key social capital measure. Thus our study takes place in districts that are similar to other parts of Sierra Leone along key dimensions.

Table S1 also shows that the sections in the study are similar along these key dimensions to other sections in the five Fambul Tok districts of operation. Although the NPS is a nationally representative survey, it does not sample every section, so this exercise is based on the 57 overlapping sections that are in the NPS and in our study. The coefficients on the indicator of whether the section is in our study are both imprecisely estimated and small in magnitude relative to the mean of the dependent variables, indicating the null effects are not simply driven by power limitations. This further bolsters the external validity of the results to other parts of the country.

## Other Related Literature

Our study examines how reconciliation influences individual healing as measured by psychological health, as well as societal healing as measured by social capital outcomes. Other related work has examined how truth and reconciliation commissions affect democracy (8, 86) and peace (87, 88). For a comprehensive review the impact of TRCs and other transitional justice programs, see (89).

Another related literature has examined how war itself affects social capital outcomes, and found varied results. Some studies have documented that trust and other forms of social capital are lowered by civil wars (90, 91) and other traumatic events (92). Other studies have shown that exposure to war can induce more pro-social behavior in the future (93-95); toward in-groups but not out-groups (96); and, when war is perceived as an external threat to the community (97). Our focus, instead, is whether reconciliation, in the aftermath of war, can influence these outcomes.

## Text 5: Additional Results

In this section, we present additional results, which are also referenced in brief in the main text. The tables are numbered by the order in which they are referenced in the main text.

Table 1-Panel B shows that the reconciliation treatment increased trust of ex-combatants and migrants. These questions were only administered to those who reported knowing members of these groups. The specifications controlling for the baseline dependent variable already restrict the sample to those who knew group members at both baseline and endline, so these effects can't be attributed to compositional changes in who knows ex-combatants and migrants at baseline and endline. Moreover, Table S15 verifies that the treatment doesn't change the likelihood of knowing someone from either of these groups in cross-sectional specifications. This is likely because these are relatively small communities, and so knowing new types of individuals is not the margin through which the reconciliation effect operates. Consistent with this, when we present cross-sectional specifications on trust at the bottom of Table S15 (which are not restricted to those who knew members of these groups at baseline and thus allow compositional changes in the sample), these estimates are statistically indistinguishable at the 5% level from those shown in Table 1.

Table 2 reports that social networks are stronger in reconciliation areas. It is possible that attending the ceremony may have generated friendships through an alternate social channel beyond its impact on community healing. However, we verified that there was no significant differential impact of ceremony attendance on this

outcome. Since attendance is endogenous, we instrumented it using interactions of the treatment indicator and whether the village is a section headquarter. (Since the ceremonies were typically held in headquarter villages, attendance was higher in these villages within treatment sections, giving rise to a strong first stage). In a specification that includes both the treatment indicator and instrumented attendance, the coefficient on attendance captures the additional impact of attending the ceremony, specifically. This coefficient is statistically indistinguishable from zero and negative in sign, suggesting that ceremony attendance and its social impact is not the main driver of the social network effect.

Table S24 examines additional economic outcomes. Here, we observe that the treatment induces a significant increase in the objective measure of household wellbeing, which is an asset index constructed by PCA. In contrast, there are negative effects on the two perceptions-based measures, and the effect is significant for the perceptions of overall household economic situation. This pattern is consistent with the idea that lower psychological wellbeing - such as depression - reduced perceptions of economic wellbeing, even while households in treatment areas experienced economic improvements. Aggregated together, these effects produce a negative impact on the index of economic outcomes. However, as shown in Tables S8 and S9 below, this effect is not significant when we adjust p-values for multiple testing.

Also, as we show in the second panel of this table, individual indicators within the economic outcomes index were imbalanced at baseline (although the aggregate index was not). When we control for these baseline indicators individually, the overall effect is insignificant. In the second part of Table S24, we also examine effects on a number of different economic activities, such as borrowing and lending and farming. While six of eight indicators are positive, the overall effect is not statistically significant. Given the negative coefficient on the household economic outcomes and relatively small positive coefficient on economic activity, our results do not show definitive evidence that the reconciliation process influences economic outcomes.

In Table S25, we examine effects on conflicts. We find no significant impacts on the index of social divisions, which focuses on conflictual divisions among groups. The coefficient on the index is small, .012 standard deviations. To examine effects on day-to-day conflicts, we sum the number of crimes and disputes over matters such as loans and land at the village level; we add this to a measure of inter-village disputes to construct a mean index. We look at this index cross-sectionally since we are missing surveys in a number of villages at baseline. Among the individual indicators, there is a positive effect on the number of inter-village conflicts. When we examine this outcome separately by wave, we observe that the effect is only present in wave one. The effect sizes are 0.15 (standard error of 0.054) and 0.08 (standard error of 0.068) in waves one and two respectively. When we disaggregate the separately over the short versus long run (Table S26), we observe the effect only in the short-run of wave one. Thus it is possible that when the reconciliation ceremony brings people together from different villages, this reignites some dormant inter-village disputes, which then get resolved or die down after 9 months. However, we place a limited emphasis on this interpretation since this result is localized to one wave. Also, the impact on the overall conflict index is statistically insignificant. Overall, these results on conflicts, crimes and tensions suggest that while reconciliation can facilitate forgiveness for violence committed during the war, it can't necessarily mitigate other types of conflicts between households and groups, or prevent people from committing crimes in their community.

While conflict incidence did not change, there appeared to be some compositional changes in how conflicts were resolved. A larger fraction of conflicts were resolved by the chiefs in treatment areas. Chiefs were often targeted by youths in their communities during the war. If the reconciliation process restored these relationships, this may influence the degree to which individuals rely on chiefs for conflict resolution. At the same time, a smaller fraction were resolved through friends and family. However, this was not an indicator we pre-specified we would examine.

In the last panel of Table S25 we examine attitudes related to gender. Our index of attitudes toward women captures views on domestic violence as well as the rights of wives to express their opinions. The wave two baseline didn't include the opinion question about wife beating, and only included three of the seven questions about domestic violence that are aggregated together on a Likert-scale, so here, we include the pared down dependent variable controls. We find that reconciliation communities experienced significant improvements in this index of gender attitudes. These outcomes could shift because the Peace Mothers groups promote female empowerment, or because the reconciliation process highlighted challenges faced by women during the war. However, this effect is

not as robust to controls as other estimates (see Table S19 and S23). For completeness, we also look at just the subset of the three questions about domestic violence (aggregated on the Likert scale), but also find no significant effect for this indicator only.

Table S26 examines the persistence of these additional outcomes over the first and second rounds of the Wave 1 communities. These results suggest that the positive impact on household assets is not restricted to the short-run. And, the negative impact of the two subjective indicators emerge across different rounds, which again shows the inconsistency in the impact of the perceptions measures. The gender attitude index is also stronger in the short run and appears to diminish over time.

In Table S23, we control for the presence of FT communal farms, since some treated areas had farms while others did not. Of course this is an over-control since it controls out for our treatment. Yet, none of the estimated effects are rendered insignificant in a meaningful manner with the inclusion of this control. The coefficient on trust of ex-combatants becomes insignificant but this coefficient is not statistically distinguishable at the 5% level from the estimate in Table 2. Overall, this suggests that the treatment effects do not stem primarily from the presence of communal farms in treatment communities.

Next, we examine heterogeneous treatment effects based on gender and history of violence. It is only meaningful to examine differential effects of individual characteristics on outcomes that vary at the individual level (versus at the household or village level). Since several of the economic activity variables were at the village level, we created another index of just individual economic activities. Also, assets are household level measures; but the subjective perceptions are asked of individuals, so we are also able to examine these outcomes.

Table S22 reports differential effects by gender. None of the coefficients on the Treatment x Female interactions are significant, with the exception of the borrowing and lending measure. These insignificant effects may reflect our limited power to detect heterogeneous effects. However, if the Peace Mother's Group were a key driver of impacts in treated areas, we would expect to observe larger treatment effects for women, especially on economic activities. In contrast, we observe a significantly *smaller* treatment effect for females on borrowing and lending (and an insignificant negative coefficient on the overall economic activity index). This table presents evidence against that account.

Next we examine heterogeneous effects based on exposure to violence. Theoretically, two effects are possible. On the one hand, someone who has experienced more violence may have a greater need for reconciliation, which implies that treatment effects should be larger for these individuals. On the other hand, a violent experience may also mean that they have more to forgive (as the baseline Table S10 shows); and this implies that the treatment effects should be smaller.

Table S20 presents estimates with two measures of violence victimization. In Panel A, an individual is violence-exposed if they were raped, maimed, beaten, saw violence or abducted. This is the definition of violence exposure we pre-specified in our PAP. However, it is arguably important to incorporate whether the person had a family member killed as a measure of violence exposure (since this seems to be an important determinant of baseline forgiveness and psychological wellbeing – see Table S14). Moreover, it is arguably important to check the robustness to leaving out abductions since this may overlap with the tendency to be an ex-combatant, since some ex-combatants were kidnapped by the RUF. In Panel B, an individual is violence-exposed if they were raped, maimed, beaten, saw violence or had a family member killed. Both panels show that there are no significant differential impacts. Also, coefficients on the interaction terms are typically small in magnitude when we look at the indices. This is consistent with the theoretical idea that violence may have two potentially offsetting effects. Also, in Table S21, we find no differential treatment effects associated with the subset of individuals who we are able to identify as ex-combatants.

We also undertake three additional robustness checks.

Since our PAP did not pre-specify that the specification with controls for the baseline dependent variable would constitute our preferred specification, we also estimate all of our main results with a cross-sectional specification, even outcomes where we have the baseline dependent variable (Table S18). The results only affect

one of our effects – the index on public goods contributions (SE=0.022, P=.0252). However, this estimate is statistically indistinguishable at the 5 percent level, from the estimate with the baseline dependent variable control.

Table S6 shows the results for the mean effect indices as constructed by the Anderson methodology. The only effect that changes in terms of statistical significance is the index for public goods contributions, which arises from a large loss of observations owing to missingness in indicators within the index. Each individual-level indicator is missing for some respondents, and aggregating this non-response over respondents produces a sizable loss on overall observations. This is precisely the circumstance under which it is useful to utilize some form of imputation as the KLK methodology does. Moreover, the public contributions index also includes the number of community projects indicator from the village level survey that is missing for 12 villages, which imposes further missingness on the overall index, and creates a different including only 188 of the 200 villages in the remainder of the study. As shown in Table S6, when the community projects variable is removed from the index, the effect on the contributions index is larger and more precisely estimated. However, even with the community projects variable included, the estimated effect for this outcome under the Anderson approach is also not statistically distinguishable from the estimated effect under the KLK approach (in Table 4) at the 5 percent level.

Finally, in Table S17, we re-do analysis for all mean effect indices constructed via the KLK approach which include a variable from the village survey (public goods provision, economic activity, and the conflict and crime index), but re-create the index *excluding* the village level variables. The odd-numbered columns show the main results from the original index including all the variables, and the even-numbered columns show the results for the new mean index that exclude the village-level variable. Note that for the public goods contributions, there is a slight reduction in the total number of observations under the new index. This is because there are 22 observations that are missing for all the variables in the reduced index that were non-missing for the village-level variable. Since missing values are imputed for observations if there is at least one indicator with a response, these observations are treated as missing in the reduced index, but not the full index. The effects on economic activity and conflicts and crime remain insignificant under the reduced index, while the coefficient on public goods contributions remains of a similar size and significant at the 10% level.

## Text 6: Multiple Comparison Corrections

Our study examines impacts on multiple outcomes which 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 False-Discovery rate (FDR) using the Benjamini-Hochberg method (54).

The FDR is the expected proportion of falsely rejected null hypotheses (or "false discoveries"). FDR-controlling procedures are advocated by many as the most appropriate means of accounting for multiple comparisons (54-56, 98), and have been used by recent experimental studies in the social sciences, such as (50).

The Benjamini-Hochberg method requires selecting a threshold rate, the *FDR*, and ranking  $n$  hypotheses in a family from the highest to the lowest *p-value*. The  $i^{\text{th}}$  hypothesis is then assigned the false discovery rate critical significance level of:  $d_i = FDR * (i/n)$ . For example, if we select a *FDR* of 5% 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* <sub>$i$</sub>  <  $d_i$ .

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. First, we apply the FDR controlling procedure to all of the hypotheses as they were grouped together in the sections of our Pre-Analysis Plan. These groupings were thematically related. For example, forgiveness and psychological wellbeing were grouped together since they were both psychology-related outcomes. Table S8 shows these results. We ranked 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 false discovery rates 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 .0003, which is significant at the 1% level, in unadjusted terms). The maximal penalty is applied to this outcome, yielding an adjusted critical significance level of .001 for a FDR of 1% (which is a tougher criteria to meet than the standard .01 for unadjusted 1% significance). Yet  $.0003 < .001$ , so the effect remains significant at the 1% level, even with this adjustment. Analogously, our index of contributions was only significant at the 10% level (in unadjusted terms) with a p-value of .055. As the table shows, it continues to remain significant under a FDR of 10% (after facing an adjusted critical significance level of .056 instead of the standard .10 for unadjusted 10% significance). This table shows that effects on social capital remain in place even after we include general trust as well as measured trust toward all various sub-groups.

Table S8 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 significant at the 10% level when considered in unadjusted terms, but do not remain so after we account for a FDR of at least 10%. This doesn't necessarily mean that there are no true effects, but simply, we cannot reject the null hypothesis of no impact after accounting for multiple comparisons.

While the results above are presented for groupings under the PAP, one could argue that a different grouping would be relevant for conceptualizing families. Most notably, we posit a close conceptual relationship between forgiveness and social capital outcomes in describing potential mechanism under our section on "Healing under Reconciliation." Under this conceptualization, social networks and associated norms such as trust, and 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 alternate grouping in Table S9. Panel A shows that our results are again unaffected. Finally, in Panel B, we further incorporate psychological wellbeing 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 do not influence our findings.

**Table S1. Characteristics of Study Sections vs. Other Sections.** This table uses data from the National Public Service (NPS) 2007 survey to compare the sections in our study to other sections in Fambul Tok's five districts of operation. The NPS data includes 57 of the 100 Fambul Tok sections. Data is collapsed to the section level. Each row is a separate regression which regresses the NPS survey outcome on an indicator of whether the section is in the study along with district fixed effects. Sections hosting the headquarter city of the district are not included in these regressions. The second row is the mean value in the sections that are not in our study. Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

NPS Outcome Variable	Mean in other sections	Coeff.	Std. Error	Obs.	R-sqr.
Household member injured/maimed	0.504	-0.010	(0.068)	176	0.047
Household member killed	0.525	-0.059	(0.063)	176	0.168
Household farms	0.887	0.051	(0.035)	175	0.052
Household head has no education	0.78	-0.040	(0.062)	176	0.077
Household access to own toilet	0.264	0.052	(0.064)	170	0.115
Trust people from own community	0.796	0.065	(0.048)	176	0.141

**Table S2. Attrition.** 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 one. Overall Attrition-Broad equals one if the respondent is missing from either endline round in wave 1 or the endline in wave 2. Overall-Attrition-Narrow equals one if 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.

	(1) Attrition-Round1	(2) Attrition-Round 2	(3) Attrition	(4) Overall-Attrition-Broad	(5) Overall-Attrition-Narrow
Treatment	0.017 (0.011)	0.022 (0.017)	-0.007 (0.014)	0.010 (0.012)	-0.003 (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

**Table S3. Missing Village-level Surveys.** Each column represents a separate regression of an indicator of missing of whether the village-level survey is missing on treatment assignment. The indicator in the first column equals 1 if the village survey is missing in baseline (5 villages). The indicator in the second column equals 1 if data is missing in the endline (a different 6 villages). The indicator in the third column equals 1 if the village-level data is missing in either baseline or endline (11 villages). The coefficient in the second column is exactly 0, because there are an identical number of missing villages in treatment and control. 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.

	(1)	(2)	(3)
VARIABLE:	Village Survey Missing in Baseline	Village Survey Missing in Endline	Village Survey Missing in Either Baseline or Endline
Treatment	-0.030 (0.021)	-0.000 (0.020)	-0.030 (0.024)
Observations	200	200	200
R-squared	0.343	0.227	0.355



**Table S4. Sample size by wave and round.** Table displays number of individuals and villages surveyed in each wave and round of data collection.

<b>Panel A. Individual Survey</b>		
	<b>Baseline Wave 1</b>	<b>Baseline Wave 2</b>
Target number of individuals	960	1440
Surveyed number of individuals	952	1430
	<b>Endline Wave 1 Round 1</b>	<b>Endline Wave 2</b>
Target number of individuals	952	1430
Surveyed number of individuals	885	1273
	<b>Endline Wave 1 Round 2</b>	
Target number of individuals	952	
Surveyed number of individuals	850	
<b>Panel B. Village Survey</b>		
	<b>Baseline Wave 1</b>	<b>Baseline Wave 2</b>
Target number of villages	80	120
Number villages surveyed	75	120
	<b>Endline Wave 1</b>	<b>Endline Wave 2</b>
Target number of villages	80	120
Number villages surveyed	80	116
	<b>Endline Round 2</b>	
Target number of villages	80	
Number villages surveyed	78	

**Table S5. Baseline Balance.** 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.

VARIABLES	T-C	Std. Error	Obs.
Market in village	0.021	(0.034)	2,171
Village size	-8.796	(32.413)	2,171
No formal education	0.011	(0.015)	2,199
Forgive perpetrators	-0.107	(0.153)	1,862
Ex-combatants would not Fight	0.011	(0.034)	2,191
Trust of Rebel Ex-combatants	-0.022	(0.054)	1,546
Trust of Migrants	-0.059*	(0.032)	1,962
Trust of former CDF	-0.068	(0.049)	1,652
Trust of former SLA	-0.073	(0.049)	1,552
Index of Generalized 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 Wellbeing	0.023	(0.038)	2,202
Attitude toward 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

**Table S6. Impacts using Indices as Constructed by Anderson (2008).** Each row represents a separate regression of the outcome shown in the first column on treatment assignment. Variables not shown in all regressions include section pair fixed effects and the second round indicator. The final column indicates if 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.

VARIABLES	Coeff.	Std. Error	Obs.	R-sqr.	Baseline dependent variable controls?
Index of Attitudes toward Ex-Combatants	-0.004	(0.029)	2,960	0.075	Y
Index of War Attitudes	-0.014	(0.028)	3,000	0.045	N
Index of Generalized Trust in Community	0.015	(0.029)	2,915	0.124	Y
Index of Network Strength	0.076**	(0.029)	3,005	0.113	N
Index of Participation in Community Groups	0.035**	(0.017)	1,930	0.159	Y
Index of Public Goods Contributions	0.019	(0.026)	1,809	0.223	Y
Index of Public Goods Contributions (Indicators in both baselines)	0.021	(0.026)	1,817	0.219	Y
Index of Public Goods Contributions (Without village-level variable)	0.042*	(0.023)	1,904	0.138	Y
Index of Psychological Wellbeing (All indicators)	-0.143***	(0.034)	2,635	0.121	Y
Index of Psychological Wellbeing (Indicators in both baselines)	-0.133***	(0.031)	2,667	0.120	Y
Index of Economic Outcomes	-0.039*	(0.020)	2,831	0.134	Y
Index of Economic Activity	-0.020	(0.035)	1,872	0.253	Y
Index of Social Divisions	0.026	(0.022)	2,709	0.088	Y
Index of Conflict and Crime	-0.001	(0.064)	273	0.261	N
Index of Attitudes toward Women	0.044*	(0.026)	2,920	0.038	Y

**Table S7. Summary of Hypotheses in Pre-analysis Plan.** The first two columns list the number and hypothesis as listed in the PAP. The third column lists the table number in the main paper or supplement in which the hypothesis is tested. The fourth column indicates if there was a change in how the hypothesis was examined relative to what was pre-specified. The fifth column provides a brief description of the reason for the change, each of which is detailed in the SOM Text 2.

No.	Hypothesis	Table	Change?	Reason for Change
1	Program implemented according to objectives	S9		
2	Forgiveness	1	Y	Original aggregated conditional outcome
3	Psychological Wellbeing	5		
4	Attitudes toward Ex-Combatants	S16		
5	War Attitudes	S16		
6	Attitudes toward Gender	S25		
7	Incidence of Conflict and Crime	S25		
8	Conflict Resolution	S25	Y	Original aggregated conditional outcome
9	Entrenchment of traditional sources of power	S25	Y	Original aggregated conditional outcome / no control group variation in 1 indicator
10	Trust	1	Y	Original aggregated conditional outcome
11	Social Divisions	S25	Y	Change to social networks question
12	Social Networks	2	Y	Change to social networks question
13	Group Membership	3		
14	Public Goods Provision	4		
15	Economic Activity	S24		
16	Economic Outcomes	S24		

**Table S8. Adjusting for Multiple Comparisons.** This tables adjusts for multiple comparisons through a FDR controlling procedure, applied to the conceptually related hypotheses around forgiveness and social capital (Panel A) as well as psychological wellbeing (Panel B). Each hypotheis 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 a FDR of 1%, ++ is significant with a FDR of 5%, + is significant with a FDR of 1%.

Comparison <i>i</i>	Variable	Coef.	P-value	Adjusted critical level of significance ( $d_i$ )		
				FDR = 0.01	FDR = 0.05	FDR = 0.1
<b>Panel A. Forgiveness and Social Capital</b>						
9	Index of Generalized Trust in Community	0.006	0.816	0.010	0.050	0.100
8	Trust of former SLA	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 Contributions to Public Goods	0.042+	0.055	0.007	0.033	0.067
5	Trust Rebel Ex-combatants	0.177++	0.027	0.006	0.028	0.056
4	Forgive Perpetrators	0.571++	0.0134	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
<b>Panel B. Forgiveness, Psychological Wellbeing, and Social Capital</b>						
10	Index of Generalized Trust in Community	0.006	0.816	0.010	0.050	0.100
9	Trust of former SLA	0.019	0.783	0.009	0.045	0.090
8	Trust of former CDF	0.029	0.638	0.008	0.040	0.080
7	Index of Contributions to Public Goods	0.042+	0.055	0.007	0.035	0.070
6	Trust Rebel Ex-combatants	0.177++	0.027	0.006	0.030	0.060
5	Forgive Perpetrators	0.571++	0.0134	0.005	0.025	0.050
4	Index of Network Strength	0.112+++	0.002	0.004	0.020	0.040
3	Index of Community Group Participation	0.058+++	0.001	0.003	0.015	0.030
2	Trust Migrants	0.123+++	0.0003	0.002	0.010	0.020
1	Index of Psychological Wellbeing	-0.147+++	0.00002	0.001	0.005	0.010

**Table S9. Adjusting for Multiple Comparisons with Pre-Analysis Plan Groupings.** This tables adjusts for multiple comparisons through a FDR controlling procedure, based on the groups in the pre-analysis plan. Each hypotheis 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 a FDR of 1%, ++ is significant with a FDR of 5%, + is significant with a FDR of 1%.

Comparison <i>i</i>	Variable	Coef.	P-value	Adjusted critical significance level ( $d_i$ )		
				FDR = 0.01	FDR = 0.05	FDR = 0.1
<i>Social Capital</i>						
9	Index of Generalized Trust in Community	0.006	0.816	0.01	0.050	0.100
8	Trust of former SLA	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
<i>Forgiveness and Psychological Well-being</i>						
2	Forgive Perpetrators	0.571++	0.0134	0.01	0.050	0.100
1	Index of Psychological Wellbeing	-0.147+++	0.00002	0.005	0.025	0.050
<i>Attitude and Beliefs</i>						
3	Index of Attitudes toward 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 toward Women	0.044	0.074	0.003	0.017	0.033
<i>Conflict and Conflict Resolution</i>						
7	Fined by chief	-0.006	0.527	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	Satisfactory 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
<i>Economic Activity and Welfare</i>						
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

**Table S10. List of Forgiveness Questions.** These questions are answered on a 4-point Likert scale, where the responses are strongly agree, agree, disagree and strongly disagree. \* Questions are reversed scored.

---

---

Respondent is asked how much s/he agrees with the following statements:

***Cessation of Negative***

Do you think constantly about how you were wronged by this person? \*

Do you wish for good things to happen to the person who wronged you?.

Do you spend time thinking about ways to get back (get revenge) at the person who wronged you? \*

Do you feel bad whenever you think about the person who wronged you? \*

Do you feel hatred whenever you think about the person who wronged you? \*

Do you avoid certain people and/or places because they remind you of the person who wronged you? \*

Have these person's wrongful actions kept you from enjoying your life? \*

Do you become depressed whenever you think of how you were mistreated by this person? \*

***Presence of Positive***

Is the person who wronged you worthy of respect?

Would you help this person if this person were in need?

If you had the chance, would you be willing to be friends with the person who wronged you?

Have you been able to let go of your anger towards this person?

---

**Table S11. List of Psychological Wellbeing Questions.** These questions are answered on a 4-point Likert-scale where the responses are never, yes small small, yes sometimes and yes often. All questions except those indicated by \* are reversed scored.

---

### Post-Traumatic Stress Disorder

---

#### *Rexperience*

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?

In the last month, did you have some bad bad dreams from all you have been through?

In the last month, can it look like the bad things are just happening again?

#### *Arousal*

In the last month, have you easily gotten angry at people?

In the last month, when you were thinking about bad things, is it difficult for you to concentrate on what you are doing (enumerator instructions: like when you are trying to work)?

In the last month, when some noise happens, can it make you jerk quick quick?

#### *Avoidance*

In the last four weeks, have you sometimes tried to stay away from places or certain people or talk about things because it makes you think about bad things that happened to you?

In the last month, how often did you feel like keeping to yourself?

In the last month, do you not feel like doing the things you usually like to do with friends?

In the last month, did you ever not want to plan for your future?

In the last month, how much did you feel you had no hope for the future?

---

### Depression

---

In the past month, did you feel down-hearted and blue?

In the last month, how much did you cry?

In the last month, how much difficulty have you had falling asleep or sleeping through the night?

In the past month, did you eat as much you used to (i.e. no loss in appetite)? \*

In the last month, how much did you have difficulty making decisions?

In the last four weeks, how often have you felt like you not important to nobody?

In the last month, how often did you think that other people would be better off if you were dead?

---

### Anxiety

---

In the last month, how much did you feel nervous or anxious or worried?

In the last month, did you feel fear without cause?

In the last month, did you often feel upset or feel sudden panic?

In the last month how often have you felt that everything is alright and nothing bad will happen in the future? \*

In the last month, how much did your legs and arms shake and tremble?

In the last month, how much did you have headaches or pain in your neck and back?

In the last four weeks, how often have you felt tired even if you not doing nothing?

During the last month, how much did you feel restless?

In the last month, has your heart been pounding fast?

In the last month, how much were you bothered by pain in your stomach?

---



**Table S12. Descriptive Statistics**

VARIABLES	Obs.	Mean	Std. Dev	Min	Max
<i>Individual and Village Characteristics (Baseline):</i>					
Gender	2212	0.550	0.498	0	1
No formal education	2208	0.717	0.451	0	1
Occupation farmer	2178	0.746	0.436	0	1
Market in village	2075	0.085	0.279	0	1
Village size (number households)	2158	183.763	330.126	0	2811
Beaten	2097	0.329	0.470	0	1
Maimed	2099	0.020	0.138	0	1
Raped	2092	0.030	0.170	0	1
Family member killed	2157	0.535	0.499	0	1
Saw Violence	1749	0.440	0.496	0	1
Ex-combatant	2110	0.052	0.227	0	1
<i>Panel Outcomes (Baseline and endline):</i>					
Forgive Perpetrators	4296	-0.217	3.901	-10.5	10.5
Ex-combatants would not Fight	5191	2.552	0.794	1	4
Trust of Rebel Ex-combatants	3016	1.823	0.981	1	4
Trust of Migrants	4484	3.074	0.775	1	4
Trust of former CDF	3490	2.460	1.081	1	4
Trust of former SLA	3051	2.435	1.058	1	4
Index of Generalized Trust	5212	0.014	0.686	-1.987	1.932
Index of Community Group Participation	5218	0.011	0.429	-0.573	2.437
Attitude toward Wife Beating	5185	10.468	1.930	4	12
Index of Economic Outcomes	5222	-0.017	0.555	-1.545	6.350
Index of Economic Activity	5222	0.002	0.481	-1.052	11.839
Index of Group Tensions	5212	0.004	0.581	-2.778	1.470
Index of Psychological Wellbeing	5205	-0.035	0.839	-5.506	1.907
Less PTSD	5067	26.769	5.746	0	33
Less Anxiety	5141	13.356	3.929	0	21
Less Depression	5158	10.988	2.380	0	15
<i>Cross-sectional Outcomes (Endline):</i>					
Forgive Perpetrators	2434	2.502	5.408	-18	18
Index of War Attitudes	3000	-0.011	0.692	-1.675	2.526
If another war, people would not fight	3000	0.770	0.421	0	1
People would not be a part of another rebellion	3000	0.838	0.368	0	1
If another war, you would not fight	3000	0.045	0.207	0	1
Index of Network strength	3008	0.047	0.817	-1.144	27.597
Number of people respondent would approach for advice / help	3005	2.961	2.193	0.000	47.000
Number of people respondent would ask to collect money	3005	3.214	5.223	0	244
Number of times respondent listed by others as good friend	3008	2.236	2.023	0	13
Number of times respondent listed by others for advice / help	3008	3.419	2.887	0	16

**Table S13. Program Implementation.** 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.

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 Mother's Group	0.406***	(0.046)	3,008	0.381

**Table S14. War Exposure, Baseline Forgiveness and Psychological Health.** 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.

VARIABLES	(1) Forgive Perpetrators	(2) Less PTSD	(3) Less Anxiety	(4) Less Depression
Raped	-1.211** (0.544)	-2.358*** (0.849)	-0.512 (0.575)	-0.529 (0.462)
Observations	1,470	1,918	1,986	1,999
Maimed	-0.564 (0.803)	-2.471*** (0.928)	-1.193** (0.494)	-0.613 (0.536)
Observation	1,475	1,925	1,990	2,005
Beaten	-0.407 (0.259)	-1.715*** (0.311)	-0.416* (0.233)	-0.272* (0.158)
Observations	1,481	1,925	1,990	2,001
Family member killed	-0.920*** (0.232)	-1.140*** (0.286)	-0.402** (0.195)	-0.330*** (0.120)
Observations	1,500	1,972	2,039	2,056

**Table S15. Robustness: Trust of Ex-Combatants and Migrants.** See Table S6 for notes.

VARIABLES	Coeff.	Std. Error	Obs.	R-sqr.	Baseline dependent variable controls?
Do you know any rebel ex-combatants?	-0.034	(0.024)	2,970	0.186	Y
Do you know any former SLA members?	-0.021	(0.024)	2,950	0.206	Y
Do you know any former CDF members?	-0.042	(0.025)	2,970	0.171	Y
Do you know any migrants?	-0.012	(0.013)	3,008	0.116	Y
Do you know any rebel ex-combatants?	-0.024	(0.024)	3,000	0.179	N
Do you know any former SLA members?	-0.017	(0.024)	3,000	0.198	N
Do you know any former CDF members?	-0.036	(0.025)	3,000	0.166	N
Do you know an migrants?	-0.011	(0.013)	3,008	0.109	N
How much do you trust rebel ex-combatants?	0.145**	(0.066)	1,470	0.177	N
How much do you trust former CDF members?	0.036	(0.054)	1,838	0.227	N
How much do you trust former SLA members?	0.101	(0.064)	1,499	0.258	N
How much you trust migrants to this community?	0.083***	(0.032)	2,522	0.167	N

**Table S16. Attitudes Related to War.** See Table S6 for Notes.

VARIABLES	Control mean	Coeff.	Std. Error	Obs.	R-sqr.	Baseline dependent variable controls?
<i>Index of Attitudes toward Ex-Combatants</i>	0	-0.007	(0.029)	2,980	0.075	Y
Indicators:						
Those who did bad things during the war would do it again	2.582	0.018	(0.030)	2,966	0.060	Y
Rebels are not responsible for their actions	2.832	-0.025	(0.029)	2,966	0.089	Y
<i>Index of War Attitudes</i>	0	-0.024	(0.030)	3,000	0.057	N
Indicators:						
If another war, people would not fight	0.780	-0.023	(0.016)	3,000	0.099	N
People would not be a part of another rebellion	0.853	-0.030**	(0.015)	3,000	0.070	N
If another war, you would not fight	0.038	0.013*	(0.007)	3,000	0.040	N

**Table S17. Accounting for missing village-level variables.** Each column represents a separate regression. All specifications include section pair fixed effects. Specifications including baseline dependent variable controls also include the second round indicator, the baseline outcome variable, and its interaction 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.

VARIABLE:	Index of Public Goods Contributions		Index Public Goods Contributions (excluding women's groups)		Index of Economic Activity		Index of Conflicts and Crime	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treatment	0.042*	0.040*	0.046**	0.043*	0.034	0.013	0.112	0.030
	(0.022)	(0.021)	(0.023)	(0.022)	(0.026)	(0.021)	(0.072)	(0.071)
Exclude village-level variables		Y		Y		Y		Y
Baseline dependent variable controls?	Y	Y	Y	Y	Y	Y		
Observations	3,008	2,986	3,008	2,986	3,008	3,008	274	274
R-squared	0.171	0.138	0.184	0.149	0.182	0.156	0.275	0.246

**Table S18. Comparison with or without controlling for baseline dependent variable.** Each row represents a separate regression of the outcome shown in the first column on treatment assignment. All specifications include section pair fixed effects and the second round indicator. The first two columns report regressions that include as a control the baseline outcome variable, and its interaction with both the second round indicator and the second wave indicator. The second two columns report regressions using the cross-sectional specification. 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.

<b>Baseline Dependent Variable Controls?</b>	<b>YES</b>		<b>NO</b>		<b>Obs.</b>
	<b>Coeff.</b>	<b>Std. Error</b>	<b>Coeff.</b>	<b>Std. Error</b>	
Forgive perpetrators	0.571**	(0.227)	0.483**	(0.238)	2,010
How much do you trust ex-combatants?	0.177**	(0.079)	0.189**	(0.082)	900
How much you trust migrants to this community?	0.123***	(0.033)	0.118***	(0.033)	2,203
Index of Generalized Trust in Community	0.006	(0.027)	0.008	(0.028)	2,996
Index of Participation in Community Groups	0.058***	(0.017)	0.052***	(0.018)	3,008
Index of Public Goods Contributions	0.042*	(0.022)	0.026	(0.022)	3,008
Index of Psychological Wellbeing (All indicators)	-0.147***	(0.033)	-0.145***	(0.033)	2,982
Index of Economic Outcomes	0.034	(0.026)	0.026	(0.027)	3,008
Index of Economic Activity (Individual level)	-0.036*	(0.019)	-0.034	(0.021)	3,008
Index of Social Divisions	0.028	(0.021)	0.027	(0.022)	2,996
Index of Attitudes toward Women	0.044*	(0.025)	0.041*	(0.024)	2,982
Resolved	-0.057	(0.050)	-0.045	(0.050)	172
Satisfactory resolved	-0.107	(0.067)	-0.100	(0.062)	172
Resolved without third party	-0.036	(0.035)	-0.037	(0.033)	172
Resolved with mediation from family/friends	-0.141**	(0.055)	-0.136**	(0.053)	172
Resolved by chief	0.103*	(0.058)	0.100*	(0.055)	172
Fined by chief	-0.006	(0.009)	-0.007	(0.009)	280

**Table S19. Controlling for Baseline Imbalance.** Each row represents a separate regression of the outcome shown in the first column on treatment assignment. Indeces are constructed using the Kling et al. (2007) methodology. All regressions include section pair fixed effects and the second round indicator. The final column indicates if 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 (See Table S17). 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.

VARIABLES	Coeff.	Std. Error	Obs.	R-sqr.	Baseline dependent variable controls?
Forgive Perpetrators	0.548**	(0.239)	1,919	0.143	Y
Trust Rebel Ex-combatants	0.222***	(0.076)	860	0.240	Y
Trust Migrants	0.110***	(0.034)	2,084	0.181	Y
Index of Generalized Trust in Community	0.003	(0.024)	2,832	0.145	Y
Index of Attitudes toward Ex-Combatants	-0.012	(0.030)	2,818	0.085	Y
Index of War Attitudes	-0.025	(0.029)	2,831	0.062	N
Index of Network Strength	0.130***	(0.035)	2,839	0.077	N
Index of Community Group Participation	0.060***	(0.018)	2,836	0.174	Y
Index of Contributions to Public Goods	0.044*	(0.022)	2,839	0.180	Y
Index of Psychological Wellbeing	-0.142***	(0.033)	2,820	0.130	Y
Less PTSD	-0.664***	(0.196)	2,628	0.135	Y
Less Anxiety	-0.395***	(0.118)	2,738	0.149	Y
Less Depression	-0.279***	(0.069)	2,752	0.097	Y
Index of Economic Outcomes	-0.026	(0.020)	2,839	0.228	Y
Assets	0.149***	(0.054)	2,836	0.418	Y
Perception of household needs	-0.115	(0.071)	2,835	0.102	Y
Perception of economic situatoin	-0.130***	(0.037)	2,831	0.088	Y
Index of Economic Activity	0.037	(0.026)	2,839	0.192	Y
Index of Social Divisions	0.030	(0.022)	2,832	0.094	Y
Index of Conflicts and Crime	0.117	(0.083)	259	0.333	N
Index of Attitudes toward Women	0.043	(0.026)	2,818	0.044	Y



**Table S20. Impacts by Exposure to Violence - Saw Violence, Abducted, Beaten, Raped, Maimed, Family killed.** See Table S6 for notes. Also, T x Violence Exposed is the interaction of Treatment with the violence exposure measure in panel headings.

VARIABLES	T		T x Violence-exposed		Obs.	R-sqr.	Baseline dependent variable controls?
	Coeff.	Std. Error	Coeff.	Std. Error			
<b>Panel A. Violence Exposure: Saw Violence, Raped, Maimed, Beaten, Abducted</b>							
Forgive Perpetrators	0.552	(0.424)	-0.105	(0.587)	1,945	0.136	Y
Trust Rebel Ex-combatants	0.229*	(0.116)	-0.045	(0.136)	873	0.226	Y
Trust Migrants	0.172***	(0.049)	-0.084	(0.066)	2,108	0.175	Y
Index of Generalized Trust in Community	-0.016	(0.036)	0.003	(0.046)	2,861	0.144	Y
Index of Attitudes toward Ex-Combatants	-0.009	(0.040)	-0.011	(0.053)	2,848	0.075	Y
Index of War Attitudes	-0.039	(0.039)	0.013	(0.049)	2,861	0.060	N
Index of Network Strength	0.084**	(0.034)	0.032	(0.057)	2,868	0.066	N
Index of Community Group Participation	0.072***	(0.026)	-0.021	(0.033)	2,865	0.163	Y
Index of Contributions to Public Goods	0.037	(0.028)	0.011	(0.032)	2,868	0.178	Y
Index of Psychological Wellbeing	-0.160***	(0.052)	0.011	(0.064)	2,852	0.121	Y
Less PTSD	-0.871***	(0.309)	0.298	(0.391)	2,662	0.123	Y
Less Anxiety	-0.476**	(0.213)	0.003	(0.268)	2,778	0.144	Y
Less Depression	-0.270**	(0.127)	-0.044	(0.162)	2,788	0.094	Y
Index of Economic Outcomes	-0.071*	(0.037)	0.050	(0.045)	2,868	0.169	Y
Index of Economic Activity (Individual level)	-0.006	(0.033)	0.016	(0.046)	2,868	0.164	Y
Index of Social Divisions	0.014	(0.032)	0.017	(0.042)	2,861	0.084	Y
Index of Attitudes toward Women	0.019	(0.039)	0.027	(0.053)	2,847	0.039	Y
<b>Panel B. Violence Exposure: Saw Violence, Raped, Maimed, Family Member Killed, Beaten</b>							
Forgive Perpetrators	0.350	(0.452)	0.257	(0.568)	1,975	0.133	Y
Trust Rebel Ex-combatants	0.144	(0.118)	0.073	(0.140)	875	0.227	Y
Trust Migrants	0.228***	(0.054)	-0.139**	(0.065)	2,155	0.175	Y
Index of Generalized Trust in Community	0.017	(0.049)	-0.020	(0.056)	2,937	0.139	Y
Index of Attitudes toward Ex-Combatants	0.045	(0.051)	-0.083	(0.057)	2,921	0.075	Y
Index of War Attitudes	-0.029	(0.051)	0.013	(0.056)	2,936	0.059	N
Index of Network Strength	0.071	(0.050)	0.034	(0.071)	2,944	0.062	N
Index of Community Group Participation	0.063**	(0.030)	-0.009	(0.037)	2,941	0.160	Y
Index of Contributions to Public Goods	0.037	(0.035)	0.004	(0.040)	2,944	0.173	Y
Index of Psychological Wellbeing	-0.148**	(0.061)	0.010	(0.072)	2,923	0.116	Y
Less PTSD	-0.935**	(0.360)	0.402	(0.418)	2,723	0.120	Y
Less Anxiety	-0.494**	(0.248)	0.089	(0.288)	2,839	0.145	Y
Less Depression	-0.222*	(0.134)	-0.077	(0.165)	2,856	0.093	Y
Index of Economic Outcomes	-0.074**	(0.037)	0.062	(0.043)	2,944	0.164	Y
Index of Economic Activity (Individual level)	-0.018	(0.035)	0.036	(0.042)	2,944	0.159	Y
Index of Social Divisions	0.023	(0.033)	0.012	(0.041)	2,937	0.086	Y
Index of Attitudes toward Women	0.036	(0.047)	0.006	(0.060)	2,923	0.039	Y

**Table S21. Impacts by Ex-Combatants.** See Table S6 for notes. Also, T x Ex-Combatant is the interaction of Treatment with a binary variable indicating if respondent is an ex-combatant

VARIABLES	T		T x Ex-Combatant		Obs.	R-sqr.	Baseline dependent variable controls?
	Coeff.	Std. Error	Coeff.	Std. Error			
Forgive Perpetrators	0.462*	(0.234)	0.828	(0.854)	1,930	0.138	Y
Trust Rebel Ex-combatants	0.209**	(0.084)	-0.057	(0.209)	868	0.227	Y
Trust Migrants	0.128***	(0.035)	-0.009	(0.123)	2,081	0.178	Y
Index of Generalized Trust in Community	-0.014	(0.025)	-0.006	(0.094)	2,819	0.141	Y
Index of Attitudes toward Ex-Combatants	-0.015	(0.030)	0.001	(0.123)	2,806	0.076	Y
Index of War Attitudes	-0.029	(0.028)	-0.092	(0.126)	2,819	0.062	N
Index of Network Strength	0.093***	(0.029)	0.381	(0.439)	2,826	0.073	N
Index of Community Group Participation	0.060***	(0.019)	0.007	(0.095)	2,823	0.164	Y
Index of Contributions to Public Goods	0.043*	(0.023)	0.064	(0.082)	2,826	0.176	Y
Index of Psychological Wellbeing	-0.154***	(0.037)	0.136	(0.157)	2,810	0.125	Y
Less PTSD	-0.703***	(0.223)	0.355	(0.914)	2,626	0.126	Y
Less Anxiety	-0.495***	(0.127)	0.997	(0.604)	2,736	0.145	Y
Less Depression	-0.301***	(0.076)	0.317	(0.348)	2,747	0.095	Y
Perception of household needs	-0.112	(0.075)	-0.150	(0.313)	2,700	0.085	Y
Perception of economic situation	-0.134***	(0.041)	0.026	(0.213)	2,704	0.086	Y
Index of Economic Activity (Individual level)	0.004	(0.023)	0.064	(0.099)	2,826	0.187	Y
Index of Social Divisions	0.024	(0.023)	-0.064	(0.097)	2,819	0.082	Y
Index of Attitudes toward Women	0.040	(0.028)	0.093	(0.105)	2,805	0.041	Y

**Table S22. Impacts by Gender.** See Table S13 for notes. See Table S6 for notes. Also, T x Female is the interaction of Treatment with a binary variable indicating if respondent is female

VARIABLES	T		T x Female		Obs.	R-sqr.	Baseline dependent variable controls?
	Coeff.	Std. Error	Coeff.	Std. Error			
Forgive Perpetrators	0.249	(0.341)	0.534	(0.568)	2,009	0.138	Y
Trust Rebel Ex-combatants	0.165	(0.109)	0.024	(0.144)	900	0.222	Y
Trust Migrants	0.132***	(0.045)	-0.021	(0.061)	2,203	0.174	Y
Index of Generalized Trust in Community	-0.013	(0.035)	0.032	(0.048)	2,995	0.136	Y
Index of Attitudes toward Ex-Combatants	-0.054	(0.043)	0.083	(0.055)	2,979	0.077	Y
Index of War Attitudes	-0.024	(0.044)	-0.001	(0.051)	2,999	0.058	N
Index of Network Strength	0.113**	(0.050)	-0.038	(0.064)	3,004	0.076	N
Index of Community Group Participation	0.069***	(0.025)	-0.024	(0.031)	3,003	0.164	Y
Index of Contributions to Public Goods	0.035	(0.029)	-0.001	(0.033)	3,004	0.196	Y
Index of Psychological Wellbeing	-0.133***	(0.042)	-0.034	(0.066)	2,981	0.123	Y
Less PTSD	-0.526**	(0.260)	-0.328	(0.394)	2,775	0.124	Y
Less Anxiety	-0.534***	(0.174)	0.127	(0.266)	2,894	0.149	Y
Less Depression	-0.270***	(0.097)	-0.066	(0.156)	2,912	0.103	Y
Perception of household needs	-0.222*	(0.112)	0.153	(0.152)	2,856	0.086	Y
Perception of economic situatoin	-0.083	(0.060)	-0.092	(0.093)	2,860	0.084	Y
Index of Economic Activity (Individual level)	0.029	(0.035)	-0.032	(0.046)	2,099	0.157	Y
Frequency of borrowing and lending	0.093	(0.064)	-0.102	(0.084)	3,004	0.465	Y
Monetary value of borrowing and lending	0.770**	(0.338)	-0.860**	(0.430)	2,911	0.106	Y
Respondent belongs to an <i>osusu</i> (savings group)	-0.018	(0.028)	0.010	(0.042)	2,949	0.151	Y
Respondent buys from trader	-0.003	(0.018)	-0.016	(0.023)	2,955	0.078	Y
Respondent belongs to a labor gang	-0.004	(0.026)	0.002	(0.035)	2,738	0.178	Y
Days spent working on other's farms	0.830	(1.106)	-1.124	(1.385)	2,414	0.145	Y
Index of Social Divisions	0.020	(0.029)	0.009	(0.046)	2,995	0.088	Y
Index of Attitudes toward Women	0.034	(0.036)	0.005	(0.058)	2,981	0.054	Y

**Table S23. Controlling for FT Communal Farm.** See Table S6 for notes.

VARIABLES	Coeff.	Std. Error	Obs.	R-sqr.	Baseline dependent variable controls?
Forgive Perpetrators	0.640**	(0.248)	2,010	0.132	Y
Trust Rebel Ex-combatants	0.122	(0.080)	900	0.227	Y
Trust Migrants	0.129***	(0.038)	2,203	0.172	Y
Index of Generalized Trust in Community	0.038	(0.030)	2,996	0.139	Y
Index of Attitudes toward Ex-Combatants	-0.001	(0.031)	2,980	0.075	Y
Index of War Attitudes	-0.015	(0.034)	3,000	0.058	N
Index of Network Strength	0.132***	(0.036)	3,008	0.064	N
Index of Community Group Participation	0.052***	(0.017)	3,004	0.161	Y
Index of Contributions to Public Goods	0.043*	(0.024)	3,008	0.172	Y
Index of Psychological Wellbeing	-0.161***	(0.036)	2,982	0.116	Y
Less PTSD	-0.732***	(0.224)	2,776	0.115	Y
Less Anxiety	-0.455***	(0.128)	2,895	0.139	Y
Less Depression	-0.300***	(0.074)	2,913	0.090	Y
Index of Economic Outcomes	-0.042**	(0.020)	3,008	0.161	Y
Assets	0.105*	(0.055)	2,991	0.403	Y
Perception of household needs	-0.167**	(0.080)	2,857	0.083	Y
Perception of economic situation	-0.125***	(0.042)	2,860	0.081	Y
Index of Economic Activity	0.029	(0.029)	3,008	0.182	Y
Index of Conflict and Crime	0.112	(0.072)	274	0.275	Y
Index of Social Divisions	0.056**	(0.023)	2,996	0.089	Y
Index of Attitudes toward Women	0.041	(0.027)	2,982	0.035	Y

**Table S24. Economic Activity and Outcomes.** Each row represents a separate regression of the outcome shown in the first column on treatment assignment. The control mean is the mean in the control group at endline. All specifications include section pair fixed effects and the second round indicator, as well as the baseline outcome variable, and its interaction 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.

VARIABLES	Control Mean	Coeff.	Std. Error	Obs.	R-sqr.
<i>Index of Economic Outcomes</i>	0	-0.036*	(0.019)	3,008	0.161
<i>Index of Economic Outcomes - control for baseline of individual economic indicators</i>	0	-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
<i>Baseline balance on individual economic outcome indicators</i>					
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	
<i>Index of Economic Activity</i>	0	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 <i>osusu</i> (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 labor gang	0.333	0.002	(0.016)	2,738	0.164
Days spent working on other's farms	7.96	0.473	(0.618)	2,418	0.130

**Table S25. Societal Conflicts.** Each row represents a separate regression of the outcome shown in the first column on treatment assignment. Variables not shown in all regressions include section pair fixed effects and the second round indicator. The final column indicates if 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.

VARIABLES	Coeff.	Std. Error	Obs.	R-sqr.	Baseline dependent variable controls?
<b>Social Divisions</b>					
<i>Index of Social Divisions</i>	0.012	(0.017)	3,008	0.081	Y
<i>Index of Social Divisions (indicators in both baselines)</i>	0.028	(0.021)	2,996	0.085	Y
Indicators:					
Community is not divided between social groups	-0.049	(0.041)	2,996	0.106	Y
Dominant groups do not benefit more from community resources	0.026	(0.040)	2,963	0.094	Y
Marginalized groups benefit from community resources	0.053	(0.056)	2,809	0.092	Y
Respondent feels included and respected in the community	0.000	(0.005)	2,960	0.050	Y
Perception that social divisions escalated into conflict	0.037	(0.037)	2,943	0.140	Y
<b>Conflict and Crime</b>					
<i>Index of Conflict and Crime</i>	0.112	(0.072)	274	0.275	N
Indicators:					
Number of conflicts	0.002	(0.019)	274	0.320	N
Number of crimes	-0.005	(0.007)	274	0.226	N
Number of violent crimes	0.003	(0.003)	273	0.172	N
Number of inter-village conflicts (village level indicator)	0.122***	(0.042)	274	0.295	N
<b>Conflict Resolution</b>					
Resolved	-0.057	(0.050)	172	0.330	Y
Satisfactory resolved	-0.107	(0.067)	172	0.456	Y
Resolved without third party	-0.036	(0.035)	172	0.603	Y
Resolved with mediation from family/friends	-0.141**	(0.055)	172	0.547	Y
Resolved by chief	0.103*	(0.058)	172	0.326	Y
Fined by chief	-0.006	(0.009)	280	0.257	Y
<b>Gender Attitudes</b>					
<i>Index of Attitude toward Women</i>	0.044*	(0.025)	2,982	0.036	Y
Indicators:					
Belief that a wife has a right to her own opinion	0.019**	(0.008)	2,957	0.055	Y
Attitude toward wife beating	0.081	(0.115)	2,957	0.036	Y
Attitude toward wife beating (questions in both baselines)	0.033	(0.069)	2,957	0.0454	Y



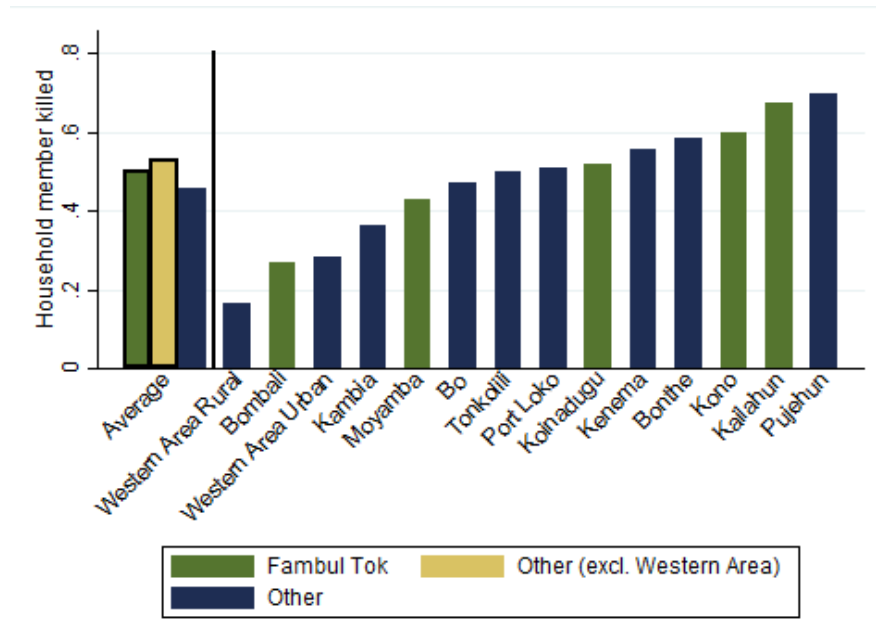
**Table S26. Persistence of Additional Outcomes.** These results present separate estimates for the two endline rounds in Wave One. Each row represents a separate regression of the outcome shown in the first column on treatment assignment. Variables not shown include section pair fixed effects and the second round indicator. The final column indicates if the specification also includes the baseline outcome variable, and its interaction 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.

VARIABLES	Sample: Round 1			Round 2			Baseline dependent variable controls?
	Coeff.	Std. Error	Obs.	Coeff.	Std. Error	Obs.	
Index of Attitudes toward Ex-Combatants	0.115**	(0.052)	875	-0.065	(0.055)	841	Y
Index of War Attitudes	0.015	(0.033)	828	-0.039	(0.063)	789	Y
Index of Economic Outcomes	-0.014	(0.029)	885	-0.014	(0.023)	850	Y
Assets	0.195**	(0.092)	879	0.182**	(0.083)	842	Y
Perception of household needs	-0.052	(0.142)	806	-0.146*	(0.080)	780	Y
Perception of economic situation	-0.113*	(0.057)	811	-0.067	(0.060)	784	Y
Index of Economic Activity	-0.025	(0.041)	885	0.021	(0.035)	850	Y
Index of Conflict and Crime	0.201	(0.139)	80	0.130	(0.122)	78	N
Number of conflicts	-0.018	(0.027)	80	0.030	(0.035)	78	N
Number of crimes	-0.006	(0.014)	80	-0.008	(0.011)	78	N
Number of violent crimes	0.009	(0.008)	80	0.002	(0.002)	78	N
Inter-village conflicts	0.231*	(0.134)	75	0.078	(0.063)	78	N
Index of Social Divisions	0.021	(0.026)	878	0.060	(0.039)	845	Y
Index of Attitudes toward Women	0.068*	(0.038)	877	0.007	(0.045)	844	Y

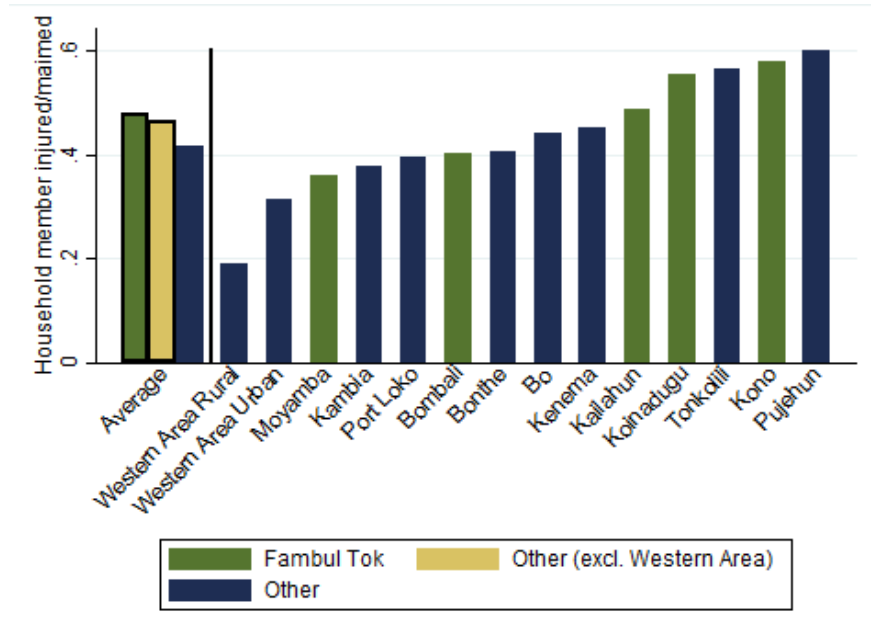


**Fig. S1: War Violence in Fambul Tok Districts and Non Fambul Tok Districts**

*Panel A: Household Member Killed*



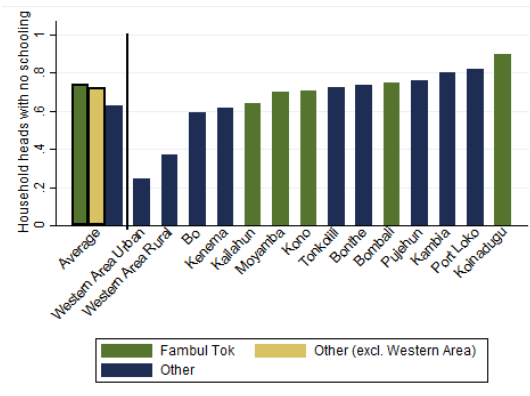
*Panel B: Household Member Injured*



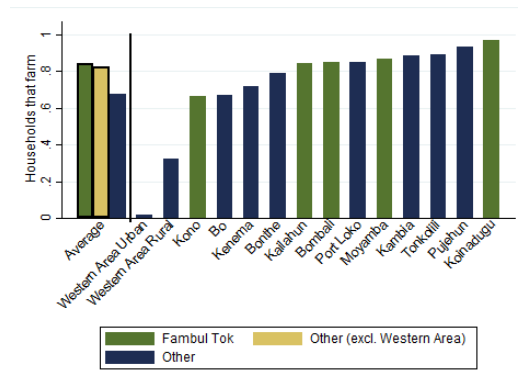
*Notes:* This figure shows average levels of war violence as reported in the 2007 NPS Survey in Fambul Tok districts (in green) and in other, non-Fambul Tok districts (in blue). The means for all Fambul Tok districts together are shown in green with a black border to the left of the graph. The means for all other districts together, including Western Area, which contains the capital of Freetown, is shown in yellow with a black border, while the mean for other districts excluding Western Area is shown in blue with a black border.

**Fig. S2: Socio-economic Indicators in Fambul Tok Districts and Non Fambul Tok Districts**

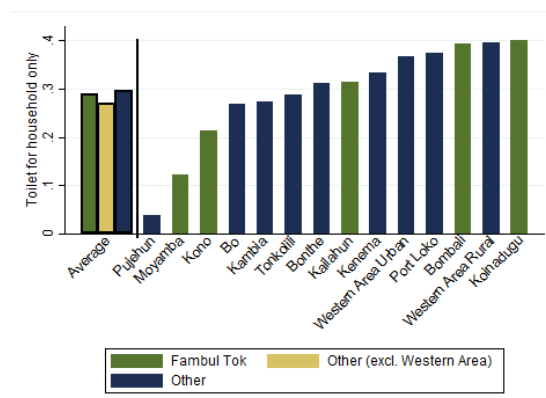
*Panel A: Formal Education*



*Panel B: Farming Households*

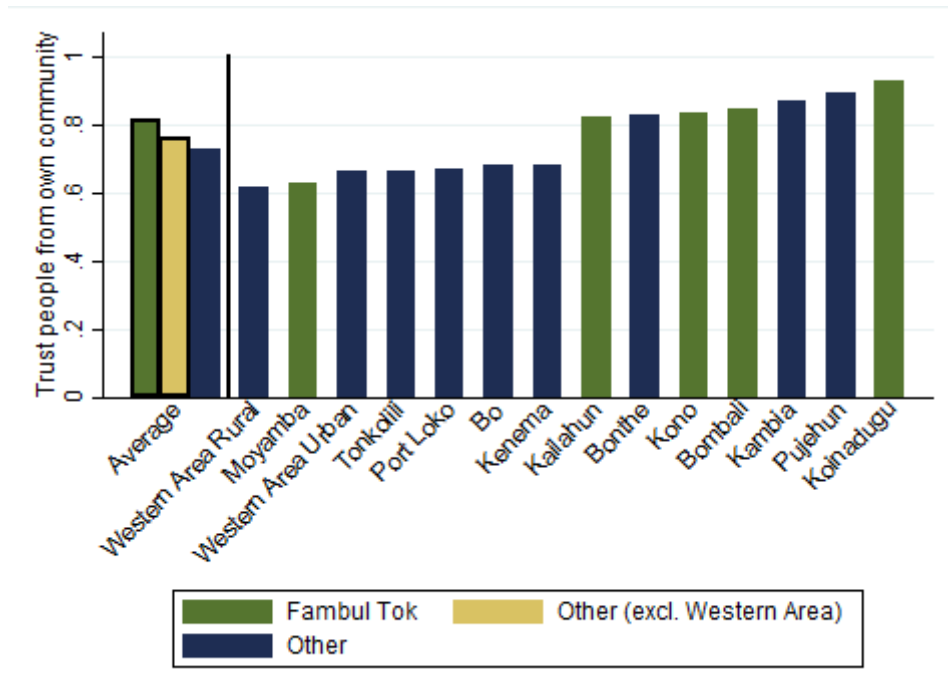


*Panel C: Toilet Facility in Household*



*Notes:* This figure shows average socio-economic indicators as reported in the 2007 NPS Survey in Fambul Tok districts (in green) and in non-Fambul Tok districts (in blue). The means for all Fambul Tok districts together are shown in green with a black border to the left of the graph. The means for all other districts together including Western Area, which contains the capital of Freetown, is shown in yellow with a black border, while the mean for other districts excluding Western Area is shown in blue with a black border. Panel A shows the fraction of households in which the head has no formal education. Panel B shows the fraction of farming households, and Panel C shows the fraction of households that have access to their own toilet facility.

**Fig. S3: Trust Levels in Fambul Tok Districts and Non Fambul Tok Districts**



*Notes:* This figure shows average trust levels reported in the 2007 NPS Survey in Fambul Tok districts (in green) and in the other, non-Fambul Tok districts (in blue). The means for all Fambul Tok districts together are shown in green with a black border to the left of the graph. The means for all other districts together, including Western Area, which contains the capital of Freetown, is shown in yellow with a black border, while the mean for other districts excluding Western Area is shown in yellow with a black border.

**Fig. S4: Timeline of the Study**

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