Moral Violence:
At the Heart of one of Congo's Oldest and Largest Armed Groups*

Lewis Dunia Butinda[†] **Raúl Sánchez de la Sierra**[‡] **Hilary Yu**[§]

September 2024

Abstract: We receive daily phone calls from the human resource commanders of one of Congo's largest armed groups, allowing us to observe who volunteered to join it and who did not, in real time. Borrowing notions and measurements from the field of moral psychology we unbundle altruism in a simple organizing framework and examine its role in violence participation. Those who exhibit more prosociality to the community have at the same time less *empathy* for the enemy—they are also significantly more likely to join the group, even controlling with an array of economic characteristics. This correlation and selection appears to be explained by a prevailing "morality of retaliatory/just violence:" civilians dominantly perceive joining the armed group to be morally right because of the consequence of joining and perceive killing members of the enemy, including civilians, as a moral duty, justified based on the prolonged violence by the enemy. These results underscore the value of un-bundling generic notions of altruism, suggest that a "morality of violence," which can emerge in contexts of existential threat, can govern violence participation especially by the most prosocial individuals, alongside economic motivations. **Keywords:** Morality, Violence, Genocide
JEL Codes: D23, O1, D73

*IRB approval was obtained from the University of Chicago. This project was supported by JPAL, IPA, the IGC, the Pearson Institute for the study of conflicts at the Harris School of Public Policy (University of Chicago), the Becker Friedman Institute, the Kiphart Center for Public Health, and the Harry Frank Guggenheim Foundation. We are thankful to the researchers and individuals who collected the data, to Marakuja Kivu Research, Kivu Agro Pastorale and to Akanksha Kala, Margot Barranco, Jan Fasnacht, Sebastian Hernandez Leiva, Zerai Wu, and Stan Xie, for excellent research assistance. We thank Nathan Nunn, James Robinson, and participants of BFI's Institutions Organizations and Growth, ERINN's conference, Brown development seminar, Boston University's development seminar, London School of Economics/Bocconi crime seminar, Stockholm IIES development seminar, and Harvard Political Economy seminar as well as Jean-Francois Bonnefon and Jean Decety for valuable comments.

[†]Kivu Agro Pastorale, Mpofi, Democratic Republic of the Congo. Email: lewis.dunia.butinda@gmail.com

[‡]University of Chicago and NBER. Email: raul@uchicago.edu

[§]University of California, Berkeley. Email: hilary.yu@berkeley.edu

1. Introduction

A growing literature in economics has made progress in explaining the causes of violent conflict. Much of that literature is based on so-called rationalist explanations: just as crime, violence would emerge when parties contest some economic prize (Fearon, 1995, Blattman and Miguel, 2010). This focus contrasts a large literature in anthropology, emphasizing the role of altruism broadly defined in violence and war (Kelly, 1985, Nisbett and Cohen, 1996, 1994, Nordstrom, 1997, Gluckman, 2009, Lagrange, 2016, Mcgovern, 2011). Protecting, raiding for, or avenging their community appear to be recurrent motivations. This omission is important. First, humans even stand out from other animals in their unusually high rate of intra-species homicides (Pinker, 2011, Gómez, Verdú, González Megías, and Méndez, 2016). Second, a predisposition for altruism or morality is an important motivator of other domains of behavior (Enke, 2019). Third, our altruism is, in fact, intimately related to our history of violence (Bowles and Gintis, 2011).

Are those who volunteer to take part in deadly conflict distinctively altruistic? Progress toward answering this question is complicated by two challenges. First, current notions of altruism in economics bundle pro/anti-sociality with *disposition to inter-personal harm*. For example, individuals who do not make a donation in a dictator game, or even choose to reduce the endowment of the opponent, will not necessarily be the ones most comfortable, or even enjoying, killing them. Similarly, violence often contributes to a community's public good, hence potentially provides pro-sociality rewards on some dimensions. Second, non-state organizations are secretive, thus existing research is limited to interviewing ex-combatants, with the survivor biases this implies; for state organizations, data of both joiners and non-joiners measuring altruism is hard to find.

In this paper, we analyze the volunteers' peculiar altruism in a war setting after addressing these challenges. First, borrowing from neuroscience, moral psychology and psychophysiology, we unbundle altruism into measurable notions relevant to violence participation. Second, drawing on eight years of preparation, we collect those measures directly on actual volunteers to one of the largest armed groups of the Democratic Republic of the Congo (DRC), the Nduma Defense of Congo-Rénové (henceforth, NDC-R), right *before* being hired (henceforth, *joiners*) and *simultaneously*, for each one, on randomly selected members of their households and villages.

The armed group is well-suited to study the role of altruism broadly defined in violent conflict. First, unlike other non-state armed groups in the region or mandatory conscription in armies at war, the group's recruitment is based on volunteers because there is a large flow

of volunteers and the group perceives involuntary recruitment as associated to low motivation. Furthermore, volunteers usually walk from where they live, which is typically a village in the state-controlled zone, far away from the group's area of influence, to the front where the armed group is. This allows us to abstract from selection into the group that may be induced through recruitment by force. Second, this recruitment model is a direct reflection of the fact that the armed group formed with the participation of civil society in North Kivu's district of Walikale in 2014 as a split from another armed group, with a mission to advance the interests of the Walikale community. Regardless of whether the mission is genuine, this group is suited to test whether altruistism plays a role in deciding to join it. Third, the armed group frames itself as *protecting* the community by *combating* an enemy—an armed group composed of Rwandan-Hutus. Thus, the setting is suited to a proof of concept for unbundling altruism. Fourth, the group is large (2,600 active members in 2024) and gets, on average, 35 volunteers monthly.

To motivate our analysis, we present an organizing framework that borrowing from the field of moral psychology (Decety and Jackson, 2004, Decety, 2011), unbundles altruism into: a. willingness to affect the material payoff of others, be it positively or negatively (henceforth, *pro/anti-sociality*) vs. disposition to physical inter-personal harm through violence (henceforth, *aversion(empathy)/inclination to inter-personal harm*); b. across targets (just as empathy is "selective"). Individuals' payoffs from joining depend on two idiosyncratic additive draws: (potentially positive) prosociality for contributing to one's community, akin to mission intrinsic motivation, and (positive or negative) aversion to harm the enemy, akin to a marginal cost of effort. The framework implies that, when the correlation between these two draws is negative or even weakly negative (that is, when unbundling is informative), conditional on joining, prosociality to the community and inclination to harm the enemy are larger on average. Moral values are important to interpret such correlation. First, whether individuals perceive joining to be prosocial to the community depends on the community members prevailing moral convictions that it is moral to join. Second, moral convictions about the appropriateness of harming others can over-ride humans' aversion to inter-personal harm (Workman, Yoder, and Decety, 2020); thus, if those who are more prosocial to the community uphold them more, or tend to believe it is more appropriate, prosociality to the community might be negatively correlated with aversion to harm the enemy.

The first key input into our analysis is a set of measurements informed by this framework. To measure prosociality, we use standard games and survey tools validated globally (Falk, Becker, Dohmen, Huffman, and Sunde, 2022). To measure disposition to inter-personal harm, we borrow

tools from moral psychology (Ferguson and Inzlicht, 2022, Chen, Chen, Decety, and Cheng, 2013) and measure emotional arousal in electro-dermal activity (henceforth, *EDA*) to validate them (Fowles, Christie, Edelberg, Grings, Lykken, and Venables, 1981). To measure the moral convictions about appropriateness of *joining* and about *killing*, we adapt existing moral dilemmas (Thomson, 1984, Greene, 2013). The subjects implemented these privately and confidentially, towards an unknown member of the district (henceforth, the community) and towards an unknown Rwandan Hutu (henceforth, *the enemy*) in randomized sequence.

The second key input into our analysis is a system that we developed between July 2019 and April 2020, interviewing and training the 55 human resource commanders of the group (henceforth, S1s). The system produced real time phone calls since April 2020 about when and where volunteers meet an S1 to declare their intention of joining, as well as their village of origin. This allowed us to survey the universe of joiners between January and September 2022 (n=213), prior to them being onboarded—as well as, simultaneously, a random sample of joiners' household members (n=213) and of adults in their villages of origin (n=213, from 133 villages). Using the group's monthly administrative records, which we gather since April 2015, we could corroborate that this system covers the universe of joiners. To corroborate our measures, we gathered, seven months after joining, contemporaneous measures of disposition to inter-personal harm and supervisor and peer assessments of the joiner's behavior in their first seven months.

Our first result is that prosociality to the community is *negatively* correlated with empathy to the enemy. We establish this result in two steps. First, we empirically validate the significance of unbundling altruism for the study of violence. For the joiners, we find that dispositions for inter-personal harm are hugely predictive of supervisor and peer evaluations of actual behavior in the group. The exercise reveals huge levels of parochial disposition to inter-personal violence and prosociality. For all respondents from whom we measured EDA at the time of joining, we find that self-reported measures of empathic response in response to witnessing the pain of others are significantly correlated with actual emotional arousal in that instant. Emotional arousal also reveals a bias in the empathic response to witnessing the pain of others. Second, we unbundle altruism and analyze the correlations between its relevant components. We find that high prosociality to the district is associated with low aversion to harm members of the enemy group—again, driven by cognitive empathy. This relationship is starker within households than within villages; thus, a testable implication of the framework is that joiners should have more prosociality to the community and less empathy to the enemy especially within households.

Our second result is that, consistent with the framework's prediction implied by this correlation, those who join the violence are, at the time of joining, both more prosocial to the community *and* less empathetic/more inclined to harm the enemy. Importantly, joiners have just as much empathy towards members of their community than the rest, and are just as likely to test positive for psychopathy in standard measurements than the community, i.e., they do not have lower capacity for empathy. As one should expect, this selection exists alongside selection on economic characteristics: joiners are more likely to be unemployed and have less income. However, selection on altruism is unaffected (and in fact strengthened), controlling for economic characteristics such as income, education, and unemployment, suggesting that selection on economic characteristics cannot explain selection on altruism. Corroborating the role of prosociality in the decision to join, the joiners' reports about why they joined, why others joined (first-order beliefs) and what others think about why they joined (second-order beliefs) are dominated by contribution to their community's territorial integrity.

Our third result is that a prevailing *morality of war* can explain this correlation and selection and is not simply motivated reasoning to justify an inclination to harm Rwandan Hutus. First, underscoring the central role of prosocial motivations in the decision to join, the civilian population in the district's villages regards joining as morally right because of the consequences of joining for society. At the same time, the civilians also hold a "morality of retaliatory violence:" most hold the moral conviction that it is right to kill a member of the enemy group when there are no other casualties, with most holding that conviction regardless of whether it is a civilian, a pregnant mother, or an infant, and most considering their murder morally right even if there are 100 additional casualties among the enemy group. This violent morality is upheld justified based on the violence perpetrated by the enemy in the past, and through substantial de-humanization and moral disengagement more generally. Those who hold it have the same understanding of the pain it causes than the rest, but care less about the consequence. Second, joiners' lower levels of empathy to the enemy cannot be explained by stronger moral values—prosociality to the community is negatively correlated to empathy to the enemy, regardless of moral values. Third, motivated reasoning and emotions related to the war do not help explain why prosocial individuals have less empathy to the enemy, nor why they are also more likely to join—the relationship is robust to controls for motivated reasoning and emotions. Taken together, these suggest a prevailing morality of retaliatory violence might govern the correlation between prosociality and disposition to inter-personal harm, and the selection into violence.

These findings reveal, to our knowledge for the first time, a "morality of violence" in which prosociality to the community is associated with an inclination to murder the enemy, supported by moral beliefs about the righteousness of murdering them. This morality appears to be associated to an environment of war and perceived existential threat. With such circumstances, basic presuppositions about morality from peaceful settings may not hold, the romanticized bundle of altruism breaks down, and violence may be perpetrated by ordinary men who are just more prosocial (to the community) who also reduce their empathy to the enemy.

Our findings contribute to various areas of knowledge. First, they suggest altruism is a bundle worth unbundling. While a growing literature in economics unbundled prosocial preferences across targets (eg Enke (2019), Fouka and Tabellini (2022), Habyarimana, Humphreys, Posner, and Weinstein (2007)), we borrow from moral psychology and neuroscience to also unbundle altruism into the conventional notion of prosociality and into *distinct* notions of empathy that have been shown to capture responses in entirely different different neural networks (Decety and Jackson, 2004, Zaki and Ochsner, 2012). In addition to demonstrating that this unbundling is necessary to explain, for example, violence participation and moral values, our examination complements the distinction between universalist and communitarian altruism. Specifically, Enke, Rodríguez-Padilla, and Zimmermann (2022) describe moral universalism as a "more uniform allocation of a given altruism or trust "budget." In our setting, we document a violent form of parochial altruism, consistent with violent communitarian morality, yet the "marginal propensity to spend the altruism budget" is positive in all identities for prosociality, empathic concern, and affective empathy, while it is negative for cognitive empathy. This suggests that altruism "spending" across targets could be complements or substitutes depending on the notion.

Second, our findings expand the literature on who participates in violence and who does not, by observing, for the first time, the morality of those who join and those who do not (Humphreys and Weinstein (2008); Mugaruka, Marchais, Sanchez de la Sierra, and Wu (2024)).

Third, our study extends the literature on the causes of genocide, which burgeoned following the Holocaust (Arendt and Elon (2006); Zimbardo (2008); Browning (2017)). Taking measurements of genocidal motivations among actual perpetrators of violence, we extend the literature beyond laboratory settings analyzing behaviors that might not be predictive of genocide. Applying measurements from moral psychology, we also extend the strand of the literature that has been limited to using qualitative methods among actual perpetrators (Browning, 2017). We further extend the literature by analyzing (to-be) perpetrators and non-perpetrators.

2. Background: The Conflict in Walikale and the Nduma Defense of Congo-Rénové

Guidon Shimiray's Nduma Defence of Congo-Renové is today one of the most powerful and oldest armed groups in all of the DRC. Predominantly controlling vast areas of its home district of Walikale (and part of other districts) in North Kivu province, its territorial control today covers more area than any other armed group in the DRC.[1] The significance of this comparison is hard to exaggerate, given that there are approximately 122 active armed groups today in eastern Congo alone (Vogel, Salvaggio, Boisselet, and Stearns, 2021), in one of the longest modern conflicts.

The predecessor of the NDC-R was the Nduma Defence of Congo (henceforth, NDC), created in 2008. The origins of the NDC combine opportunism by some of its leaders (Ntabo Ntaberi Sheka) related to the control of tin mining, alongside a mission to advance the interests of the Walikale community, in the face of unjust allocation of mineral rents in the global supply chain and of real and perceived threat to its territorial integrity from Rwandan-speaking populations.[2] Regardless of whether the protection motivation to form the group was genuine on the part of its creator, the framing of the movement as advancing the development of Walikale in the face of unjust allocation of mineral rents in the global supply chain, and as providing protection against Rwandan encroachment was locally very appealing, garnering increasing popular support among the population of Walikale. First, reflecting its purported motivation to advocate for the interests of Walikale, Sheka's own slogan at the time was: "*the minerals of Walikale shall first serve Walikale, before serving elsewhere*." In a context of extreme poverty, where children did not have school infrastructure or a functioning health system while valuable minerals were being exported to be sold at much higher prices abroad (Nest, 2011), the NDC presented itself as the only representative of the interests of the Walikale community in the face of a major injustice. Second, in a district with a long history of conflicts with Tutsi and Hutu populations from Rwanda dating back to a century before, and involved in systematic dispossession, settler colonialism, and gruesome violence against the Walikale population especially after 1994, the framing of the movement in terms of protecting Walikale against Rwandan aggression to its territorial integrity sparked pre-existing sentiments and fueled its popular support. The NDC had a spectacular success at freeing most of Walikale's historical homeland from the Rwandan Hutu militia Front de Liberation du Rwanda (*henceforth*, FDLR), who, after perpetrating the

[1] Figure A1 presents a map of Walikale's district against the DRC.

[2] The predecessor of the NDC, the MC45, was born in April 2007 with just a dozen collaborators, with the purported motivation to combat the resettlement of 45,000 Rwandan Tutsi families to Walikale (Congo Research Group, 2020).

Rwandan genocide, have committed countless abuses in Walikale alongside their predecessors since the 1990's. This success was supplemented by regular "sensitization campaigns" (in which the NDC explained the objectives of the group to the population); by its reliance on customary institutions of Walikale (regularly consulting customary chiefs) and by the regular provision of public goods (such as repairing roads and even in some cases maintaining health centers in Walikale (Congo Research Group, 2020)), all contributing to nurture high popular legitimacy.[3]

The NDC-R emerged as a split from inside the NDC in 2014, which took over the NDC in 2016. In September 2014, a coalition of NDC officers led by one of the former leaders, Guidon Shimiray, rallied against the NDC's leadership. Soon, the split attacked the positions of the NDC, a conflict which ended with the surrender of Sheka to the United Nations in July 2016.[4] The creation of the NDC-R was framed as a direct response to the deviation of Sheka from the objectives of the group to advance the interests of the community of Walikale. The NDC-R's mission statement indicates that its goal is to provide security in Walikale against Rwandan armed groups (FDLR), who have repeatedly brutalized the population.[5] Its operations are fueled by a sophisticated state apparatus. Their fiscal documents, to which we have access since 2015, indicate that the group generates of 1.5 million US Dollars per month—from poll/capitation taxes, mining fees, trade licenses, and from liquor and cigarette distribution monopolies. Most of these revenues appear to be spent in the wage bill of the combatants, logistics, and social and medical support for its staff. By 2024, the NDC-R is composed of 2,600 combatants, controls 110 villages in most of the Walikale district and other districts (source: authors' calculations based on NDC-R's administrative records), the largest area controlled by any armed group today (see Figure A3).

How do people join the NDC-R? First, the NDC-R hinges on people joining voluntarily (henceforth, *volunteers*). According to its leadership: "*we accept anyone but we don't like opportunists, they do not fight hard in the front*" (Source: interview with NDC-R's commanders, April 2015). To join, volunteers typically walk a few days from the state-controlled zone of Walikale, where most of the population lives, to the front controlled by the NDC-R. They are first referred to the closest

[3]In the first few years of its existence, the NDC conducted joint operations with the FDLR. While the motives for this alliance remain unclear, the justification provided by the leadership is that this allowed to spy on their enemy. In August 2010, the alliance attacked three villages of Walikale and hundreds of civilians were raped. The NDC distanced itself from the FDLR as this represented a loss in their legitimacy in their own population. Similarly, later, the leadership connected with General Bosco Ntaganda, a Rwandan-allied Tutsi deputy commander of the national Army in the Kivus. This resulted in a major operation against one of the leaders of the FDLR, Colonel Sadiki in 2011. It is at this point that the NDC began to frame their struggle as one against the FDLR predominantly.

[4]After two years of trial, 337 victims of Sheka were identified and a Congolese Military trial sentenced Sheka and his accomplice Lionceau of war crimes by murder, sexual slavery, enrolment of child soldiers and pillage.

[5]Those were presented in their foundational document (a copy of which is provided in Figure A2).

Human Resource commander of the NDC-R for onboarding (there are 55 of them, called S1s and headed by the Human Resource commander of the NDC-R, the T1).[6] In 2023, 5% and 2% of the adult male population and of the total population in the catchment area are NDC-R combatants, respectively (Figure A5, Panel A). Each month, 30 to 40 new volunteers join (Figure A5, Panel B). Second, while the probability to remain alive after ten years remains above 95% for civilians, it drops to the 55-70% range for joiners, i.e., joining drastically increases the probability of death (Figure A5, Panel C). Comparing administrative records of staff wages to the average income of civilians suggests that joining offers some economic benefits but economic incentives may not be enough to explain why people join (Figure A5, Panel D). Third, the prosocial mission of the NDC-R figures centrally in the personal accounts of hundreds for why they joined it (Dunia Butinda, 2021).[7] Many who joined emphasized their desire to contribute to: a. the security of the community of Walikale; b. revenge against the past perpetrators, often with genocidarian undertones. Regardless of whether the organization's purported goals are genuine and whether the reader deems them moral, these are motivations of an altruistic and moral nature.

3. Unbundling Altruism: Organizing Framework and Measurement

A. *A Simple Roy Model with Unbundled Altruism*

This section presents a simple organizing framework that illustrates how altruism might matter in the decision to join. The framework serves two purposes. First, it demonstrates that the selection of who joins the violence, even if it is entirely driven by altruism, is not trivial, and depends crucially on the context and specifically on the prevailing morality. Second, it produces testable implications for the relevance of altruism in violence participation, linking key empirical correlations in the population to the selection of "types" into violence participation.

Consider the decision of whether to participate in violence against an enemy group, that potentially benefits one's own community (henceforth, *joining*). Suppose that the payoff from not

[6]Figure A4 presents the organizational structure of the NDC-R.

[7]Dunia Butinda (2021)'s decision to join appeared to be an extension of his commitment to his family and community and in order to contribute to the survival of the community. "The killings started in all the villages of Walikale. The Hutus, on their way, killed the villagers. We heard the shots from our village ... My grandmother took us to the forest to find refuge ... On the way back, we found corpses starting to rot. It a horrible spectacle, a feast for the crocodiles ... The death of my grandmother left in me a deep emptiness. I still remember today the fables she use to tell us ... One morning [11 years old], I thought enough is enough, if we don't act, we are going to be exterminated ... So, I took a machete and covered myself a bit, and joined the youth tracking the armed group attacking us ... This was the first day I used the AK47."

joining is: $U_0 = w_0$ where $w_0 \in \mathbb{R}^+$ is the outside wage, and that the payoff of joining is:

$$U_1 = w_1 + \underbrace{\bar{\theta} + \tilde{s}_C - \tilde{\lambda}_E}_{\text{Altruism}}$$

Where $w_1 \in \mathbb{R}^+$ is the going wage upon joining and $\bar{\theta}$ is the average altruism payoff, constant across individuals for parsimony; $\tilde{s}_C$ and $\tilde{\lambda}_E$ are two idiosyncratic components: the intrinsic motivation payoff from contributing to one's own community, and the cost of effort arising from the effort to exert violence on members of the enemy group, respectively, both with mean zero.

The two idiosyncratic components represent two central notions from the field of moral psychology. The first component, $\tilde{s}_C$, captures the utility (intrinsic motivation) derived from knowingly contributing to one's community (henceforth, *prosociality* towards one's community). Key to this formulation is that individuals *believe* that joining contributes to the community. Thus, in the empirical analysis, we analyze the the prevailing beliefs in the population about the appropriateness to join. The second component, $\tilde{\lambda}_E$, captures the cost of effort (exerting violence against an enemy) arising from aversion to inter-personal harm against members of the enemy group, henceforth denoted *empathy* (Decety and Jackson, 2004). Key to this unbundling are not only these two distinct notions of empathy, but also the notion that both empathy and prosociality are "selective:" even for one individual, it varies across targets (Decety, 2011). Overall, this captures that while prosociality and empathy are related, they are distinct notions; furthermore, individuals who are more prosocial towards their community are not necessarily more empathetic towards the enemy. This framework abstracts from the established fact that empathy, a catch-all term, captures different notions, which we consider separately in the analysis.

Assumption A (Joint normal distribution)

$$\begin{pmatrix} \tilde{s}_C \\ \tilde{\lambda}_E \end{pmatrix} \sim N\left(\boldsymbol{0}, \begin{pmatrix} \sigma_s^2 & \sigma_{s\lambda} \\ \sigma_{s\lambda} & \sigma_\lambda^2 \end{pmatrix} \right)$$

Let us define the correlation coefficient between these two terms as $\rho_{a\lambda} = \frac{\sigma_{s\lambda}}{\sigma_\lambda \sigma_s}$. Thus, when $\rho_{s\lambda}$ is large, unbundling altruism will not be very informative, as both $\tilde{s}_C$ and $\tilde{\lambda}_E$ measure the same thing; in contrast, when $\rho_{s\lambda}$ is low or negative, this corresponds to cases where it is informative.

We assume that individuals maximize their payoff and only make one choice: whether to join the armed group. Therefore, an individual chooses to join if, and only if:

$$U_1 > U_0 \iff \tilde{s}_C > w_0 - w_1 - \bar{\theta} + \tilde{\lambda_E}$$

This framework induces a taxonomy of three types of selection into violence along the altruism dimensions. The three cases are presented in Figure A6. First, if prosociality towards one's own community and empathy towards members of the enemy are *strongly positively* correlated, there are two cases. The intuition for it is straightforward: some individuals tend to care about others more, regardless of their identity. These cases appear to be reasonable for situations of peace, in which in-group and out-group distinctions are mild. *Case I*. This case is induced is the variance of prosociality towards the community is large relative to empathy towards the enemy. The intuition is that, as prosociality to the community increases (across individuals), this increase dominates the increase in the cost of empathy to the enemy. In this case, joiners are more prosocial towards the community as well as more empathetic towards the enemy, corresponding to "soft" social movements of prosocial individuals universally empathetic. *Case II*. This case is induced if the variance of prosociality towards the community is *low* relative to empathy towards the enemy. The intuition here is that those who are prosocial to the community are disproportionately averse to harming the enemy—and those who are least prosocial to the community care disproportionately less about harming the enemy. In this case, joiners are *less* prosocial towards the community and *less* empathetic towards the enemy, corresponding to a "mercenary" army of members with low prosociality and empathy.

Second, if prosociality towards one's own community weakly or negatively correlated with empathy towards the enemy, there is only one case. *Case III*. Those who join have more prosociality towards the community and *less* empathy to the enemy. This corresponds to cases in which those who choose to participate in violence care unusually for their community or cause, and have unusually low empathy to the enemy. This correlation, and hence selection, could emerge in cases of retaliatory violence—in which those who care most about their community are more willing to harm those who wronged their community (rather than just inter-personal vengeance). More importantly, regardless of the proposed framing for justifying the violence, this characterizes ethnic supremacist and genocidal violence, in which the community is perceived to be in existential threat. In that case, violence with the specific intent to destroy, in whole or in part, a national, ethnic, racial, or religious group, be it by elimination of their members or the pillars of their life as a community, could be perceived as prosocial to the community—as documented in the narratives that enable voluntary participation in genocidal violence.[8]

[8]For example, in the Rwandan genocide, radio propaganda inculcating the collective fear that the Hutus were under threat of extermination by the Tutsis appears to have played an important role in mobilizing large segments of the population for genocide—to contribute to the survival of the community (Yanagizawa-Drott, 2014).

Yet, while this can explain why violence towards an enemy can be prosocial towards the community, it is insufficient to explain why prosocial individuals would have (weakly) less *fundamental aversion* to harm the enemy. Research in neuroscience has shown that moral values have the potential to "override natural aversion to interpersonal harm" (Workman, Yoder, and Decety, 2020). Thus, a morality of retaliatory violence, upholding that it is morally right to harm members of the enemy group, could induce this correlation through three plausible channels. First, if prosocial individuals apply the prevailing moral convictions about violence more strongly than the rest, this would suppress their empathy to the enemy more than the rest—Appendix B formalizes this mechanism. Second, if prosocial individuals hold more violent moral values towards the enemy than the rest, this too could lead them to have lower empathy to the enemy than the rest. Third, if prosocial individuals have an inclination to harm the enemy regardless of morality, for example because they care more about the suffering caused by the enemy to their community and feel anger and revenge, they may develop moral values justifying their inclination, as motivated beliefs, selected to appease the psychological costs (Bandura, 2011).

In sum, this framework produces one testable implication for the role of altruism in the decision to participate in violence, linking the correlation between prosociality towards the community and empathy towards the enemy to the characteristics of the individuals who join the violence. Governing the altruism parameters and their correlation, we introduced two key moral convictions. First, the moral conviction that it is right to join the violence, regardless of whether one chooses to act according to such conviction, underscores the significance of prosociality in the decision to join. Second, the moral conviction that it is right (or wrong) to hurt an enemy governs the correlation between prosociality and empathy towards the enemy (and can invert it).

B. *Measuring Unbundled Altruism and Moral Convictions*

The survey design, based on six years of piloting, finalized in Fall 2021 and rolled out in January-August 2022, induces survey respondents to implement the following measurements. Each of the following survey modules were implemented as follows: a. it was the respondents themselves who completed them in complete privacy, after the respondent was trained sequentially to implement each module; b. the respondents were informed that their answers were immediately encrypted upon entry (which was true) and hence no individual who could trace the answers back to them; c. they had to provide answers separately regarding their choices towards each of different target identities: a member of the district of Walikale (henceforth, the community),

a Rwandan Hutu (henceforth, *the enemy*), and where appropriate we distinguish Rwandan Hutu civilians from combatants; these target identities were randomized.

We measure prosociality using the following unincentivized survey games and modules, which we borrow off-the shelve from Falk, Becker, Dohmen, Enke, Huffman, and Sunde (2018), Falk, Becker, Dohmen, Huffman, and Sunde (2022), who validated their design globally.[9] Given that most of these measures are sensitive to income effects, they may not be suitable for inter-personal comparisons; hence, in the course of the study, we also added a "time dictator game," in which the respondent is asked to donate up to ten hours of their own time in a way that benefits a stranger. To isolate income effects that might arise through opportunity cost of time, players are asked to answer under three different assumptions for the counterfactual use of their time: a. however they imagine they would use their time; b. if they had planned to use to work to generate income; c. if they had planned to only socialize with their family.[10]

We measure disposition for inter-personal harm in two ways. First, we borrow from moral psychology for definitions and measurement tools of three distinct notions of aversion to interpersonal harm: cognitive empathy, affective empathy, and empathic concern. While measurements tended to rely on inventories, an "empathy task" recently became widespread (Eres, Decety, Louis, and Molenberghs, 2015).[11] The task presents respondents with an image of a person suffering, such as being hurt by a knife (while the image is staged, it is impossible to tell that it is) and asks them to report: how much they think the person is suffering (cognitive empathy); how much sorrow they feel for the person (affective empathy); how much they feel the same pain (emotional empathy). We adapted the task to the social context, piloted it with the community, staging scenes of pain by Congolese individuals, for two types of accidents: someone's ear being cut and someone falling from a tree. As before, a randomized sequence of different photos

[9]First, we implement an (anonymous) dictator game (including with the option to reduce the endowment of the opponent). The player has an endowment of 10 tokens (with each token being worth 10 cents of USD), and has to choose to (anonymously) allocate any amount of tokens from 0 to 10 to an "opponent" of identity [X], or instead to reduce the opponent's 5 token endowment, but each token reduced costs half a token of the player's own endowment. The player first chooses whether to offer tokens, do nothing, or reduce the opponent's endowment, and then the amount to give (or to reduce). Respondents are informed that some games will be selected for the realized payoff to be given (we always compensated the respondents for the outcome of a cheating game). Second, we implement off-the-shelve hypothetical donations: "*If we were to give you 1,000 USD, how of that would you give (anonymously): to benefit the development of the community of Walikale?/ to benefit the development of the community of the Hutus? to benefit the growth of a large popular militia in Walikale?*" Since there was talk of a development project, it was believed that the answers to this question could be binding. Third, we collected measures of off-the-shelve self-reported declared prosociality (for example: "*in a scale from 0 to 10, how likely are you to make actions to the benefit of the Walikale community that might be costly to you today, without expecting anything in return?*").

[10]Reflecting qualitative learning during implementation, we only started collecting it after half the study.

[11]We are grateful for Professor Decety for sharing with us the tools validated for use in the United States.

presents the individuals as belonging to each of same three identities. The photos were taken so the identity of the person could not be guessed from the photo (Figure A7, Panel A). To allow us to validate this tool, respondents wore empatica E4 bracelets designed to measure electrodermal activity (henceforth, EDA) while they were completing the task, allowing us to detect emotional arousal upon witnessing the pain of others, thus to correlate empathy task self-reported measures with emotional arousal—the latter cannot be manipulated.[12]

Second, we also develop our own measure of inclination to inter-personal physical harm. To measure the extent to which respondents have a *taste* for the suffering of other individuals, respondents were also asked to report in the empathy task how much they *enjoyed* the image. The question was always asked after the empathy questions and pertaining to the same image. To measure respondents' *preferences* towards harming other individuals, we developed a video game in which respondents encounter another person on a bridge, and can make a choice about what to do to such person. Some of these choices are positive or do not result in physical harm (such as giving them a candy, doing nothing, stealing from them), but others engender their death (such as shooting them, pushing them off the bridge, or stabbing them). Upon watching an animated series of examples as training, the respondents then choose in a sequence of scenarios again with varying identities of the person on the bridge.[13] The games were tailored to be appropriate to community norms, after we realized that there was a large demand to play the game especially when the option was there to kill members of the enemy group (Figure A7, Panel B).

The framework presented in this section suggested that two moral convictions about the righteousness of joining and of murdering members of the enemy group might be central in governing the role of altruism in violence participation. First, to elicit whether respondents believe it is morally right to join the violence, we designed a joining moral dilemma, in which a person encounters a recruiter and has the option of voluntarily join. Respondents are asked: "*Choose what, according to you, is the morally right action in each scenario: should the person join, or not join?*" (see Figure A12). To prevent influencing the respondent, the game is played in private after the respondent was trained. Joining can be morally right because of the intrinsic moral value of the action (*deontological* moral value of joining), or because of the beliefs about the consequences of the action for society (*consequentialist*, or teleological, moral value of joining). Let

[12]Due to the difficulties to implement the bracelet and the prohibitive nature of the equipment, the bracelets were carried by a non-random subset of the sample. We therefore use the bracelets as validation.

[13]To enhance the realism of the game, sounds for each of the actions play on the tablet, and the consequences of the actions can be seen in an animated cartoon.

$s_C = s_C^d + s_C^c$, where s_C^d : is a deontological motivation and s_C^c : is a consequentialist motivation:

$$U_0(n) = w_0 + s_C^c * n$$
$$U_1(n) = w_1 + s_C^c * n + \left[s_C^c + s_C^d - \lambda_E\right]$$

The differential payoff from joining can be written as the following expression, where J is an indicator taking value 1 if the individual joins and zero otherwise:

$$U_J(n) = \underbrace{w_0 + s_C^c * n}_{\text{baseline}} + J * \left(\underbrace{(w_1 - w_0 - \lambda + s_C^d)}_{\text{d: benefit to self}} + \underbrace{s_C^c}_{\text{benefit to society}} \right)$$

If the person joins, this switches on two terms: the benefit to *self* and to the community. The former bundles the economic benefit (the wage differential $w_1 - w_0$), the (dis)utility from hurting the enemy, and the deontological motivation for joining. The latter isolates the consequentialist moral value of joining. To empirically pin down this term, we induce a trade-off between joining and the number of others who would join. At increments of x, respondents were then asked: "*if the person joins, another x young men cannot join. Should the person join?*" If, as assumed, the utility is additive in n, then respondents' choices will be monotonic: they *may* initially choose joining, but when x is above a threshold (which could be zero), they switch to not joining.

Second, to measure beliefs about the moral righteousness to kill members of the enemy group, we adapt trolley dilemmas (Thomson, 1984, Greene, 2013, Bénabou, Falk, and Henkel, 2024). The respondent faces the following dilemma (Figure A7, Panel C):[14]

> You are on this bridge [see the screen]. Under you, there is a tram, going towards a group of people from the Walikale community tied to the rails. Next to you, there is Rwandan Hutu. Choice A: if you do nothing, the tram stops, because it is not going fast, and none will die; Choice B: if you push the Rwandan Hutu off the bridge, the person dies, but this also creates a short-circuit that results in the death of the members tied to the rails. Regardless of what you would do, what should you do? Regardless of your choice, it is important to remind that your choice would carry no consequence (e.g., legal) and that none, in the scenario or in the real world, will be able to identify that the choice was made by you.

[14]The choice is elicited through the strategy method, in the following increments for the number of Walikale civilians sacrificed: [0; 5; 10; 25; 50; 100]; after the last choice, the respondent is asked to make their strategy explicit.

To separately identify a *consequentialist* moral motivation to kill them arising from the fact that they are perceived to represent a threat to one's own community (in terms of number of fellows that they could kill) from a *deontological duty* to kill them, in one version, the person on the bridge is a Rwandan Hutu that has been integrated in the community for decades, is a civilian, and will never harm the community for sure: the duty to kill them can only be explained by deontological morality and not through the consequences for the survival of one's own community members. To identify whether a morality of killing is genocidarian, in randomly sequenced versions, the identity of the Rwandan Hutu on the bridge is a pregnant mother, or an infant. These population sub-groups are often targets of violence in genocidal contexts (Graça Machel, 1996, Ndahayo and Dufatanye, 2015).[15] Finally, we also implement a "placebo" in which the individuals on the rails are also Rwandan Hutus—this allows to estimate the "moral willingness to kill" Rwandan Hutus (up to a number), unaffected by the appropriateness of killing members of the community.

4. Data Collection

To know who joins in real time, we created a communication system comprising all of the 55 human resource commanders of the group (the S1s), that is, the universe of the 55 "ports of entry" of volunteers into the group. We developed this system between January and July 2019, and we installed it between July 2019 and April 2020. The system consists in daily phone calls between the study coordinator and all of the S1s, in which the study coordinator learns about all new volunteers and their village of origin. To install the system, we conducted a survey to all S1 commanders between July and August 2019, collecting their basic demographics and trajectory, their contact information, as well as a phone number if they had one. This system allowed us to create two samples: the universe of new joiners before they were registered and onboarded; a random sample of civilians in the villages of origin of these new joiners. For each new joiner, we could then gather data from the new joiner and from their village of origin.

[15] For example, Graça Machel, the Secretary-General's Special Representative, declared in the first report on children at war in 1996 (Graça Machel, 1996) that "*Children are not just getting caught in the crossfire, they are also likely to be specific targets*," and provided an explanation based on her analysis that is germane to our design: "*When ethnic loyalties prevail, a perilous logic clicks in. The escalation from ethnic superiority to ethnic cleansing to genocide, as we have seen, can become an irresistible process. Killing adults is then not enough; future generations of the enemy—their children—must also be eliminated.*" The reasons for why these groups have tended to be targeted in genocide are complex but are intuitively not difficult to grasp. First, a pregnant mother represents the reproductive capacity of the community, carry a second human, and have the potential to procreate more. Therefore, when violence is deployed with the intent of destroying a group, pregnant mothers constitute an effective target. Second, targeting the youth might reduces more the years of life exterminated, and future generation. Third, pregnant mothers and infants represent an intimate core of what makes us human, thus are symbolic targets of attacks aiming to degrade the humanity of another group.

To assess whether the system captures all of the joining events, we take advantage of the fact that, since April 2015, we gather mondly the NDC-R's internal administrative records, which contain information about the new recruits. Since the information of joining is centralized with one month delay and arrives in our possession with at least two months delay, the administrative records are not a suitable source for conducting this study. However, it is an ideal independent source to verify the reliability of the S1 communication system. Figure 1, Panel A, shows the monthly total number of phone calls received reporting a new volunteer.

We deployed ten teams of two surveyors through Walikale immediately following all 213 calls about new volunteers that we received between January 22nd and August 25th 2022. The surveyors had been trained by us to the design and were known to the Walikale community. For each call, we deployed two teams simultaneously. One team was dispatched to the volunteer to administer an individual survey comprising the above measurements as well as basic socio-demographics (henceforth, *individual survey*). Another team was dispatched to the volunteer's village of origin, where they visited: a. a randomly selected household from the village census created by the surveyors with the village administrators; b. the volunteer's own household. In each of these two households, the surveyors implemented individual survey to a randomly selected youth from the household with the same gender as the volunteer, as well as a household survey with the head of household measuring wealth, health, income, and history of violence.

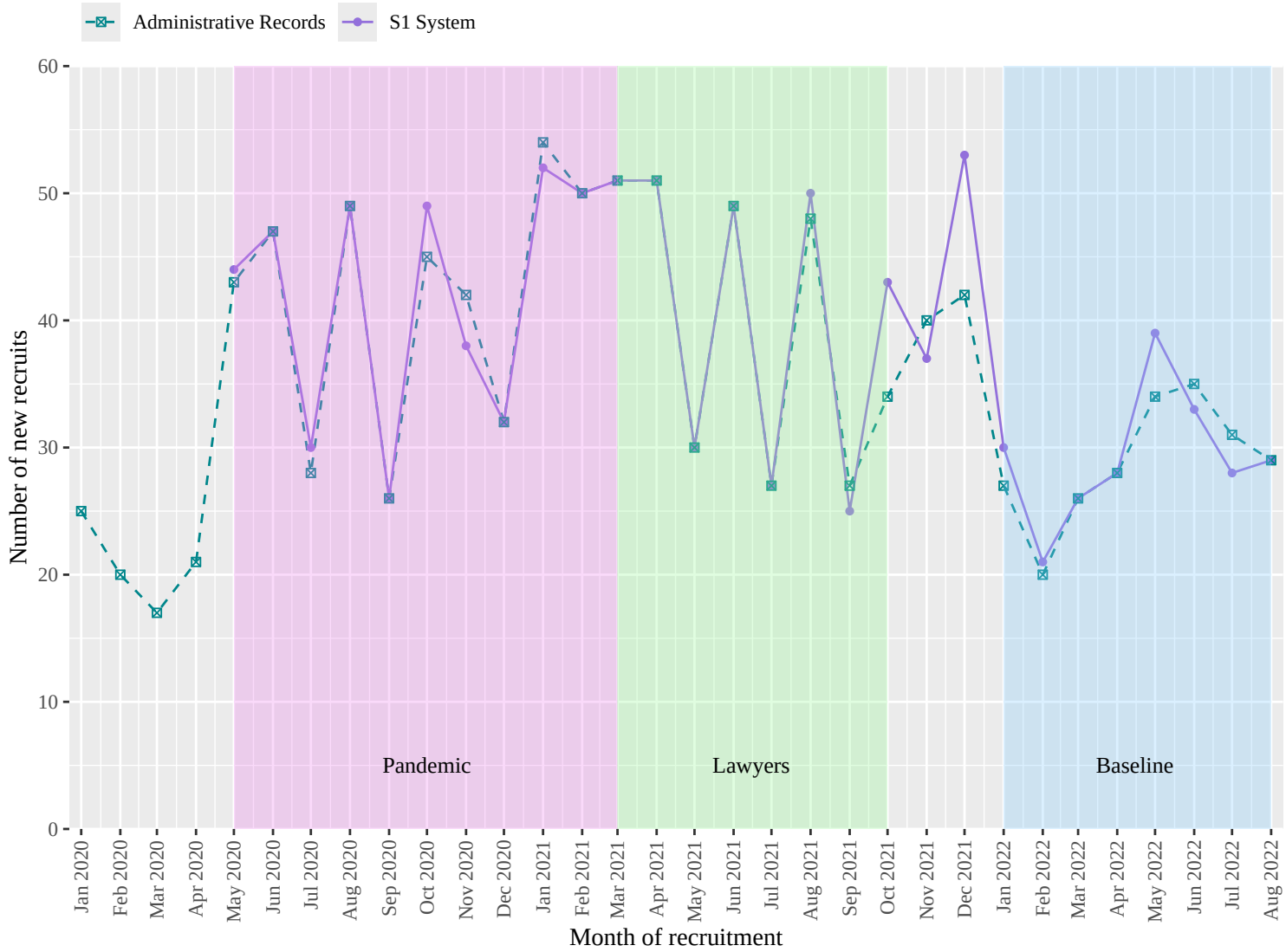
These interviews were conducted in private. For the measurements of disposition for interpersonal harm, prosociality, and moral dilemmas, they were implemented by the respondent alone in the absence of the surveyor (after the surveyor trained the respondent and verified their capacity to answer). Respondents were informed that the data would be encrypted (preventing access by any third-party) and would note be shared with the NDC-R nor with any person in the country. The data were collected by a reputed NGO trusted by the population, reassuring about the neutrality of the interaction and securing the respondents' trust.[16]

The sample comprises data from 213 households of new volunteers and 213 households randomly selected in the villages of origin of these volunteers, constituting individual data on 639 individuals and on 426 households. Figure 1, Panel B, presents a map of the location of the new volunteers first contact with an S1, their villages of origin and the NDC-R's area. A majority

[16]The data collection was overeseen by the three authors, administered by Marakuja Kivu Research (henceforth, Marakuja), and implemented by Kivu Agro Pastorale (KAP). Marakuja is a DRC-based NGO founded in February 2015 by 7 Congolese and one of the authors. KAP is the only non-governmental organization enjoying civilian and military trust to operate in NDC-R area. Importantly, KAP enjoys respect among civilians, as it is perceived as neutral and committed to the development of the Walikale community, mostly through activities that support the youth.

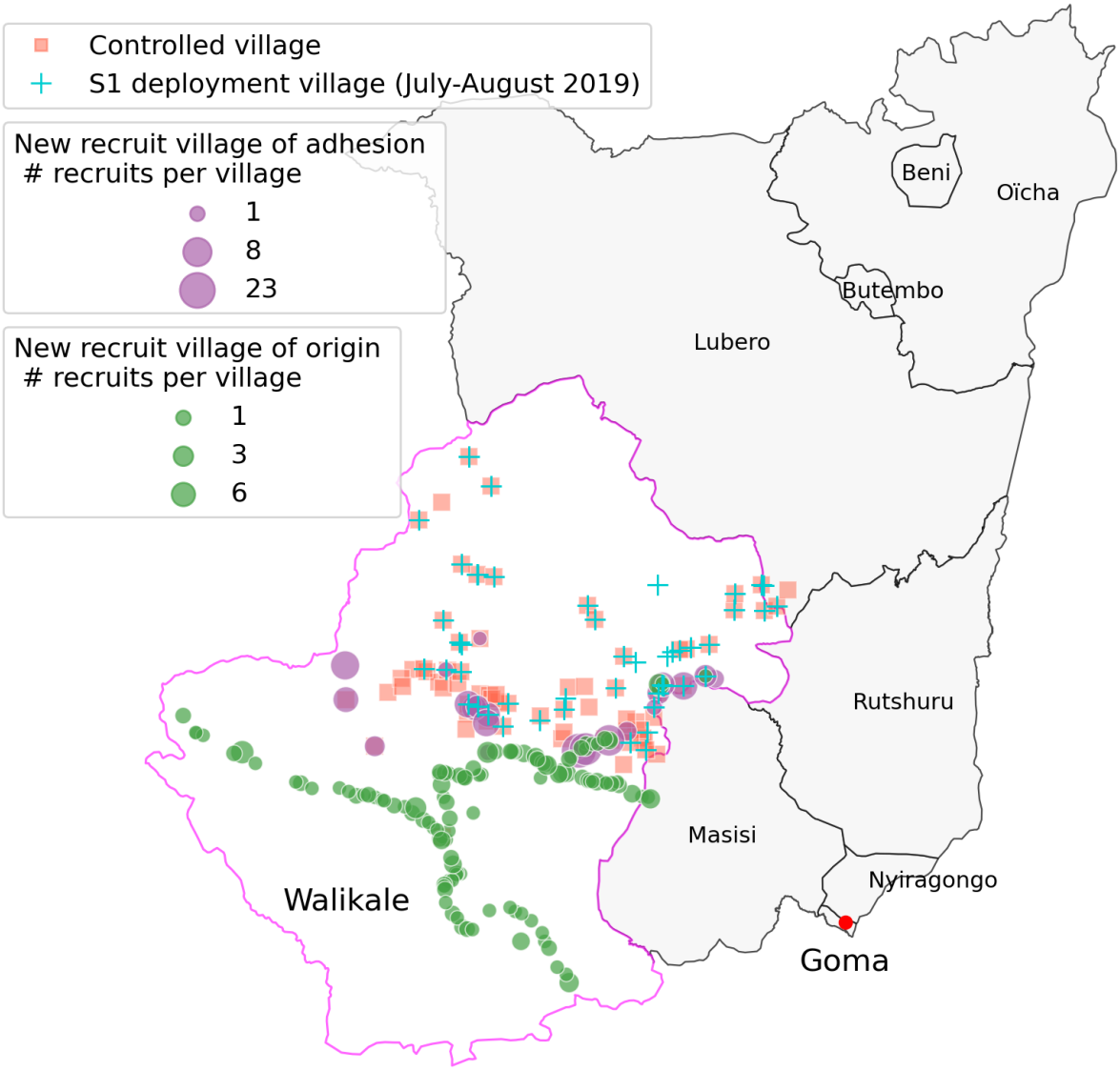
Figure 1: The S1 Communication System and Locations of New Volunteers and their Villages

Panel A. New Volunteers: S1 Communication System vs. NDC-R's own Adminstrative Records



Notes: This figure presents the times series for the number of calls received from the S1 commanders alerting of a new joiner, against the number of new joiners found in the NDC-R's administrative records (which are produced with one month lag and collected with three months lag and are thus not useful for the purpose of this study but can be used as valiadation). The S1 communication system provided us with real-time information about the events of joining, and allowed us therefore to interview the joiners before they were influenced by participation in the group. Since the joiners always had to wait a few days before they could be onboarded, this ensured that we could interview them in that window. The communication system was activated in April 2020, and was active through the pandemic (as it involved no human interaction), until we begun following-up on the calls on January 18th 2022, upon months of preparations with legal counsel at the University of Chicago through 2021, as well as ethics committee consultations, design of surveys and training, all in the fall of 2021. For S1s who had a cell phone, we relied on their phone for the daily communications. For S1s who did not have a cell phone, we examined whether the assistant to the S1 had a cell phone (all S1s had an assistant). In that case: if the assistant had a cell phone, the assistant was the point person for the corresponding port of entry; if they did not, we identified a civilian in the location of the S1 and bought a cell phone for the civilian, who then became our point of contact upon training the civilian and the S1 to communicate daily the information about new volunteers. This ensured that we covered all ports of entry, and prevented us from acquiring cell phones for active militia commanders.

Panel B. Locations of New Volunteers and their Villages



Notes. This figure presents the map of the location of new joiners adhesion into the group (purple dots) against the villages of origin of the new joiners (green dots). The GPS information to create this figure was obtained at the time of the interview conducted in response to the S1 call. These locations are mapped against the backdrop of the locations controlled by the NDC-R in 2022 (orange squares) as well as the location of the S1 commanders (blue crosses), which was obtained during our S1 onboarding survey in July 2019.

of volunteers came from villages far away from the NDC-R's area of control, reassuring that their decision to join is unlikely to be coerced, as suggested from our qualitative evidence.

Then, upon collecting these data, we immediately begun a follow-up of the joiners, for each joiner, eight months after they joined, in order to collect (again) the measures of unbundled altruism and to corroborate those with actual behavior. Measuring their behavior in the group

is an ideal setting to validate the measures of altruism and disposition for inter-personal harm because, unlike civilian life in the state-controlled zone (which is where most joiners come from), the environment in the group creates the opportunity to reveal the actual disposition in behavior. We gathered supervisor (commander) and peer (combatant) assessments of the joiner's behavior, for each joiner (as well as a self-report by the joiner as benchmark), alongside alongside measures of joiners' altruism collected in a survey to the joiners directly.[17] To measure tendency to be violent across targets in a context-relevant fashion, we specifically asked the evaluator to report ow high is the tendency of the joiner being evaluated to, on a scale from 0 to 10: a. abuse Walikale civilians; b. be violent against enemy Hutus. Upon interviewing all new volunteers in January-August 2022, we implemented this follow-up between September 2022 and May 2023.[18]

This data collection design can raise the concern that this agreement or interest may feed into an illegal armed group. Key features of the context shed light on this concern. First, the group is neither illegal nor is listed as a terrorist organization anywhere. Instead, the NDC-R is recognized as a reserve army in Congolese law: between April and September 2023, a new Congolese law that recognized various militias, including the NDC-R as "Wazalendo" (meaning "patriots" in Lingala) was adopted by the Parliament and the government, recognizing them as a national service of support to the government army and materializing a long history of collaboration (Radio France International, 2023, Actualite.cd, 2023). Second, since colonial times, a huge number of actors have formed a part of, benefited from, and been aware of organizations like this one. Their participation and the popular support of society at large, rather than the view of foreign researchers, determines whether they perceive it to be legitimate.

[17]With the only difference that, due to time, for prosociality, we only gathered the (two-sided) dictator game. For each joiner, the supervisor was the commander directly in charge of the joiner, and the peer was randomly selected from a roster that we created for each joiner listing all the peers of the same level that know the joiner well-enough to be capable of answering questions about them.

[18]The study was approved by the NDC-R's human resource agency in 2015, as well as by Walikale's customary authorities and district administrator, the National army unit controlling North Kivu,the intelligence agency in North Kivu (Agence Nationale des Renseignements, in French). Upon seeking legal advice from experts, it was determined that the design, including compensation of respondents, is in compliance with US and DRC law.[19] The design was also approved by the University of Chicago's IRB, IRB19-1102-AM002. Additional steps were taken to ensure the safety of the surveyors. The surveyors were carefully recruited civilians who were trusted by the leadership of the group and the community. They carried authorization letters and symbols making their activity public: KAP T-shirts, service cards. They traveled by day and we provided them with cutting edge security information. Furthermore, six months prior to the launch of the data collection, we created a local ethics committee, comprised of members of KAP, Marakuja, foreign researchers, as well as other Congolese researchers with the mandate to conduct a risk assessment of the design and to propose amendments: a. to reduce the risks to reasonable or otherwise as low as possible; b. mitigate the consequences of the risks; c. assess whether the amended design was below reasonable risk and ethical. The committee modified key components of the data collection (such as whether the surveyors should be publicly visible and how they ought to be introduced in the communities) and deemed the modified study below reasonable risk. We did not record any security incident to individuals or organizations directly or indirectly associated.

5. Unbundled Altruism and the Key Correlation

In this section, we present the measures of unbundled altruism and analyze the relationship between prosociality to the community and disposition for physically harming Rwandan Hutus (henceforth, *the key correlation*). We do so in two steps. First, we empirically validate the unbundling and the measures, by measuring the correlation between the measures of unbundled altruism and: a. the third-party evaluations of joiners' violent behavior inside the group; b. emotional arousal, measured by EDA. Second, we analyze the relationships between prosociality and the notions of disposition for inter-personal harm and, in particular, the key correlation.

A. *Unbundled Altruism: Validation with Observed Behavior and Emotional Arousal*

We begin our validation exercise by examining the correlation between our measures and *actual* behavior observed by third-parties: the assessments of joiners' behavior seven months after joining made by their supervisors and their peers (as well as by themselves). In Panel A, we present the estimates from regressing the (supervisor, peer, and self) reports about the joiner's tendency to be violent collected in the follow-up (henceforth, *endline*) on the endline measures of disposition for inter-personal harm, and of prosociality. To isolate the individual contribution of disposition for inter-personal harm vs. prosociality to tendency to behave violently, when analyzing disposition for inter-personal harm, we control for prosociality, and vice versa. We separate the analysis by specific measure and by target (community, enemy).

First, the measures of disposition to inter-personal harm capture relevant fundamentals: they are hugely predictive of actual observed violence, but the measures of prosociality are not. Empathy towards members of Walikale, which is higher than 8/10 on average both at baseline and at endline, are negatively correlated with the reports of violence against civilians.[20] In contrast, prosociality towards own Walikale members has weak to no predictive power of tendency to abuse of Walikale civilians in the group. Similarly, inclination to harm Hutus is positively associated with violence against Hutus and empathic concern and affective empathy of Hutus are negatively correlated with violence towards Hutus—whereas more cognitive empathy of the suffering of Hutus is associated with higher tendency to be violent towards them. Again, in contrast, prosociality towards the Hutus (the dictator game's contribution to Hutus is negative on average) is only weakly associated with tendency to harm the Hutus. Interestingly, it is precisely

[20] the measures of inclination to harm Walikale civilians are all essentially zero, hence unsurprisingly not predictive of abuse of Walikale civilians

Table 1: Validating the Unbundling of Altruism

Panel A. Correlation Between Altruism and True Behavior

	Mean Disposition		Regression coeff. Dep vars: Assessments of Joiner's Behavior						
	Baseline	Endline	Baseline	Harm of Own Civilians [0-10]. Report by			Enemy Harm [0-10]. Report by		
				Supervisor	Peer	Self	Supervisor	Peer	Self
Endline Altruism to Community									
Disposition for Inter-Personal Harm									
Empathetic Concern [0-10]	9.00	9.31***	0.010	-0.562***	-0.600***	-0.311**	0.822***	1.004***	0.888***
			(0.072)	(0.121)	(0.144)	(0.123)	(0.110)	(0.106)	(0.100)
Affective Empathy [0-10]	8.44	8.66*	-0.011	-0.465***	-0.312**	-0.427***	0.444***	0.720***	0.539***
			(0.076)	(0.101)	(0.129)	(0.098)	(0.099)	(0.100)	(0.090)
Cognitive Empathy [0-10]	8.36	8.70***	-0.111	-0.393***	-0.661***	-0.334***	0.445***	0.762***	0.518***
			(0.078)	(0.101)	(0.124)	(0.106)	(0.096)	(0.100)	(0.096)
Inclination (Taste) [0-10]	0.02	0.05	0.000	0.142	0.423	-0.322	-0.011	-0.093	-0.048
			(0.000)	(0.354)	(0.419)	(0.235)	(0.317)	(0.350)	(0.218)
Inclination (Preference) [0-10]	0.14	0.05	-0.007	0.000	0.000	0.067	0.000	0.000	0.009
			(0.122)	(0.000)	(0.000)	(0.166)	(0.000)	(0.000)	(0.155)
Prosociality (Dictator Game)	2.97	1.01***	0.096	-0.064	-0.129*	-0.049	0.012	0.070	0.098**
			(0.117)	(0.058)	(0.076)	(0.057)	(0.051)	(0.053)	(0.045)
Endline Altruism to Enemy									
Disposition for Inter-Personal Harm									
Empathetic Concern [0-10]	2.85	1.86***	0.041	0.342***	0.418***	0.320***	-0.156***	-0.174***	-0.151***
			(0.104)	(0.042)	(0.049)	(0.043)	(0.054)	(0.056)	(0.050)
Affective Empathy [0-10]	3.42	1.93***	0.014	0.330***	0.437***	0.340***	-0.114**	-0.155***	-0.072
			(0.108)	(0.039)	(0.049)	(0.038)	(0.046)	(0.055)	(0.047)
Cognitive Empathy [0-10]	4.96	6.99***	0.015	-0.064**	-0.063	-0.032	0.093***	0.119***	0.081***
			(0.074)	(0.032)	(0.038)	(0.031)	(0.028)	(0.031)	(0.028)
Inclination (Taste) [0-10]	6.30	7.11*	-0.081	-0.159***	-0.221***	-0.138***	0.136***	0.196***	0.158***
			(0.087)	(0.025)	(0.027)	(0.024)	(0.027)	(0.028)	(0.027)
Inclination (Preference) [0-10]	8.29	8.61	0.164*	-0.107***	-0.162***	-0.033	0.160***	0.317***	0.272***
			(0.091)	(0.039)	(0.042)	(0.036)	(0.040)	(0.038)	(0.032)
Prosociality (Dictator Game)	-2.31	-3.76***	-0.156	0.069	0.054	-0.028	-0.232***	-0.002	-0.100*
			(0.122)	(0.060)	(0.069)	(0.061)	(0.063)	(0.062)	(0.056)
Mean Dep. Var				1.57	1.78	1.48	8.90	9.01	9.28
Observations	211	209	209	209	209	204	209	209	204

Panel A presents the correlation between altruism and true behavior.

Panel B. Correlation Between Disposition to Inter-Personal Harm and Emotional Arousal

	Mean	Household Member	Villager	Joiner	Pooled
Disposition for Inter-Personal Harm					
Empathetic Concern	6.76	0.037	-0.042	0.040	0.023
		(0.063)	(0.058)	(0.031)	(0.025)
Affective Empathy	7.09	0.026	-0.032	0.054*	0.031
		(0.060)	(0.061)	(0.031)	(0.025)
Cognitive Empathy	8.07	0.088	-0.039	0.079**	0.059**
		(0.075)	(0.064)	(0.032)	(0.028)
Observations		303	255	608	1166

Notes. Panel B presents the coefficients of a regression of the measure of emotional arousal immediately after witnessing the prime of the person suffering in the empathy task, obtained through analysis of electrodermal activity, on the reported empathic response for affective empathy, empathetic concern, and cognitive empathy in the empathy task upon observing the same prime. Due to logistical and budgetary constraints, we obtained usable data on electrodermal activity for a non-random sample of 44 randomly selected youth (out of the 213), 59 randomly selected youth in the volunteers' households (out of the 213), and 125 new volunteers (out of the 213). The sample in this panel is the number of times an empathy task prime is played. Since emotional arousal is only meaningful for the empathy task, this panel only shows measures of the empathy task.

That it is driven by the joiners is not surprising, given that, as we noted in the description of data collection, we were only able to collect these EDA measurements in a small non-random sample of respondents, mostly joiners. The sample sizes in the table indicate the number of respondents, times the number of primes observed (6, composed of three idenities and two types of harm—fallen from a tree, and ear cut). Our validation at this stage excludes prosociality and inclination for inter-personal harm. First, we exclude prosociality and preferences for inter-personal harm because the exercise compares the report of empathic response experienced in response to an image to the (clinically measured) emotional arousal in response to the exact same image—whereas prosociality is a choice made at a later stage of the survey and is unrelated to the image. Second, while past literature has established a link between aversion to inter-personal harm and emotional arousal in response to witnessing someone else's pain, our measure of *taste* for inter-personal harm, related to the same image, was created by us, and there is no prior research on which to expect a positive, or instead negative association between emotional arousal upon observing the pain of another person and reported taste for inter-personal pain.

those who are most empathetic towards their own community who are also most violent towards the enemy and it is those who have most empathic concern/affective empathy to the enemy who are most violent against Walikale civilians reflecting parochial dispositions

Second, dispositions to inter-personal harm exhibit huge parochiality—they have high empathy towards Walikale and low empathy to Hutus, alongside zero inclination to harm Walikale community members and high inclination to harm Hutus. In terms of numbers, 83% choose to kill a Hutu in the video game against 1.4% when they face a Walikale member; 63% reported to enjoy the pain of a Hutu, against .2% enjoying the pain of a Walikale member. Similarly, while the average transfer to a Walikale at baseline is 3, it is -2.31 towards a Hutu.

Third, dispositions for inter-personal harm, but also prosociality, are not constant. On average, eight months after joining, joiners have significantly more inclination and less aversion to hurt the Hutu as well as lower prosociality to the enemy; inversely, they have higher levels of empathy to the Walikale civilians, but here too they have lower prosociality towards towards Walikale civilians than upon starting. Furthermore, analyzing individual levels reveals that baseline measures are largely uncorrelated with the measures eight months later (only the preference to hurt Hutus exhibits positive correlation between baseline and endline). This average change is consistent with de-sensitization to the harm of Hutus and homogeneization with recruits.

This provides implies that: a. measures of disposition to inter-personal harm reflect fundamentals that translate to contemporaneous violent behavior; b. just like standard measures of prosociality, those fundamentals can change over time; c. unbundling altruism to inclide dispositions for inter-personal harm is necessary to capture fundamentals related to violence.[21]

We complete our validation by examining whether cognitive empathy, empathic concern, and affective empathy correlate with emotional arousal. Past research on *empathy* has shown that aversion to inter-personal harm is associated to *emotional arousal* in response to witnessing another person's pain, which cannot be manipulated.[22] Figure 1, Panel B, pooling the responses to all identities, shows that, overall, our measures of empathic response reported by the respondents for a given identity are associated with higher emotional arousal in the moment in which they witness it. This association is driven by affective cognitive empathy by the joiners and is mirrored

[21]Table A2 presents this analysis without the corresponding controls, showing that both dispositions and prosociality appear to capture to some extent violent behavior, but only because they are positively correlated.

[22]The EDA, which tracks sweat gland activity controlled by the sympathetic nervous system, is often used as a measure of emotional or stress responses. Studies have shown that this kind of response can vary depending on the perceived identity of the individual in pain, which has been interpreted as reflecting unconscious biases in empathy (Avenanti, Sirigu, and Aglioti, 2010).

by households and by empathetic concern (although, for those, it is not statistically significant). This analysis provides confidence that the empathy task measures true empathetic response.[23]

In sum, the measures of disposition to inter-personal harm are predictive of (contemporaneous) true behavior and reflect true emotional arousal. Thus, they capture relevant fundamentals that, even if they are malleable, are relevant at the time of joining to the decision to join.

B. *Unbundling Altruism: The Correlation*

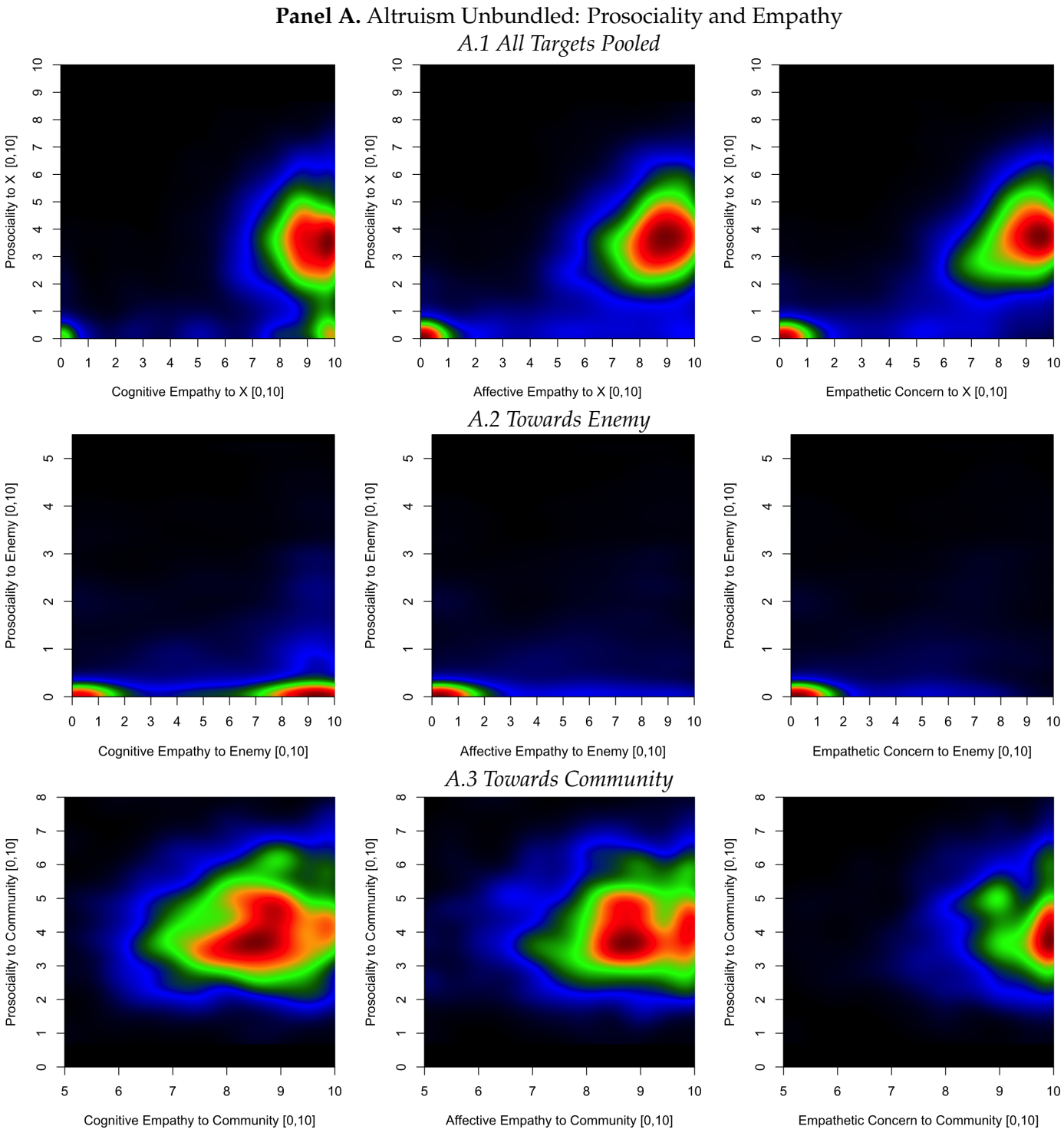
Figure 2, Panel A, presents a heatmap showing the concentration of respondents in the sample in the two-dimensional space of prosociality towards a target (y-axis) and aversion to inter-personal harm to the same target (x-axis), namely cognitive empathy, affective empathy, and empathic concern.[24] First, Panel A.1, which pools all targets reveals that there is a stark positive association between prosociality to a target and the various notions of empathy, but that this positive association is weaker for cognitive empathy. The figure shows that responses cluster in two clusters: somewhat prosocial and not prosocial, and those who are somewhat prosocial tend to have much higher levels of empathy. Thus, while empathy and prosociality are related, they appear to behave in different ways.

Second, those who are somewhat prosocial tend to have high levels of affective empathy and empathic concern, and when they are not, they have none of them. For cognitive empathy, however, in a significant mass of responses, respondents have full levels of cognitive empathy yet zero prosociality towards the target. Panels A.2 and A.3 show that this reflects that respondents tend not to care about members of the enemy group (i.e., have zero affective empathy and empathic concern) and are not prosocial to them, but are split into those who know perfectly that they suffer and those who (are able to) ignore their suffering. Responses for community members as targets are only those of the upper right cluster (those who care about, understand the pain, and are prosocial towards the target) whereas for enemy, it is the low prosociality, low empathic concern and affective empathy cluster, split by high vs. low cognitive empathy, i.e., parochial altruism (in-group bias) in all dimensions, less so for cognitive empathy.

[23]To provide further reassurance about the validity of the measures of empathy, we also benchmark the emotional responses of the respondents upon witnessing the sufferng of a Rwandan Hutu to that of a Walikale civilian. Since we know that reported empathic response is weaker to the pain of Rwandan Hutus, a test of the validity of the empathy task is if emotional arousal is also weaker. Figure 1, Panel B, confirms that civilians feel less empathy towards Hutus using the measure of emotional arousal.

[24]Inclination to harm is excluded from this figure for being a dummy but mirrors the relationships with aversion.

Figure 2: The Prosociality-Empathy Relationship

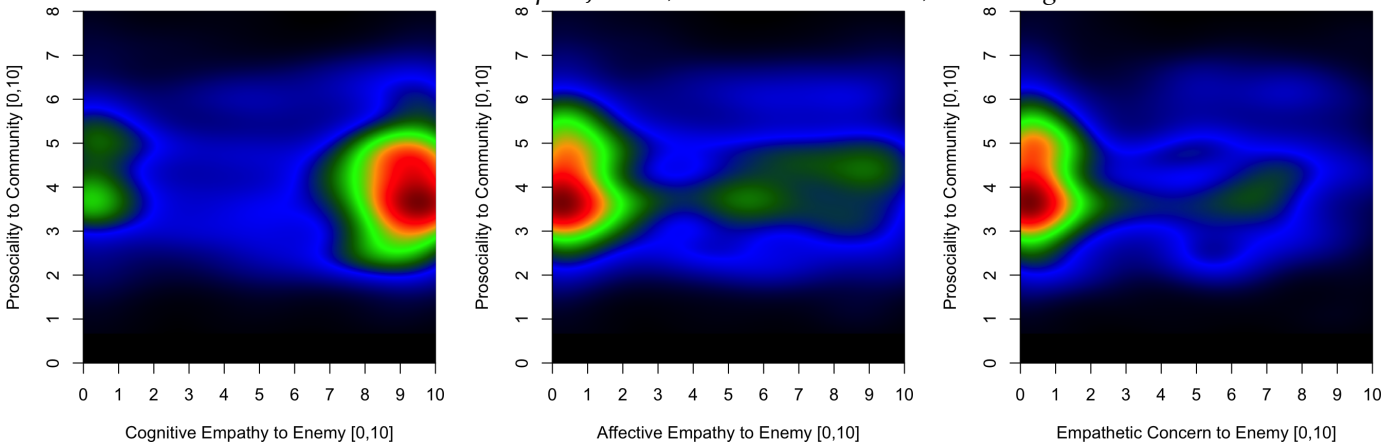


Notes. This figure presents the heatmap of prosociality (an index of declared altruism, hypothetical donation, and dictator game donation) towards target of identity X and aversion to inter-personal harm to identity X.

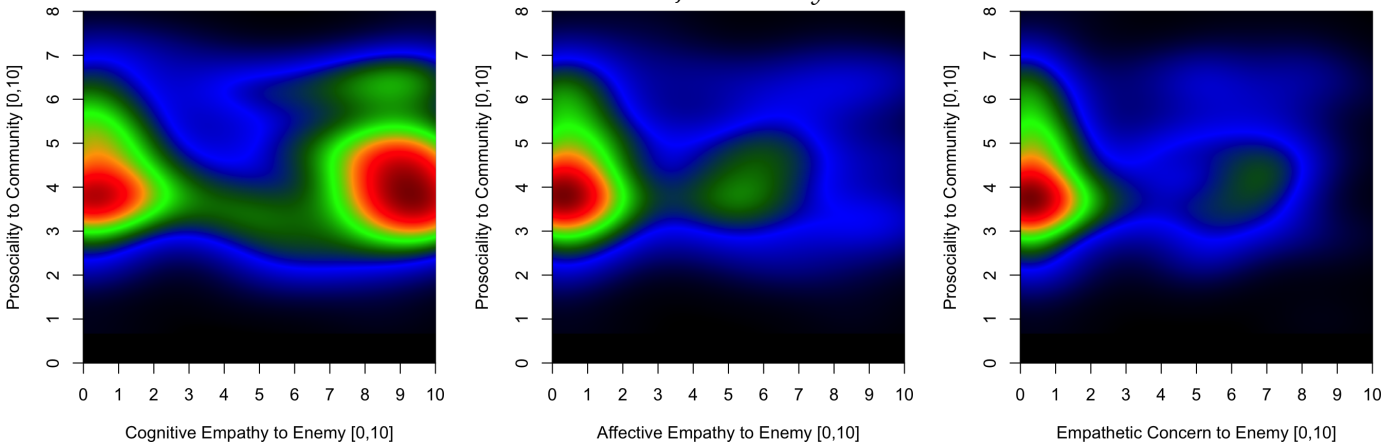
Third, Panel B shows that prosociality to community and aversion to inter-personal harm to

Panel B. They Key Correlation: Prosociality to Community and Empathy to Enemy

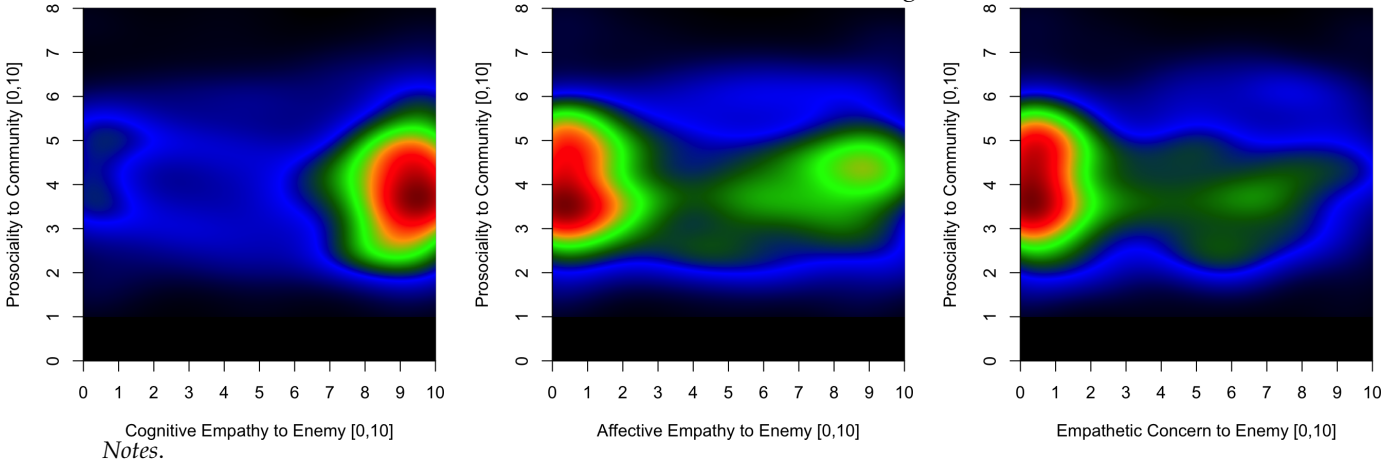
B.1 Pooled Sample (Joiners, Household Members, and Villagers)



B.2 Joiners Only



B.3 Household Members and Villagers



Notes.

members of the enemy group appear to be (weakly) negatively correlated. Interestingly, most people have reasonable prosociality towards own community and, while many tend to fully understand the suffering of the Hutus, many tend not to feel sorrow (empathic concern) nor be affected by it (affective empathy). With this empirical correlation, the framework would predict that it is those with low empathy and high prosociality who would tend to join. Panels B.2 and B.3, which separately present the heatmaps for joiners and non-joiners, show that joiners are the ones split between have full and no cognitive empathy towards the Hutu, tend to care less and be affected less by their pain, and appear to have higher prosociality to the community. Importantly, joiners tend to have much lower cognitive empathy towards the enemy than the rest.

We no turn to formally analyzing these relationships. First, in Table 2, Panel A.1., we regress measures of disposition to inter-personal harm towards a given target on the index of prosociality towards the same target. There is a stark positive association between prosociality and aversion to inter-personal harm, with it being starkest (one to one) for empathic concern, and weakest for cognitive empathy. This is consistent with the neuroscience literature which has found that empathic concern is associated to prosocial behavior. Unsurprisingly, prosociality to a target is also negatively associated to inclination to harm that target.[25]

Second, Panel A.2 examines the correlations of these notions across targets. Interestingly, prosociality to the community is positively but weakly associated to prosociality to the enemy. Empathic concern and affective empath are (weakly) positively correlated across targets; cognitive empathy is negatively correlated. This means that some individuals have more prosociality and affective empathy towards both targets than others whereas empathic concern to a target and another are uncorrelated but individuals who hold more cognitive to a target have less to the other, suggesting cognitive empathy are substitutes, akin to limited attention. This also suggests that interventions like perspective taking could induce reduction of empathy to other groups.[26]

Third, Panel B shows the key correlation separately for the sample of joiners and household members (Panel B.1.) and for joiners and randomly selected villagers (Panel B.2). The analysis always fixed effects for joiner ID, shared between each joiner, its randomly sampled household member, and its randomly sampled villager, to exclude the selection of comparison groups. The

[25]Table A3 includes individual and, separately, target fixed effects, and shows that the correlation of these measures across targets is in large part explained by variation across targets rather than variation across individuals.

[26]Replications using only the sample of joiners and household members or with different fixed effects consistently show that this coefficient is statistically significant at the 1% significance level.

Table 2: The Prosociality-Empathy Relationship

Panel A. Unbundling Altruism

A.1. Empathy / Sadism to X on Prosociality to X

	Empathy to X			Sadism to X	
	(1) Concern	(2) Affective	(3) Cognitive	(4) Preference	(5) Taste
Prosociality to X	1.073***	0.890***	0.526***	-1.303***	-1.081***
	(0.032)	(0.032)	(0.030)	(0.049)	(0.042)
Mean dep. var	68.11	68.26	77.19	30.05	19.29
R^2	0.44	0.33	0.15	0.33	0.33
Observations	1902	1902	1902	1914	1910

A.2. Unbundled Altruism

	Prosoc. E.	Empathy to Enemy			Sadism to Enemy	
	(1) Index	(2) Concern	(3) Affective	(4) Cognitive	(5) Preference	(6) Taste
Prosociality Community	0.115**					
	(0.051)					
Empathic Concern Community		0.029	0.041	1.228***	-0.443	-0.209
		(0.248)	(0.267)	(0.258)	(0.271)	(0.342)
Affective Empathy Community		0.055	0.427**	0.038	0.296	0.159
		(0.196)	(0.210)	(0.248)	(0.260)	(0.312)
Cognitive Empathy Community		0.209	0.027	-0.398	0.060	0.242
		(0.219)	(0.245)	(0.279)	(0.311)	(0.374)
Joiner ID FE	✓	✓	✓	✓	✓	✓
Mean dep. var	7.45	33.83	39.69	59.43	77.31	55.95
R^2	0.33	0.32	0.33	0.35	0.35	0.34
Observations	638	634	634	634	634	634

Notes. In Panel A.1., we regress measures of disposition to inter-personal harm towards a given target on the index of prosociality towards the same target. Panel A.2 examines the correlations of these notions across targets.

Panel B. The Key Correlation

	Empathy towards the enemy						Sadism towards the enemy	
	(1) Concern	(2) Affective	(3) Cognitive	(4) Concern	(5) Affective	(6) Cognitive	(7) Preference	(8) Taste
B.1 Joiners and Household Members								
Prosociality Community	-0.192	-0.271*	-0.554***	0.035	-0.008	-0.372**	0.126	0.292
	(0.138)	(0.156)	(0.176)	(0.080)	(0.089)	(0.145)	(0.198)	(0.210)
Empathic Concern Enemy					0.764***	0.165		
					(0.060)	(0.107)		
Affective Empathy Enemy				0.719***		0.556***		
				(0.047)		(0.108)		
Cognitive Empathy Enemy				0.059	0.210***			
				(0.038)	(0.048)			
Joiner ID FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean dep. var	33.23	39.29	59.79	33.23	39.29	59.79	76.06	56.69
R^2	0.49	0.49	0.46	0.84	0.86	0.67	0.50	0.49
Observations	418	418	418	418	418	418	424	424
B.2 Joiners and Random Villagers								
Prosociality Community	0.188	0.298	-0.021	-0.014	0.156	-0.222	-0.022	-0.388
	(0.243)	(0.304)	(0.299)	(0.168)	(0.190)	(0.211)	(0.239)	(0.334)
Empathic Concern Enemy					0.774***	0.354**		
					(0.085)	(0.150)		
Affective Empathy Enemy				0.688***		0.451***		
				(0.075)		(0.144)		
Cognitive Empathy Enemy				0.132**	0.189***			
				(0.061)	(0.071)			
Joiner ID FE	✓	✓	✓	✓	✓	✓	✓	✓
Mean dep. var	31.79	37.42	54.41	31.79	37.42	54.41	80.99	58.45
R^2	0.42	0.43	0.45	0.85	0.85	0.71	0.53	0.46
Observations	416	416	416	416	416	416	424	424

Notes. Panel B shows the key correlation separately for the sample of joiners and household members (Panel B.1.) and for joiners and randomly selected villagers (Panel B.2). Since the comparison groups are tied to each joiner, they are self selected, hence the analysis always includes fixed effects for joiner ID, which is shared between each joiner, its randomly sampled household member, and its randomly sampled villager.

analysis reveals substantial heterogeneity in the key correlation. Within the households of the joiners, individuals who are more prosocial to the community tend to have less affective empathy and especially less cognitive empathy to the enemy. However, this relationship only holds for cognitive empathy when controlling for other empathic concern and affective empathy, suggesting that the raw correlation between prosociality to the community and affective empathy to the enemy was explained by its correlation with cognitive empathy. Similarly, within households, more prosocial individuals are also somewhat more inclined to harm the enemy. In contrast, Panel B.2. shows that the same relationships are weaker.

In sum, the key correlation is consistent with case III in our framework. If prosociality to the community and aversion for harming hutus matter in the decision to join, our framework then predicts that joiners should exhibit more prosociality to the community and less aversion to inter-personal harm towards the Hutu, and that this difference should be largely driven by within household comparisons between joiners and their household members who did not join.

6. How Distinctive is the Altruism of those Who Volunteer the NDC-R?

Having shown that individuals who are more prosocial to the community are also less empathetic to the enemy, in this section, we analyze the selection implied by this correlation according to the simple framework: that, if altruism factors into the decision to join, volunteers should have more prosociality to the community and less empathy to the enemy.

A. *The Selection into Fighting*

We now examine whether new volunteers have less empathy towards the enemy and more prosociality towards the community as suggested by the framework of Section 3 and the key correlation. Let i index the individual. We estimate the following equation in OLS:

$$y_i = \alpha + \beta J_i + \epsilon_i \tag{1}$$

where y_i is a component of altruism of individual i of interest, and J_i is an indicator taking value one if individual i belongs to the set of new volunteers, zero if they belong to the set of civilians. To delineate whether the selection of altruism operates within or across households, we estimate Equation 1 separately for the sample of new volunteers and: a. the randomly selected youth in their own household; b. the randomly selected youth in the village.

Table 3 presents the means of disposition for inter-personal harm (Panel A) and of prosociality (Panel B), separately for the randomly selected villagers (Column 1) and for the randomly selected individuals in the joiners' own households (Column 3), along with the difference between the joiners and these respective two samples (Columns 2 and 4). The difference, as well as the p-value of the two-sided test of whether the difference is different from zero, is from Equation 1.

First, Panel A shows that volunteers have significantly more disposition for inter-personal harm to the Hutus than the non volunteers, but they are just as disinclined for violence against random people. Considering civilian household members, the means in Panel A show that, while civilians hold high levels of empathy towards Walikale members, they do not towards Hutus. Similarly, while civilian household members hold little to no inclination to harm members of the Walikale community, the majority experiences pleasure to the pain of Hutu (*taste* for harming Hutus), and 70% chooses to murder a Hutu in the video game (*preference* for harming Hutus). Considering the joiners, Column (2) shows that joiners have indistinguishable levels of empathy towards Walikale compared to the average civilian, but they have significantly less empathy towards the Hutu than the average civilian. Similarly, while they are just as unlikely to have inclination to harm community members, they have a larger inclination to harm the enemy than the average civilian's (especially compared to volunteers' own household members). Consistent with the auxiliary prediction presented in the previous section, the selection of inclination to harm Hutus is stronger within than across households—even if it is evidently there across households. Importantly, this does not mean that joiners have less capacity for empathy: on the contrary, they score significantly lower in primary and secondary psychopathy that the civilians.

Second, Panel B shows that volunteers are significantly more prosocial towards their own community than non volunteers (and have just as low prosociality towards the enemy than the non volunteers). Considering the civilians, they have reasonably high levels of prosociality towards members of Walikale, whereas they tend to have lower levels of prosociality towards the Hutu.[27] Column (2) shows that, compared to the average civilian joiner household member, joiners tend to be more prosocial individuals towards the community.

The distinctiveness of joiners in their disposition to harm Hutus, as well as in their prosociality towards civilians, is weaker when comparing the joiners to randomly selected villagers. This

[27]The mean donation in our dictator games towards a random person from the district is 35%. Summarizing student populations on which the dictator game is usually played, Henrich, Boyd, Bowles, Camerer, Fehr, Gintis, McElreath, Alvard, Barr, Ensminger, Henrich, Hill, Gil-White, Gurven, Marlowe, Patton, and Tracer (2005) conclude: "In the DG, modal offers are typically 0% and means usually fall in the 20% to 30% range."

Table 3: The Distinctive (Unbundled) Altruism of those who Join the NDC-R

	Household (Mean)	Joiner - HH (Difference)	Villager (Mean)	Joiner - Vill (Difference)	Pooled (Mean)	Joiner - Pooled (Difference)
	(1)	(2)	(3)	(4)	(5)	(6)
A. Aversion/Inclination to Interpersonal Harm						
A.1. Aversion/Inclination to Interpersonal Harm on the Enemy						
Empathy to enemy (Index)	50.48	-12.66 *** ***	44.74	-6.93 **	47.62	-9.80 *** ***
Cognitive Empathy	69.37	-19.06 *** ***	58.70	-8.39 *	64.05	-13.74 *** ***
Affective Empathy	44.20	-9.73 *** ***	40.50	-6.02	42.35	-7.88 *** **
Empathic concern	37.86	-9.20 *** ***	35.02	-6.36 *	36.45	-7.79 *** **
Sadism to enemy (Index)	60.45	11.72 *** ***	67.14	5.03	63.79	8.38 ** **
Taste (% enjoys pain of person)	50.94	11.33 ** **	54.46	7.80 *	52.70	9.56 ** **
Preference (% murders in video game)	69.95	12.12 *** ***	79.81	2.26	74.88	7.19 ** *
A.2. Aversion/Inclination to Interpersonal Harm on Community						
Empathy to community (Index)	85.75	0.19	87.44	-1.50	86.59	-0.65
Cognitive Empathy	82.58	0.91	84.25	-0.75	83.41	0.08
Affective Empathy	83.24	1.19	86.53	-2.11	84.88	-0.46
Empathic concern	91.43	-1.53 *	91.53	-1.63 *	91.48	-1.58 * *
Sadism to community (Index)	0.47	0.36	1.41	-0.58	0.94	-0.11
Taste (% enjoys pain of person)	0.47	-0.23	0.00	0.24	0.23	0.00
Preference (% murders in video game)	0.47	0.95	2.82	-1.40	1.64	-0.23
A.3. Psychopathy						
Primary Psychopathy	34.69	-1.52 ** **	34.41	-1.23 *	34.55	-1.37 ** *
Secondary Psychopathy	17.15	0.66	18.97	-1.16	18.06	-0.25
B. Prosociality						
B.1 Prosociality to community (index)	37.64	6.94 *** ***	44.40	0.18	41.02	3.56 *** ***
Income-sensitive prosociality measures	21.80	8.20 *** ***	31.58	-1.58	26.69	3.31 ** **
Hypothetical donation	26.67	4.51 ** **	31.78	-0.60	29.23	1.95
Dictator game (in %)	20.80	10.15 *** ***	35.96	-5.02 *	28.38	2.56
Index long sample measures	23.73	7.33 *** ***	33.87	-2.81 *	28.80	2.26 *
Time dictator game working (in %)	18.32	8.53 *** ***	22.04	4.81 **	20.18	6.67 *** ***
Time dictator game average (in %)	21.86	10.80 *** ***	28.76	3.90	25.31	7.35 *** ***
Index short sample measures	19.88	9.96 *** ***	26.51	3.33	23.20	6.65 *** ***
Income-robust prosociality measures	45.97	9.03 *** ***	49.69	5.31 ** *	47.83	7.17 *** ***
Declared prosociality	65.45	6.16 *** ***	65.45	6.16 *** **	65.45	6.16 *** ***
Socializing dictator game (in %)	21.57	8.80 *** ***	25.38	4.98 *	23.38	6.99 *** ***
B.2 Prosociality to enemy (index)	6.65	-0.14	9.09	-2.58 **	7.87	-1.36
Income-sensitive prosociality measures	4.28	-0.77	6.38	-2.87 ** *	5.33	-1.82 ** *
Hypothetical donation	2.86	-0.13	3.94	-1.21	3.40	-0.67
Dictator game (in %)	2.39	-0.13	5.26	-2.99 ** *	3.83	-1.56 *
Index long sample measures	2.63	-0.13	4.60	-2.10 ** *	3.62	-1.12 *
Time dictator game working (in %)	5.64	-1.06	6.75	-2.18	6.20	-1.62
Time dictator game average (in %)	6.98	-1.98	8.14	-3.14 *	7.56	-2.56 *
Index short sample measures	6.36	-1.47	7.48	-2.60 *	6.92	-2.04
Income-robust prosociality measures	11.43	-2.72	13.76	-5.05 **	12.60	-3.89 **
Declared prosociality	14.69	-0.17	18.08	-3.55	16.38	-1.86
Socializing dictator game (in %)	6.88	-1.69	6.75	-1.57	6.81	-1.63
Target Observations	213	426	213	426	426	639

Notes. This table presents the means of the measures of prosociality and disposition for inter-personal harm along the estimates from Equation 1.

weaker difference is also consistent with the fact that the key correlation is also weaker within villages than it is within households. Overall, pooling all civilians (joiners' household members and random civilians) together, joiners stand out in their inclination to harm the enemy and their distinctively high levels of prosociality to the community.

Robustness to economic factors. One concern with this analysis is that comparing standard measures of prosociality across individuals is sensitive to income effects, especially when comparing systematically different populations, such as joiners and non joiners. This issue is especially important in our context given that there are good reasons to expect joiners to also be more economically deprived. To tackle this identification challenge, we developed two strategies.

The table separately shows the results for income-sensitive measures as well as measures that were designed to be income-robust measures of prosociality. These include the standard declared altruism as well as the time dictator game in which the counterfactual use of time is socializing with the family. The classification reveals that, compared to their own household members, joiners are more prosocial to the community regardless of the measure and that, compared to random villagers: a. on income-robust measures of prosociality, joiners are more prosocial; b. joiners have lower income-sensitive measures of prosociality, regardless of the target than the average villager; c. accounting for the average lower level of income-sensitive prosociality by joiners across targets, they are differentially more prosocial to the community compared to their difference from civilians to the enemy. These facts are consistent with joiners having more marginal utility of income than random villagers (depressing the income-sensitive measures of prosociality) and at the same time having more prosociality to the community.

An implication from this analysis, if joiners truly are more prosocial than their random villagers, is that joiners should also be poorer, which would explain their higher marginal utility of income. Table A6 thus shows that, unsurprisingly, new volunteers were also more likely to experience unemployment and to have a low paying job before joining than the rest, reflecting that economic incentives are also likely important. Does this economic selection explain the selection on altruism? Table A5 replicates the analysis of Table 2, controlling for an array of economic variables (unemployment, past income, education, etc). The results are unaffected by the controls, which provides reassurance that, while other factors cannotbe ruled out, the pattern of selection along the lines of altruism cannot be explained by economic incentives.

In sum, this section has shown that joiners are the most prosocial individuals to their community and that they are also less averse to inter-personal harm of the Hutus, as predicted by

the negative correlation between these two components of altruism. Understanding how this unbundling of altruism and negative correlation comes about in certain contexts, as well as what motivates the most prosocial individuals to also join violent organizations, requires a serious look at the prevailing society's morality. We examine that next.

7. Explaining the Result: The Role of Morality

To interpret these findings, in this section, we ask three questions. First, we analyze society's prevailing morality regarding violence towards Hutus. Second, we ask whether the negative correlation follows from the fact that prosocial individuals apply societal moral values more strongly, or because they hold more violent moral convictions. Third, we ask whether prosocial individuals develop moral convictions to justify a stronger inclination for harming the Hutus.

A. *The Prevailing Morality of Violence*

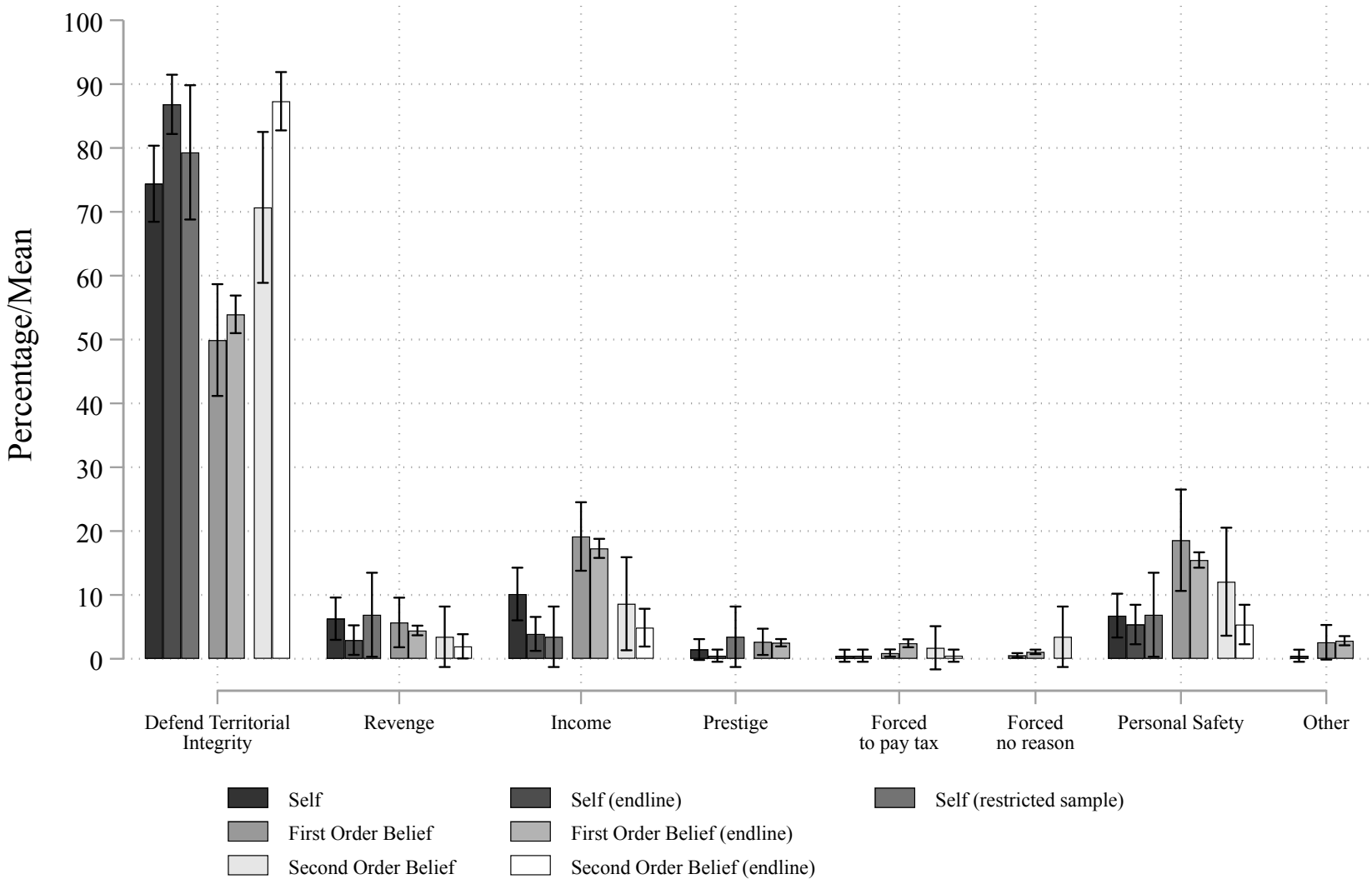
This section presents basic facts about the morality of retaliatory violence prevailing in the catchment area population. Since joiners represent only 2 to 5% of the entire population, then the analysis in this section uses only the sample of randomly selected civilians in the corresponding village of the joiners who did not join the NDC-R in the catchment villages of the NDC-R.

Fact 1. Joiners join to contribute to their district, and civilians perceive joining to be morally right. First, Figure 3 presents the fraction of respondents who report to have joined for a given reason (henceforth, self-reports). To answer this question, respondents were asked to speak freely, and the enumerator selected the answer, reducing concerns about priming during the interview. An overwhelming majority of respondents report to have joined in order to contribute to defend the community's territorial integrity. To overcome social desirability bias about the respondent's own motivation, we also asked the new joiners why they thought other volunteers joined (henceforth, first order beliefs).[28] First order beliefs corroborate that more than half of joiners are believed to have joined for the prosocial motivation to defend territorial integrity. Overall, the fraction drops from 75% to 50% indicating some social desirability bias by the respondent (or simply misperceptions about others' motives), but the dominant role of prosocial motivations remains. To complement this analysis, we also asked respondents to guess, if other joiners were asked about the reasons why the respondent joined, what would they say about why the

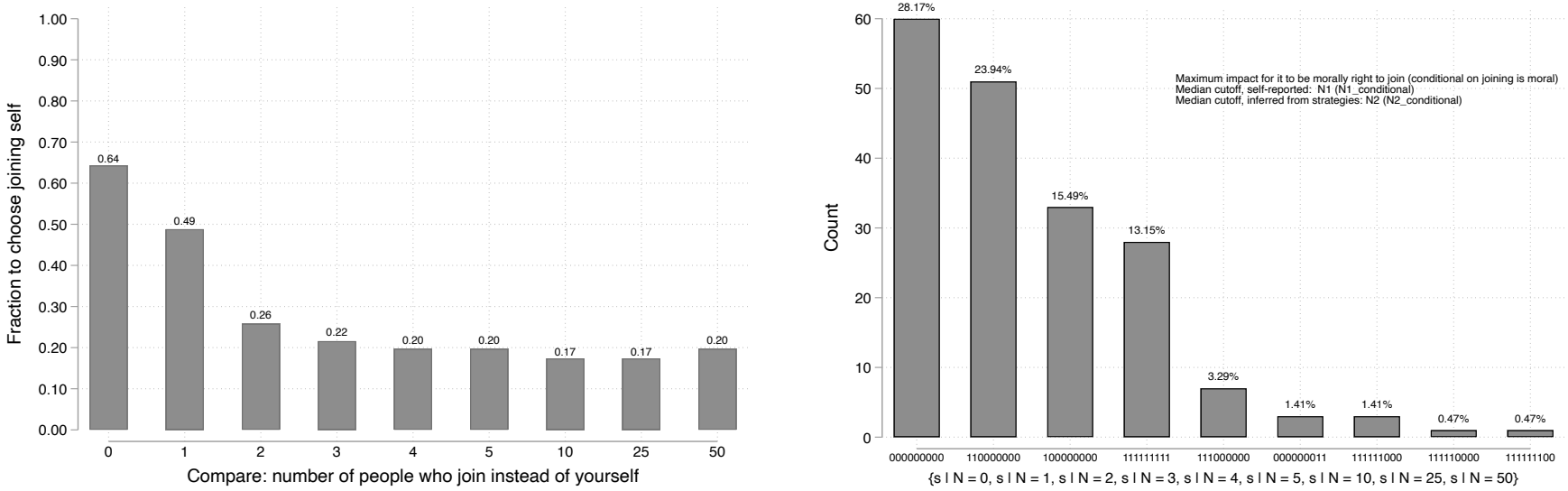
[28]We asked the respondents to allocate 100 hypothetical joiners to bins representing different motives for why these other joiners might have joined. The results are the same.

Figure 3: The Prevailing Morality of Violence

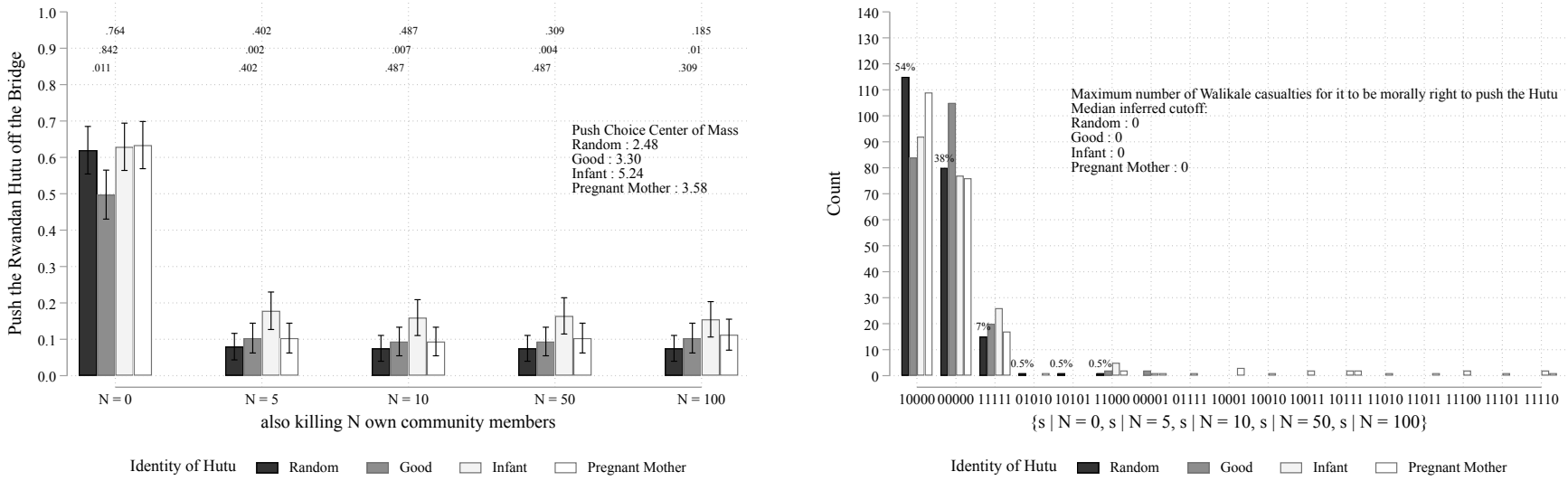
Panel A. Joiners Join to Contribute to their District (Fact 1)



Panel B. Civilians' Moral Convictions are Predominantly that Joining is Morally Right (Fact 1)



Panel C. Civilians' Moral Convictions are Predominantly that Killing Hutus is Morally Right (Fact 2)



Notes. Panel A presents the mean behavioral strategies (left) and strategies count (right) for the joining dilemma, asking respondents whether they think it is morally right to join the NDC-R, when their decision to join would imply that $N \in \{0, 1, 2, 3, 4, 5, 10, 25, 50\}$ other equally qualified members would not be able to join. The term $s|N \in \{0,1\}$ denotes the choice of whether to join ($s = 1$) or not join ($s = 0$) when joining implies that N other equally qualified others would not be able to join. Panel B presents the mean choices (left) and strategies count (right) for the killing moral dilemma (destruction bridge dilemma). It presents it for dilemma in which respondents are asked to ascertain whether pushing a Rwandan Hutu off the bridge is morally right, when it results in the death of $N \in \{0, 5, 10, 50, 100\}$ Walikale civilians. Various versions of this dilemma were implemented, with the identities of the Rwandan Hutu in the following set: {Random Rwandan Hutu, Harmless Rwandan Hutu, Infant Rwandan Hutu, Pregnant Mother Rwandan Hutu}.

respondent joined (henceforth, second order beliefs). The responses to this question again show an overwhelming majority of respondents indicating that they believe others would believe that they joined because of their motivation to defend territorial integrity, almost identical to their reported motivations. That is, prosocial motivations are important in the decision of join.

Second, Figure 3, Panel B, left panel, presents the fraction of civilians who select that the morally right answer is to join, as a function of the number of others who would not be able to join. Two-thirds of civilians consider it morally right to join the armed group (despite the fact that they themselves have not joined). Of those, 46% significantly value the *consequence* of their decision to join for society and 13.5% have essentially $s_C^c \approx 0$, that is, they consider joining as a moral duty regardless of the consequence. The right panel shows the full strategies played as a function of the number of other joiners who cannot join. It implies that, among those who consider joining morally right because of the consequence of joining: a. 15.5% are "pure prosocial consequentialists" ($-s_C^c < d \leq 0$); these are those who think joining is right when at the expense of none but not joining is right as soon as someone else would join instead; b. 24% care more about the consequences than the benefits to self, although they also carry positive "selfish" motives ($s_C^c > d > 0$) and could be labeled "mixed prosocial consequentialists;" c. 18% care about the consequences, but less so than the its benefits to self ($50s_C^c > d > s_C^c > 0$).[29] In sum, a majority of civilians, none of whom had chosen to join at the time of the interview, consider joining the NDC-R to be a morally right choice, most of whom because of the consequence of joining to society. This fact underscores conceptualizing joining as prosocial.

Fact 2. Civilians perceive murdering Rwandan Hutus to be Morally Right. First, most civilians select killing the Rwandan Hutu when there is no other consequence, and some even at the expense of Walikale civilians. Figure 3, Panel C, reports the results of the trolley bridge destruction moral dilemmas. Whereas 60% of civilians think it is morally right to kill the Rwandan Hutu if there are

[29]The remaining 2% misunderstood the game as their strategies are nonmonotonic. First, consider someone who selects that the right choice is to join when joining is at the expense of zero others. In that case $U_1(0) \geq U_0(0)$, which implies that $d + s_C^c \geq 0$. For those individuals, clearly, the bundle of "selfish" (material, empathy, and deontological moral motive) and the consequentialist motivations are large enough that joining is better. Consider now, among those individuals, the decision of whether to join when joining is at the expense of another person joining. If the person choose joining, then we have that $U_1(0) \geq U_0(1)$, which alone implies $d \geq 0$; that is, a person choosing to join when it is at the expense of another joiner for sure has a positive bundle of "selfish" economic and moral motivations. But a person who chooses to switch at the point, that is, to let another person join instead of them has $U_1(0) \geq U_0(1)$ and $U_1(0) \geq U_0(0)$, implying $-s_C^c < d \leq 0$. That is, those who would join without trade-off but would concede as soon as one person could join instead of them have necessarily a larger consequentialist prosocial component than their bundle of "selfish" motivations. Similarly, someone who chooses to join when it is at the expense of zero persons, and when it is at the expense of one, and when it is at the expense of two will have that $U_1(0) \geq U_0(2)$ implying $d \geq s_C^c$, whereas if they switch at two, they have $0 < d \leq s_C^c$. A similar logic allows us to place percentages of people whose prosocial consequentialist parameter is within various segments—up until it is insignificantly low.

no casualties from the own community, 15% deem it morally right to kill the Hutu regardless of the number of casualties of their community.[30] Second, to separately identify the conviction that it is a moral duty to kill a Hutu (*deontological* moral conviction) from the *consequentialist* conviction that it is right because of its instrumental value for an end (for example, if it is perceived to save more lives of the community), Panel A presents the "placebo" moral dilemma, in which the person to be killed is a harmless Hutu—in which case, the perceived effect of their murder on Walikale society is switched off. Among the 60% who thought that killing the Hutu was right, 10% no longer consider it moral in this version—and the decrease is statistically significant. This indicates that, among the 60%, only 10% held a *consequentialist* moral motivation; the remaining 50% held a *deontological* morality of killing Rwandan Hutus. Third, this morality is, for a small set of people, *genocidarian*. Indeed, Panel A shows that the fraction of respondents who deem it morally right to kill the Rwandan Hutu rises from 15 to 28% if the Rwandan Hutu is an infant (same if they are a pregnant mother), consistent with a "genocidarian" deontological morality. Furthermore, Figure A9 shows that this morality is informed by a moral conviction that it is right to kill a large number of Rwandan Hutus. Overall, 76% of civilians consider it morally right to kill at least 101 Rwandan Hutus with no casualties for Walikale.[31]

Fact 3. This morality is intimately connected to past violence by the Rwandan Hutus. To elicit an explanation, respondents were presented images of a Rwandan Hutu civilian expressing suffering and carrying with scars (first infants, then a pregnant mother); surveyors explained that these scars were caused by lethal violence by one Congolese armed group.[32] Respondents were then asked whether they felt good, the degree of immorality, whether they felt guilty and, in the case of the pregnant mother, they were offered an opportunity to articulate their feelings. Table 4, Panel A, shows that 40% of respondents actually feel good about the images. Of those who feel good, the mean reported immorality is in the middle of the distribution, and only 18%

[30] The moral value held by the latter group of civilians is consistent with "collective narcissism:" "*A belief that the exaggerated greatness of one's group is not sufficiently recognized by others*" (De Zavala, 2023)—it is not simply in-group bias: it entails specifically out-group prejudice to benefit the group's sense of worth even if it costs individuals of one's group. We are thankful to Jean-Francois Bonnefon for making this suggestion.

[31] Figure A10, Panel A, shows that these results are not a reflection of a generalized morality of killing anyone, and is uniquely driven by the moral conviction that killing Rwandan Hutus is moral. Panels B and C present results from the standard bridge dilemmas in which those on the rail could be saved, rather than sacrificed, by pushing the person off the bridge. They show that, while most civilians deem it morally right to kill the Rwandan Hutu as soon as their death saves a small number of community members, they consider it morally wrong to kill a community member to save Rwandan Hutus, even if the numbers of victims are large. In fact more than 60% of respondents select that they should let all Hutus die rather than sacrifice one member of their community to save all Hutus.

[32] A prime was introduced before asking the respondents anything, containing moral condemnation of such violence in order to trigger respondents' moral justification.

Table 4: The Prevailing Moral Values—Why is Killing Morally Right? (Fact 3)

	Infant		Pregnant Mother	
	Feels Good	Feels Bad	Feels Good	Feels Bad
Precentage	40.2%	59.8%	41.18%	58.82%
Mean Degree of Immorality [0-5]	2.54	4.18	2.81	4.23
Percentage Feels Guilty	18.33%	81.67%	17.24%	82.76%

Feels Good		Feels Bad
I would much rather massacre all Hutus in district		Hutus are immoral
It is good to turture them this way		They massacred our families, they deserve it
They came to our land to kill, we defended us		This follows Hutus stealing, raping, killing us
They had the same impacts in our population		This would be wrong, but given how they acted…
They did these massacres to our population		Nothing, eye for eye tooth for tooth (Bible)
Hutus did a lot of massacres in Congo		Because Hutus steal and rape our siblings
They deserve the massacres they did to us		Innocents are dying. Rwanda must take them
They made us suffer a lot		The Hutus saw an involuntary situation
They also harmed us in previous years		This is the fault of Tutsi-Hutu problem
They are bad people, massacred us		I feel very good, the Hutus aggressed us
They should leave our country		They are right, Hutus FDLR did the same
Who attacks with the sword will die with the sword (Bible)		Pushing them back so that we can be calm
Mai mais are just defending territorial integrity		Feel guilty, but it is our self-defense (x2)
When Hutus take us, they treat us the same way		This is a result of their own misconduct (x3)
Even better if it is executed by our mai mais		They are right; Hutus came as bandits
They are invaders		It's due to their own behavior
Hutus are Rwandese, they are the enemy		They also wronged our population
They are criminal people		Good or nothing (x5)
Mai mais do right, Hutus don't want to go home	*	It's a bad behavior (x2)
The Hutus did the same acts	*	Bad, bit bad, pity, nostalgy, very bad (x19)
Hutus FDLR do same agression to our people	*	They really suffered
We reciprocate the actions	*	Bible says not to kill, it's not good
Their Rwandan Hutu brothes do the same acts	*	It's bad to massacre Hutus
I feel good because they are rebels	*	Bad in heart, they are humans like other (x2)
The FDLR are the root of massacres here in our country	*	I feel pity for them (x2)
The one who kills by shooting will die by shooting	*	Bad because it is immoral (x2)
They also massacred us much more than this	*	Very bad our brothers did this to her
Our mai mai brothers helped us push them back	*	This killing doesn't help us. We need cohesion
They are the ones who started massacring us	*	What mai mais did to her is bad
They shouldn't stay here, they should go back home	*	We are all Congolese siblings
They [the Hutu] are murderers	*	They did a big sin
Tey started massacring the congolese	*	It hurts the heart really
The mai mais diminish our enemies	*	It's bad to massacre people
Very good (x5)	*	If the mai mai attacks them, it's wrong
Good (x7)	*	Bad (x10)

Notes. This table presents the answers to the module in which respondents are asked to express their opinion and potentially justify the violence perpetrated by the Congolese Mayi Mayi (in this case the NDC-R) on the enemy (the Rwandan Hutus). To trigger them expressing their opinion and justify the violence when needed, respondents are shown the image of: a. a Rwandan Hutu infant with wounds; b. a Rwandan mother with wounds, and are told that the wounds were caused by the Mayi Mayis. The respondents are then asked for both a and b to indicate whether they feel good or bad about it, to rate the degree of morality of the image, and to express whether they feel guilty. For b, they are also asked to articulate, in an open ended question, how they feel in their heart about it. The star indicates that the text to the right expresses moral condemnation of the violence. Sentences without the star are expressions of justification of the violence. The following sentence, extracted from a civilian respondent in a volunteer's household, provides an elaborate explanation, pointing at feelings of betrayal and frustration: "*I feel good because these people had come as refugees. We had helped them with fields so they could cultivate while they were still in Congo, preparing for their return to Rwanda. We had even given them food supplies and cooking pots so they could settle in the country. We even gave them our daughters in marriage just so they would feel comfortable. But later, these people turned against us, starting to act as if they were the owners of all our possessions. They had now taken up their weapons, which they had crossed the border with in 1994 during the Rwandan genocide. This is why, to this day, these people are destabilizing the eastern part of the country, the Democratic Republic of Congo. They rape, kill, and steal. So when I see such an image, I feel good because they are unconscionable, those people.*"

feel guilty about it; of those who feel bad, mean reported immorality is 4 out of 5, and 82% feel guilty, suggesting that the unease might be driven by morality and guilt. The panel also shows that the civilians' justification reflects a logic of retaliation ("*they deserve it because they this to us*") and that, while 100% of those who feel good report it to be justified, 34% of those who feel bad still find it justified.[33] Their responses imply that, overall, 67% find the violence against the Rwandan Hutus morally justified because of what they did—coinciding with the fraction of civilians who find it morally right to kill a Rwandan Hutu when there are no consequences.

B. *Mechanisms I: Stronger Upholding of Moral Values or Stronger Moral Values?*

We have shown that society's prevailing morality tends to not only consider morally right to join, but also to consider it right to kill Rwandan Hutus. Given evidence in neuroscience that moral values can override natural aversion to inter-personal harm, we have conjectured that the key correlation could be induced if prosocial individuals apply prevailing moral values more often, resulting in more often overriding the natural aversion to interpersonal harm. However, it is possible that prosocial individuals simply have more violent-friendly moral values.

Table 5, Panel A, separately examines the implications from these alternative mechanisms. Columns (1) and (2) present the coefficients of regressing dummies for the moral values on prosociality to the community. Prosocial individuals are more likely to hold the value that it is morally right to join, but they are not more likely to believe it is right to kill a Hutu. This rules out that prosocial individuals have more suppressed empathy towards the Hutus because they have more violent moral values. Columns (7)-(10) replicate the analysis of the relationship between prosociality and disposition for inter-personal harm to the enemy, but include as controls the dummies for moral values. Consistent with the existing literature in moral psychology, moral convictions that it is right to kill (or that it is right to join the armed group) are associated with weaker aversion to inter-personal harm of the Hutu. Column (11) confirms that controlling for individually-held moral values, for disposition for inter-personal harm, joiners remain more prosocial individuals who hold lower cognitive empathy to the enemy.

Two patterns are crucial to understand how moral values affect our key correlation. First, unlike all other measures of aversion to inter-personal harm, individually-held moral conviction about the appropriateness to harm Hutus are uncorrelated with cognitive empathy to the enemy. That is, the harmful implications of the moral conviction to harm Hutus are intended. Sec-

[33] These claims are based on true victimization. The median number of attacks per household is 2 (Table A4).

ond, despite the inclusion of these moral convictions as controls, prosociality remains strongly negatively associated with cognitive empathy, whereas the coefficients and significance on the other dimensions of disposition to harming Hutus disappear; that is, joiners are more prosocial regardless of their moral convictions, and they also have lower cognitive empathy to the enemy but no other different disposition to inter-personal harm. This suggests that aversion to harming Hutus that is relevant to the decision to join is governed by cognitive empathy.

Taken together, these results indicate that while the individually-held moral convictions about the appropriateness to harm enemies are related to individually-held inclination to harm enemies as well as empathic concern and affective empathy towards enemies, they cannot explain cognitive empathy to the enemy. Concomitantly, they explain away the relationship between prosociality and all dimensions of altruism to the enemy except cognitive empathy. That is, the correlation and the decision to join are governed by cognitive empathy, a distinctive mechanism than those related to individually held moral convictions.

C. *Mechanisms II: Moral Convictions or Motivated Reasoning?*

Moral values may just be motivated beliefs to justify perpetrating violence, itself motivated by anger and revenge.[34] To better understand the cognitive processes that allow the civilians to hold such moral values, we administered Bandura (2011)'s moral disengagement inventory. Research in the psychology of violence and genocide has identified cognitive processes that tend to make it less psychologically costly for individuals to support violence, and that tend to be prevalent in narratives that support violence (Bandura, 2011). Alongside de-humanization, these include moral justification, euphemistic language, advantageous comparison, diffusion of responsibility, displacement of responsibility, distorting of consequences, and blaming of the victims.[35]

We consider the following informal framework of motivated reasoning to examine its role in this relationship, inspired by Bénabou (2015): for some reason, be it past victimization, revenge motivations, anger, or other emotions, individuals may be inclined to harm Hutus. This

[34]Indeed, Figure A11 shows the text analysis for a module measuring emotions, asking respondents to describe in words the last time they felt anger. Most of the responses are related to the war, killing, and invasion.

[35]Table A7 shows the breakdown of the Bandura measurement tool administered on the civilians, separately by choices made in the killing moral dilemma. Overall, civilians engage in moral disengagement rationalizing the righteousness of hurting Hutus. The most prevalent tactics are moral justification, driven by "*it is ok to beat a Hutu because they hurt Walikealan population*," and attribution of blame to the victim (68%), driven by "*Hutu who have been abused usually did things to deserve it.*" Table A8 presents the breakdown of moral values by moral disengagement, alongside the measures of altruism, as well as of anger, and of attacks by Hutus. They show that civilians who hold violent moral values also hold much more moral disengagement, lending support to the possibility that moral values are related to motivated reasonning.

Table 5: Mechanisms I: Prosociality and Moral Values

	Moral Values		Sadism E.	Empathy to Enemy			Joining Decision
	(1) Sacrifices E.	(2) Moral Join	(3) Index	(4) Concern	(5) Affective	(6) Cognitive	(7) Join
A.1 Joiners and Household Members							
Prosociality Community	0.004	0.006***	-0.014	-0.001	-0.091	-0.409**	0.011***
	(0.003)	(0.002)	(0.159)	(0.129)	(0.148)	(0.181)	(0.003)
Sacrifices Enemy			34.472***	-24.753***	-17.744***	-8.633	0.059
			(5.135)	(4.378)	(5.169)	(5.960)	(0.103)
Moral to join			13.446***	-10.942***	-12.875***	-16.812***	0.581***
			(5.050)	(3.888)	(4.440)	(4.700)	(0.068)
Sadism Enemy							0.001
							(0.002)
Empathic Concern Enemy							0.000
							(0.003)
Affective Empathy Enemy							0.002
							(0.002)
Cognitive Empathy Enemy							-0.004***
							(0.001)
Joiner ID FE	✓	✓	✓	✓	✓	✓	✓
Mean dep. var	0.60	0.65	66.37	33.23	39.29	59.79	0.50
R^2	0.52	0.37	0.60	0.58	0.55	0.50	0.38
Observations	412	424	412	412	412	412	412
A.2 Joiners and Random Villagers							
Prosociality Community	-0.003	-0.004	-0.070	0.066	0.182	-0.089	0.003
	(0.004)	(0.003)	(0.229)	(0.231)	(0.308)	(0.303)	(0.005)
Sacrifices Enemy			17.597**	-9.439	-10.034	1.892	0.076
			(6.804)	(6.667)	(6.825)	(7.525)	(0.135)
Moral to join			18.634**	-20.824***	-19.265**	-16.385*	0.533***
			(8.135)	(7.879)	(8.295)	(9.414)	(0.129)
Sadism Enemy							-0.001
							(0.003)
Empathic Concern Enemy							-0.001
							(0.004)
Affective Empathy Enemy							0.001
							(0.004)
Cognitive Empathy Enemy							-0.002
							(0.002)
Joiner ID FE	✓	✓	✓	✓	✓	✓	✓
Mean dep. var	0.64	0.75	69.72	31.79	37.42	54.41	0.50
R^2	0.48	0.46	0.54	0.48	0.47	0.47	0.13
Observations	416	424	416	414	414	414	414

Notes. As in Table 2, all regressions include joiner ID fixed effects to exclude comparisons across individuals that might capture self-selection into the sample induced by self-selection of joiners and therefore of their own comparison samples.

inclination can generate psychological discomfort because it stands in contradiction moral values of most humans. To appease such psychological pain, individuals can choose to believe (hence in a motivated fashion) that it is actually morally right to harm the Hutu, thus appeasing the psychological pain while allowing to sustain such inclination.

We formally analyze whether moral disengagement/motivated reasoning can explain our correlation and the patterns of selection in Table 6. In Columns (1)-(4), we regress the measures of disposition for inter-personal harm against the Hutu on the two dummies for moral values,

Table 6: Mechanisms II: Moral Convictions or Motivated Reasoning?

	Sadism E.	Empathy to Enemy			Joining Decision	
	(1) Index	(2) Concern	(3) Affective	(4) Cognitive	(5) Joins	(6) Joins
B.1 Joiners and Household Members						
Attacked by Hutu	0.584	5.309	9.730*	1.581	-0.191**	-0.177*
	(5.034)	(4.030)	(5.205)	(5.705)	(0.086)	(0.090)
Anger x Anger about war	8.880	-9.417**	-6.448	-4.617	-0.111	-0.111
	(5.825)	(4.600)	(3.943)	(4.840)	(0.120)	(0.117)
Anger intensity	-0.704	1.595	2.365	0.127	-0.012	-0.029
	(1.971)	(1.579)	(1.726)	(2.089)	(0.040)	(0.038)
Anger about war	-79.097	79.429**	46.507	41.619	0.977	0.972
	(49.796)	(39.698)	(35.284)	(40.329)	(1.036)	(1.014)
Bandura	52.327***	-56.164***	-57.882***	-17.934	0.306	0.208
	(13.895)	(11.042)	(13.240)	(15.781)	(0.225)	(0.215)
Sacrifices Enemy	26.168***	-17.294***	-13.094**	-10.921	0.211*	0.192*
	(6.717)	(5.110)	(6.598)	(8.075)	(0.112)	(0.108)
Moral to join	16.808***	-14.287***	-18.069***	-27.807***	0.616***	0.584***
	(5.606)	(4.396)	(5.037)	(5.692)	(0.088)	(0.084)
Prosociality Community						0.012***
						(0.003)
Joiner ID FE	✓	✓	✓	✓	✓	✓
Mean dep. var	66.37	33.23	39.29	59.79	0.50	0.50
R^2	0.66	0.68	0.66	0.54	0.30	0.36
Observations	294	294	294	294	294	294
B.2 Joiners and Random Villagers						
Attacked by Hutu	-1.679	4.583	7.951	0.100	-0.226	-0.230
	(9.380)	(8.504)	(9.762)	(10.850)	(0.160)	(0.160)
Anger x Anger about war	15.692**	-16.219***	-16.225***	-10.482	-0.200	-0.194
	(6.438)	(5.050)	(5.894)	(7.178)	(0.158)	(0.158)
Anger intensity	0.312	1.121	1.020	0.691	0.028	0.027
	(2.601)	(2.159)	(2.464)	(2.985)	(0.042)	(0.042)
Anger about war	-130.020**	138.878***	131.841**	94.715	1.628	1.581
	(58.027)	(44.810)	(51.912)	(62.943)	(1.473)	(1.470)
Bandura	48.809**	-51.968***	-59.536***	-39.672*	0.757**	0.757**
	(20.191)	(18.193)	(18.402)	(21.779)	(0.296)	(0.293)
Sacrifices Enemy	11.611	-5.136	-5.720	2.455	-0.024	-0.017
	(8.082)	(7.983)	(8.142)	(9.755)	(0.172)	(0.170)
Moral to join	10.391	-15.104*	-14.429*	-11.894	0.502***	0.515***
	(8.587)	(7.686)	(7.450)	(10.447)	(0.147)	(0.147)
Prosociality Community						0.003
						(0.006)
Joiner ID FE	✓	✓	✓	✓	✓	✓
Mean dep. var	69.72	31.79	37.42	54.41	0.50	0.50
R^2	0.62	0.58	0.58	0.50	0.24	0.25
Observations	308	308	308	308	308	308

but we also include as controls variables susceptible of inducing an emotional *motivation* to justify the belief that it is morally right to be violent against the Hutus (or to join). Our objective is to assess whether moral values influence disposition to inter-personal harm (other than cognitive empathy) despite the controls. First, we include a dummy for whether family members of the respondent were victimized by the Hutu in the past. We had shown in Table A4 that there is a high mean level of attacks with more than half the sample not having been attacked, hence the variable has significant variation. Second, we include a measure of anger intensity, multiplied with whether such emotion was associated to the war.[36] One concern with this control is that maybe this captures that angry people more generally or that anyone talking about the war have less empathy to the enemy. Our last control is the Bandura (2011)'s moral disengagement index.

The coefficients in Columns (1)-(4) indicate that motivated reasoning cannot explain the relationship between individually-held moral values and disposition for inter-personal harm of Hutus. Indeed, despite inclusion of all these controls, the coefficients on moral values remain strong, of the same sign as in Table 5, and statistically significant (except, as before, for cognitive empathy to the enemy). Interestingly, of all the controls, only the index of moral disengagement is statistically significant and starkly associated with lower aversion to inter-personal harm of the Hutu, except, again, for cognitive empathy.

Columns (5)-(8) examine the same relationship but using a dummy for whether the person joins as dependent variable. Column (5) shows that moral values remain associated to whether the person joins even controlling for motivated reasoning. Furthermore, while the rest of controls are insignificant, having family members attacked by the Hutus is negatively associated to joining, rather than positively. This underscores that joiners are driven by moral motivations that connect them to their community as a whole rather than parochial motivations about avenging their family. Column (6) includes back prosociality to the community, and the results are unaffected, while prosociality is significant at the 1% level. This means that the selection of prosocial individuals is not explained by them having more motivations to harm the Hutus that would not stem of moral values.

In sum, enabled by moral disengagement, most civilians deem it morally right to join the NDC-R and to kill Rwandan Hutus, with some holding "genocidarian" values, because of the

[36] We measured anger intensity in emotions module that we administered following Scherer, Wranik, Sangsue, Tran, and Scherer (2004). To determine whether the emotion was associated to the war, we use a training sample of 50 observations to classify the open-ended text responses describing the emotion into war and predict whether the text is about the war. We obtain 98% accuracy in the training sample.

violence and oppression exerted by Rwandan Hutus did to the Walikale population. This has three implications. First, joining the NDC-R is rightly conceptualized as a prosocial act, from the perspective of the relevant society, Walikale. Second, these moral convictions that deem it right to kill Rwandan Hutus could potentially shut down empathy towards them, consistent with past research (Workman, Yoder, and Decety, 2020). Three, the key correlation we have documented and the selection into the group are governed by prosociality to the community and cognitive empathy to the enemy, and not by other notions of disposition for violence, heterogeneous moral values, or motivated reasoning.

8. Conclusion

In order to study the role of altruism in violence perpetration, we unbundled altruism into prosociality and empathy as well as across targets and studied those in a context in which a Congolese established armed group has been receiving volunteers for 15 years to provide protection from, and fight an enemy. Two unbundled distinct notions of altruism are crucial to the study of violence: prososiality towards the community and aversion to inter-personal physical harm to the enemy. If the altruism components matter in the decision to join violence, then when these components are weakly positively or even negatively correlated correlation, those who engage in violence against the enemy would have to be the most prosocial (towards the community) and the least empathetic towards the enemy. We found that aversion to inter-personal physical harm is distinct from prosociality and that it is starkly predictive both of actual violent behavior, as well as with emotional arousal. Then, we found that, consistent with this prediction, those who join the militia fighting the enemy are more prosocial to their community and less empathetic to the enemy, even before participating in the militia. Those who join the militia were also shown to hold prosocial motivations associated to the decision to join motivated by the *consequence* of joining for society. Potentially governing this negative association between prosociality to the community and aversion to inter-personal harm agains the enemy, we provided evidence of a "morality of violence" in a context following decades of abuses by a perpetrator, in which civilians hold moral convictions that it is morally right to join a militia against the perpetrator, as well as to kill members of the perpetrator group.

We were able to establish these results by undertaking three key steps. First, we had to unbundle altruism into new notions that are relevant for violence and that are not necessarily the

same as prosociality; to do so, we borrowed from the field of moral psychology and introduced these notions in a simple framework drawing from existing frameworks in labor economics. Second, we had to measure these notions, for which we also borrowed from the field of moral psychology as well as physiopsychology to validate some of the measures, especially those pertaining to disposition for inter-personal harm; we also had to adapt existing moral dilemmas to a social context of war where classical moral presuppositions do not apply. Third, and most importantly logistically, we nurtured during 9 years the trust of the largest and oldest armed group in the Congo. This allowed us to develop a communication system with their 55 human resource commanders to obtain in real time the information of who joins, where they are, and where is their village of origin. This step was crucial in circumventing a fundamental challenge that has made it so far impossible to answer this question: interviewing those who join (and those who do not) *before* participation into violence changes them and before survivor biases arising from violence participation kick in; the related literature so far was forced to rely on interviewing ex-combatants in de-mobilization centers, limiting progress on this question.

The framework we developed showed that, depending on the underlying correlations between the components of altruism and their variances, different cases of selection could emerge if these components mattered in the decision to join. For example, in some cases, those who join violence might be the most empathetic and prosocial towards anyone, as in the case of peaceful social movements, whereas in others, it will be those who have neither empathy for anyone nor any prosocial motivation, as in the case of mercenary armies. Our finding does not in any way provide evidence against any type of selection, as those are going to be dependent on the correlations in the context, and, as we argued, on the underlying morality that sustains such correlations. Instead, our study provides a possibility result, a proof of concept showing that altruism can be a relevant tool of analysis when studying the causes of violence perpetration and that, in some cases such as ours, unbundling altruism is necessary to reveal its role. This possibility result that prosociality towards the community might be associated to lower empathy to an enemy is in turn consistent with a large literature in anthropology, which has underscored how the moral duty to help the community attack others for revenge, honor, or raiding is an important motivator of individuals' participation in violence. Our results also do not contain any normative analysis regarding the righteousness of such beliefs, preferences, or behavior. This unbundling can be a useful tool to study violence across contexts, in particular conflict contexts with de-humanizing ethno-supremacist violence with desire to oppress an another group.

Overall, our study sheds humbling light on human nature, eliciting an intimate connection between prosocial motivations and a inclination for inter-personal harm of others. This connection is supported by a prevailing morality held by civilians, rather than by psychopathy of those who participate in violence, reminding us that "*the line separating good and evil passes . . . right through every human heart—This line shifts. Inside us, it oscillates with the years. And even within hearts overwhelmed by evil, one small bridgehead of good is retained*" (Solzhenitsyn, 2011). Regardless of the power dynamic and of the normative merit of the support for violence, our results lend support to psychologist Gabor Maté observation that: "*We each have a Nazi within," the Auschwitz survivor Edith Eger has written – pointing, in my observation, to a near-universal reality. Many of us harbor the seeds for hatred, rage, fear, narcissistic self-regard and contempt for others*" (Maté, 2024).[37] Our findings suggest that, in some circumstances, we are capable of switching off empathy towards some people, and even enjoy killing them, while at the same time holding empathy towards others. Our findings' unintended take-away is therefore to break down notions of generalized altruism that presuppose that there are good and bad people: it depends towards whom and on what dimension. Furthermore, our study closed by demonstrating that our disposition for inter-personal harm is malleable suggests "we" as an object of analysis might be just as much "situations" as we are "persons:" our potential for violence can be driven by the circumstances and social context.

Our steps, motivated by a decade of qualitative work, were intended to interrogate and to characterize the role of altruism in violence participation. The role of prosocial motivations in violence participation is well-known to liberatory violence but it is also necessary in the selfless disposition in taking part ethno-supremacist violent movements and oppression, where the individual takes a personal cost to benefit a cause beyond the individual, often when the dominance of a group is perceived to be under threat. However, the role of inclination to harm the enemy appears to be more characteristics of conflict situations of retaliatory violence and especially of etho-supremacist violence, in which contempt for the humanity of others is crucial in suppressing compassion and forms the seeds of genocidarian violence, intentionally geared towards their elimination (in part or in whole) as a community, society. This complements the current accepted knowledge in economics that economic forces might be sufficient to explain the phenomenon

[37]The original quote is from Eger (2020): "We're not all descendants of Nazis. But we each have a Nazi within. Freedom means choosing, every moment, whether we reach for our inner Nazi or our inner Gandhi. For the love we were born with or the hate we learned. The inner Nazi is the part of you that has the capacity to judge and withhold compassion, that denies you the permission to be free and victimizes others when things don't go your way. I'm still learning to let go of my inner Nazi."

of violence. But, inadvertently, first through qualitative work and then quantitatively, our study also found evidence for a morality of retaliatory violence in which it is morally right to kill the enemy—even if it comes at a cost to the community, if that has no effect on security, and especially if the enemy belongs to a population sub-group that is often targeted in in genocidal violence. Thus, inadvertently, our findings also open the door to questions about our shared potential for genocide. They suggest such potential resides more in our moral and prosocial common core than in any presumed (and rare) psychopathic tendency.

We hope that this study will inspire more research into our potential for violence and for genocide and contribute to deconstruct simplistic and assumed notions that classify humans into hierarchies of goodness, lending themselves to racialized narratives—such as the prevailing notion that non "Western" peoples fight because of greed, of minerals, or animated by some sort of morally-empty aims, often based on the de-humanizing conscious or unconscious presupposition that there exists a "they" and an "us" and that "they" lack a capacity for morality and empathy. Just like in the Spanish Civil War, in the Second World War, in liberatory struggles, as well as in many conflicts today, there is a moral dimension to perpetrating violence that has hitherto not been possible to empirically analyze in economics—and that exists regardless of whether the reader as a subject shares the moral convictions of the participants.

References

Actualite.cd (2023): "Loi instituant la Réserve Armée de la Défense en RDC: pour Claudel Lubaya, ce texte va à l'encontre de la résolution de l'Assemblée nationale sur les groupes armés et légalise les milices au sein des FARDC," .

Arendt, H., and A. Elon (2006): *Eichmann in Jerusalem: A Report on the Banality of Evil*, Penguin Classics. Penguin Publishing Group.

Avenanti, A., A. Sirigu, and S. M. Aglioti (2010): "Racial bias reduces empathic sensorimotor resonance with other-race pain," *Current Biology*, 20(11), 1018–1022.

Bandura, A. (2011): *Moral Disengagement*John Wiley and Sons, Ltd.

Bénabou, R., A. Falk, and L. Henkel (2024): "Ends Versus Means: Kantians, Utilitarians, and Moral Decisions," *SSRN Electronic Journal*.

Blattman, C., and E. Miguel (2010): "Civil War," *Journal of Economic Literature*, 48(1), 3–57.

Bowles, S., and H. Gintis (2011): *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton University Press.

Browning, C. (2017): *Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland*. HarperCollins.

Bénabou, R. (2015): "The Economics of Motivated Beliefs," *Revue d'économie politique*, 125(5), 665–685.

Chen, Y.-C., C.-C. Chen, J. Decety, and Y. Cheng (2013): "Aging is associated with changes in the neural circuits underlying empathy," *Neurobiology of aging*, 35.

Congo Research Group (2020): "For the Army, With the Army, Like the Army?," Discussion paper, Congo Research Group.

De Zavala, A. (2023): *The Psychology of Collective Narcissism: Insights from Social Identity Theory*. Taylor & Francis.

Decety, J. (2011): *Empathy: From Bench to Bedside*. The MIT Press.

Decety, J., and P. Jackson (2004): "The Functional Architecture of Human Empathy," *Behavioral and cognitive neuroscience reviews*, 3, 71–100.

Dunia Butinda, L. (2021): *D'Enfant de Walikale a la Prise de l'Arme: Un Voyage Aller-Retour*. Unpublished Autobiography.

Eger, E. (2020): *The Gift: 14 Lessons to Save Your Life*. Scribner.

Enke, B. (2019): "Kinship, Cooperation, and the Evolution of Moral Systems," *The Quarterly Journal of Economics*, 134(2), 953–1019.

Enke, B., R. Rodríguez-Padilla, and F. Zimmermann (2022): "Moral Universalism: Measurement and Economic Relevance," *Management Science*, 68(5), 3590–3603.

Eres, R., J. Decety, W. R. Louis, and P. Molenberghs (2015): "Individual differences in local gray matter density are associated with differences in affective and cognitive empathy," *NeuroImage*, 117, 305–310.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018): "Global Evidence on Economic Preferences," *The Quarterly Journal of Economics*, 133(4), 1645–1692.

Falk, A., A. Becker, T. Dohmen, D. Huffman, and U. Sunde (2022): "The Preference Survey Module: A Validated Instrument for Measuring Risk, Time, and Social Preferences," *Management Science*, 69.

Fearon, J. D. (1995): "Rationalist Explanations for War," *International Organization*, 49(3), 379–414.

Ferguson, A., and M. Inzlicht (2022): "Reliability of the empathy selection task, a novel behavioral measure of empathy avoidance," *Behavior Research Methods*, 55.

Fouka, V., and M. Tabellini (2022): "Changing In-Group Boundaries: The Effect of Immigration on Race Relations in the United States," *American Political Science Review*, 116(3), 968–984.

Fowles, D. C., M. J. Christie, R. Edelberg, W. W. Grings, D. T. Lykken, and P. H. Venables (1981): "Publication Recommendations for Electrodermal Measurements," *Psychophysiology*, 18(3), 232–239.

Gluckman, M. (2009): *31. The Peace in the Feud*pp. 306–315. Stanford University Press, Redwood City.

Graça Machel (1996): "Impact of Armed Conflict on Children: Report of the expert of the Secretary-General, Ms. Graça Machel, submitted pursuant to General Assembly resolution 48/157," Discussion paper, Human Rights Watch.

Greene, J. (2013): *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. Penguin Publishing Group.

Gómez, J., M. Verdú, A. González Megías, and M. Méndez (2016): "The phylogenetic roots of human lethal violence," *Nature*, 538.

Habyarimana, J., M. Humphreys, D. Posner, and J. Weinstein (2007): "Why Does Ethnic Diversity Undermine Public Goods Provision?," *American Political Science Review*, pp. 709–725.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, R. McElreath, M. Alvard, A. Barr, J. Ensminger, N. S. Henrich, K. Hill, F. Gil-White, M. Gurven, F. W. Marlowe, J. Q. Patton, and D. Tracer (2005): ""Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies," *Behavioral and Brain Sciences*, 28(6), 795–855.

Humphreys, M., and J. M. Weinstein (2008): "Who Fights? The Determinants of Participation in Civil War," *American Journal of Political Science*, 52(2), 436–455.

Kelly, R. (1985): *The Nuer Conquest: The Structure and Development of an Expansionist System*. University of Michigan Press.

KST (2021): "Kivu Security Tracker," Discussion paper.

Lagrange, H. (2016): "Diego Gambetta and Steffen Hertog: Engineers of Jihad: The Curious Connection between Violent Extremism and Education," *European Sociological Review*, 33(1), 161–164.

Maté, G. (2024): ""We each have a Nazi in us." We need to understand the psychological roots of authoritarianism," *The Guardian*.

Mcgovern, M. (2011): "Popular Development Economics—An Anthropologist among the Mandarins," *Perspectives on Politics*, 9.

Mugaruka, M., G. Marchais, R. Sanchez de la Sierra, and Q. Wu (2024): "Social Origins of Militias: The Extraordinary Rise of the Outraged Citizens," *Working paper*.

Ndahayo, E., and A. Dufatanye (2015): *La violence politico-militaire contre les femmes au Rwanda – De Ndabaga a Ingabire*. Editions Sources du Nil ed. Lille: Collection Pyramides.

Nest, M. (2011): *Coltan*. Polity Press, Cambridge, UK.

Nisbett, R., and D. Cohen (1994): *Feuding and Warfare: Selected Works of Keith F. Otterbein*, New Directions in Social Psychology: Self, Cognition & Collectives. Langhorne, PA, USA: Gordon and Breach.

——— (1996): *Culture Of Honor: The Psychology Of Violence In The South*, New Directions in Social Psychology: Self, Cognition & Collectives. Avalon Publishing.

Nordstrom, C. (1997): *A Different Kind of War Story*, The Ethnography of Political Violence. University of Pennsylvania Press, Incorporated.

Pinker, S. (2011): *The Better Angels of Our Nature: Why Violence Has Decline*. Viking Books, New York.

Radio France International (2023): "RDC: l'adoption d'un projet de loi sur un corps de réservistes au sein de l'armée se précise," .

Scherer, K. R., T. Wranik, J. Sangsue, V. Tran, and U. Scherer (2004): "Emotions in everyday life: Probability of occurrence, risk factors, appraisal and reaction patterns," *Social Science Information*, 43(4), 499–570.

Solzhenitsyn, A. (2011): *The Gulag Archipelago*. Random House.

Thomson, J. J. (1984): "The trolley problem," *Yale LJ*, 94, 1395.

Vogel, C., G. Salvaggio, P. Boisselet, and J. K. Stearns (2021): "The Landscape of Armed Groups in Eastern Congo: Missed Opportunities, P-rotracted Insecurity and Self-Fulfilling prophecies," Discussion paper, Kivu Security Tracker - Congo Research Group - NYU Center on International Cooperation, New York.

Workman, C., K. Yoder, and J. Decety (2020): "The Dark Side of Morality - Neural Mechanisms Underpinning Moral Convictions and Support for Violence.," *AJOB Neurosci*, (Oct-Dec;11(4):269-284).

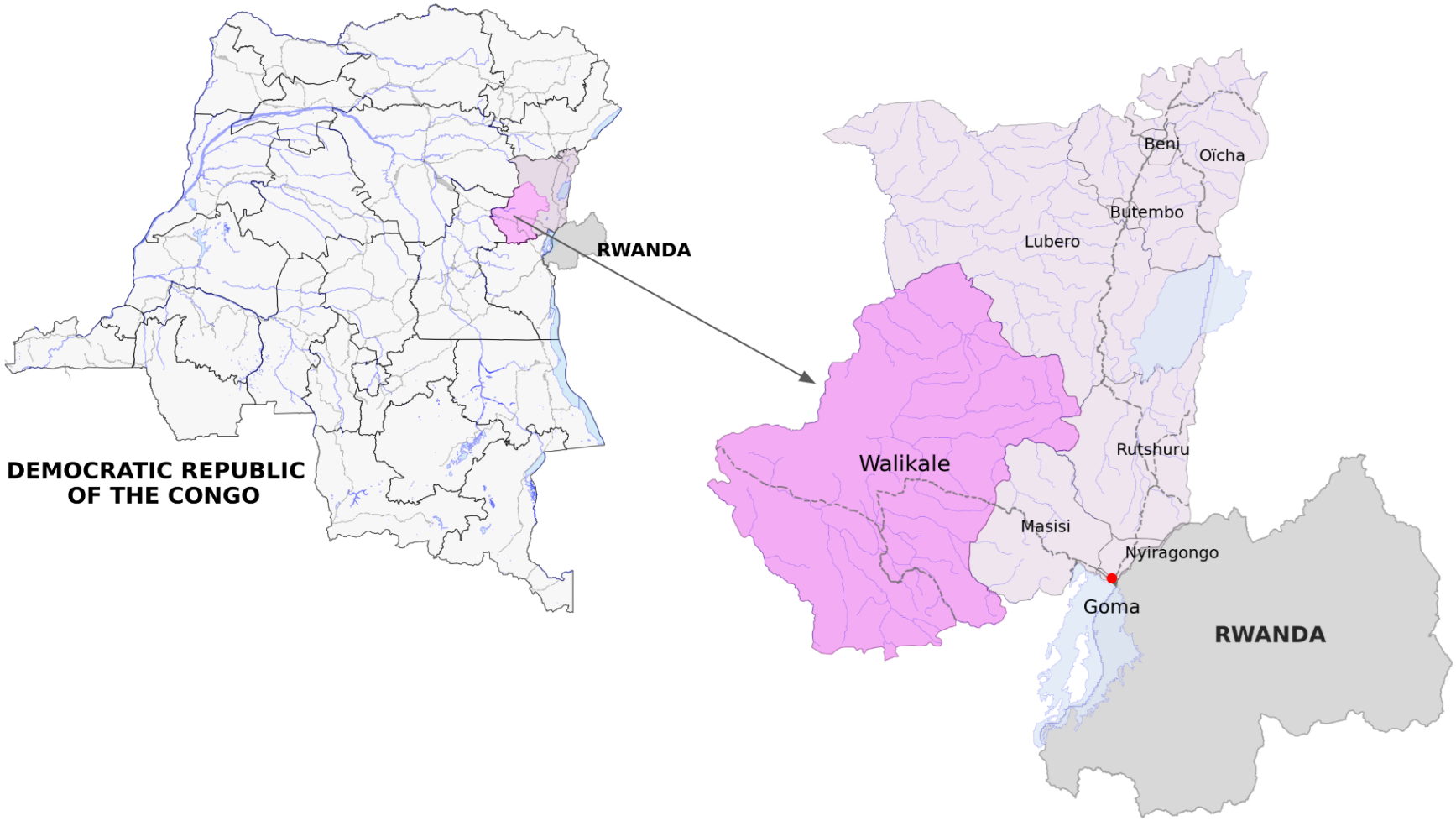
Yanagizawa-Drott, D. (2014): " Propaganda and Conflict: Evidence from the Rwandan Genocide," *The Quarterly Journal of Economics*, 129(4), 1947–1994.

Zaki, J., and K. N. Ochsner (2012): "The neuroscience of empathy: progress, pitfalls and promise," *Nature neuroscience*, 15(5), 675—680.

Zimbardo, P. (2008): *The Lucifer Effect: How Good People Turn Evil*. Rider.

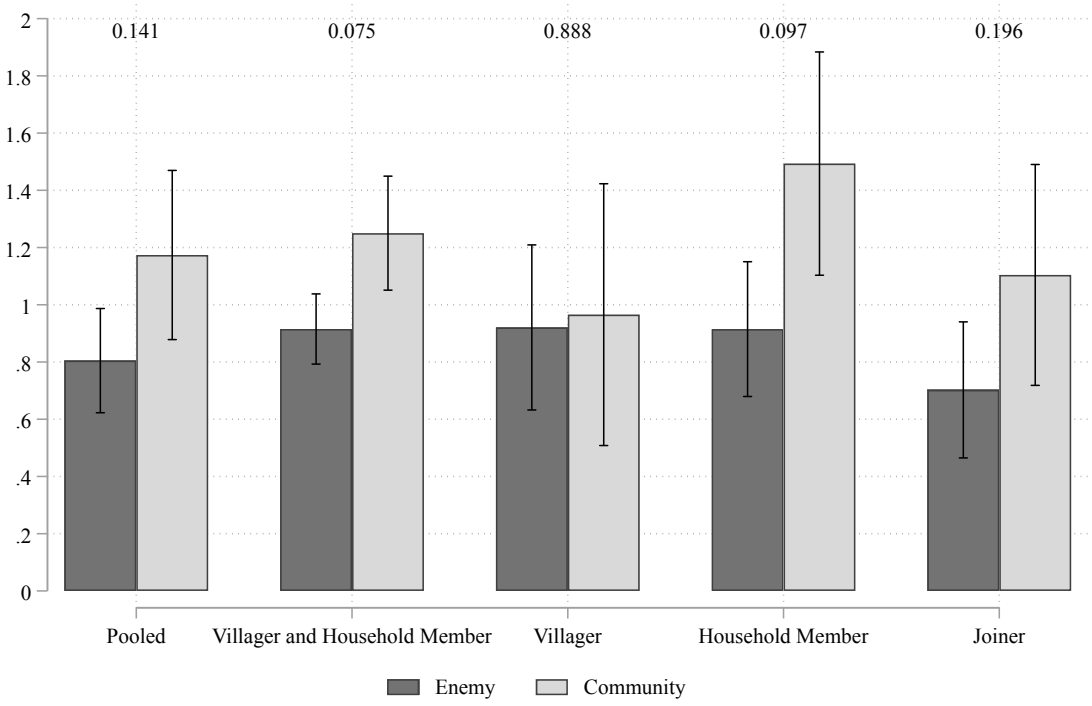
Appendix A. Appendix Tables and Figures

Figure A1: Map of the District of Walikale



Notes. This map presents the district of Walikale in its geographical and administrative context. Source: Authors using GIS software and public shapefiles.

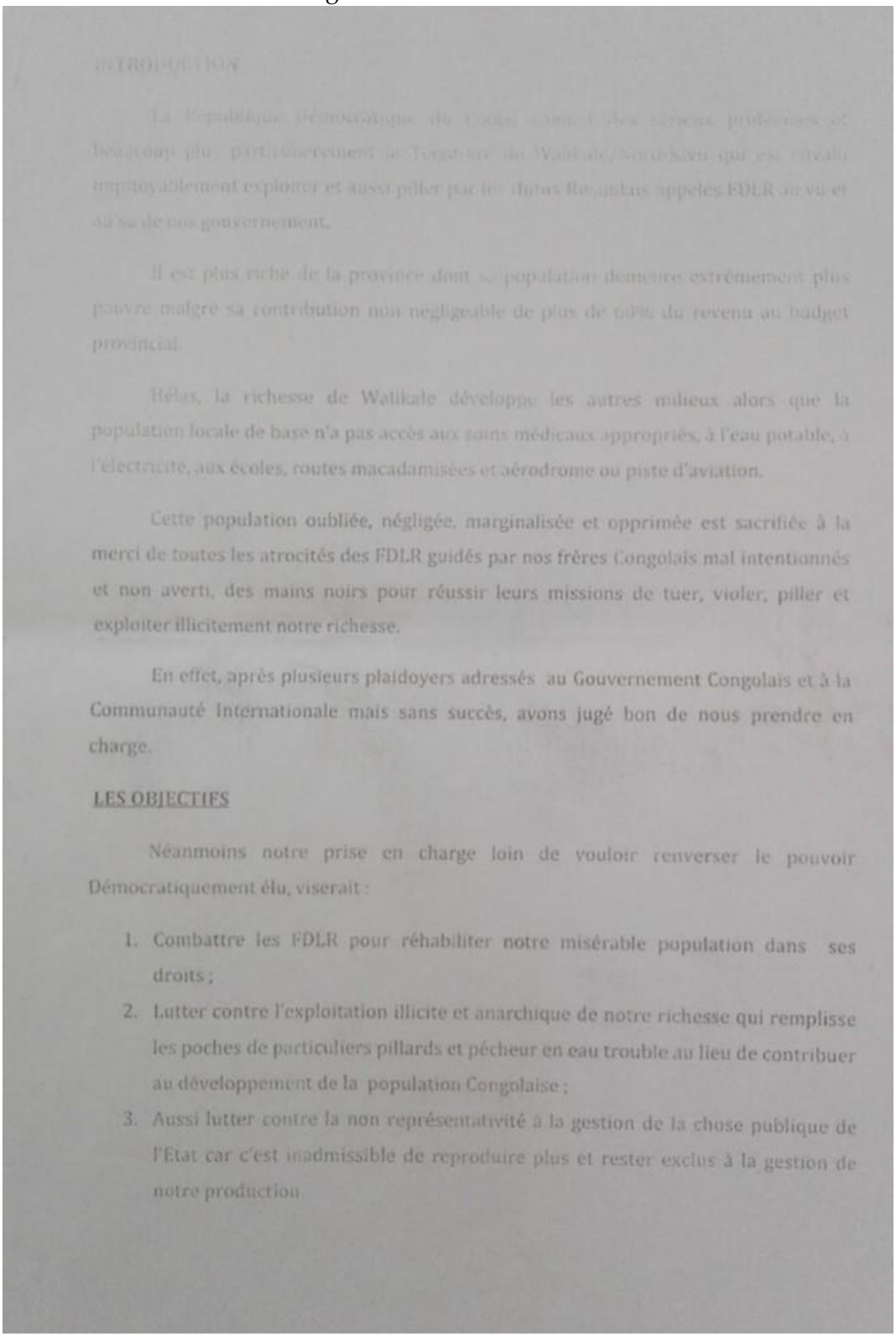
Table A1: Additional Validation of the Measurement of Aversion to Inter-Personal Harm: Identities



Notes. This figure presents the mean measured emotional arousal intensity for different respondent samples, whenever they observe the suffering of a Rwandan Hutu, of a Walikale civilian, and their difference. Target observations indicates the size of the corresponding sample used in the analysis of the corresponding column, ignoring whether individual variables have idiosyncratic missing data (which varies by column).

Figure A2: Nduma Defense of Congo-Rénové Political Demands

Panel A. "Cahier des Charges:" Political Demands at the NDC-R's Creation



INTRODUCTION

La République Démocratique du Congo connaît des revers prolongés et beaucoup plus particulièrement au Territoire de Walikale/Nord-Kivu qui est vivalu impitoyablement exploiter et aussi piller par les chefs Rwandais appelés FDLR au vu et au su de nos gouvernement.

Il est plus riche de la province dont sa population demeure extrêmement plus pauvre malgré sa contribution non négligeable de plus de 60% du revenu au budget provincial.

Hélas, la richesse de Walikale développe les autres milieux alors que la population locale de base n'a pas accès aux soins médicaux appropriés, à l'eau potable, à l'électricité, aux écoles, routes macadamisées et aérodrome ou piste d'aviation.

Cette population oubliée, négligée, marginalisée et opprimée est sacrifiée à la merci de toutes les atrocités des FDLR guidés par nos frères Congolais mal intentionnés et non averti, des mains noirs pour réussir leurs missions de tuer, violer, piller et exploiter illicitement notre richesse.

En effet, après plusieurs plaidoyers adressés au Gouvernement Congolais et à la Communauté Internationale mais sans succès, avons jugé bon de nous prendre en charge.

LES OBJECTIFS

Néanmoins notre prise en charge loin de vouloir renverser le pouvoir Démocratiquement élu, viserait :

1. Combattre les FDLR pour réhabiliter notre misérable population dans ses droits ;
2. Lutter contre l'exploitation illicite et anarchique de notre richesse qui remplisse les poches de particuliers pillards et pêcheur en eau trouble au lieu de contribuer au développement de la population Congolaise ;
3. Aussi lutter contre la non représentativité à la gestion de la chose publique de l'Etat car c'est inadmissible de reproduire plus et rester exclus à la gestion de notre production.

Source: NDC-R.

Panel .B NDC-R's Statement of Objectives in 2016

MOUVEMENT POLITICO MILITAIRE NDUMA DEFENSE OF CONGO RENOVE

NDCR

RAPPORT DES OPERATIONS DU MOUVEMENT NDCR

PRESENTATION DU MOUVEMENT

Apres avoir constaté le comportement anti révolutionnaire de Mr CHEKA NTABO, caractérisé par la megestion des ressources du mouvement, l'ivresses exagérée et attitudes dictatoriale, déviation des objectifs du mouvement, violations des lois des protections des populations civiles par le NDC, le tribalisme et le desordre sexuel indescriptible au point de tenir l'image du mouvement aupres des partenaires internationaux et surtout de la population civile que nous devons proteger, le Haut Commandement du Mouvement NDC en concertation avec tous les officiers superieurs ont decidé de déchoir Mr NTABO CHEKA et ont résolu de créer un nouveau mouvement de rectification dénommé NDUMA DEFENSE OF CONGO RENOVE depuis janvier 2014

OBJECTIFS

1) Combattre et neutraliser les FDLR et leurs alliés jusqu'à leur dernier retranchement ;

2) Sécuriser les communautés de l'EST de la RDC et des pays des Grands Lacs contres l'activisme des groupes armés terroriste à la base des violences dans la région ;

3) Contribuer à la pacification des communautés locale à travers des opérations ciblées contre les ennemis de la paix ;

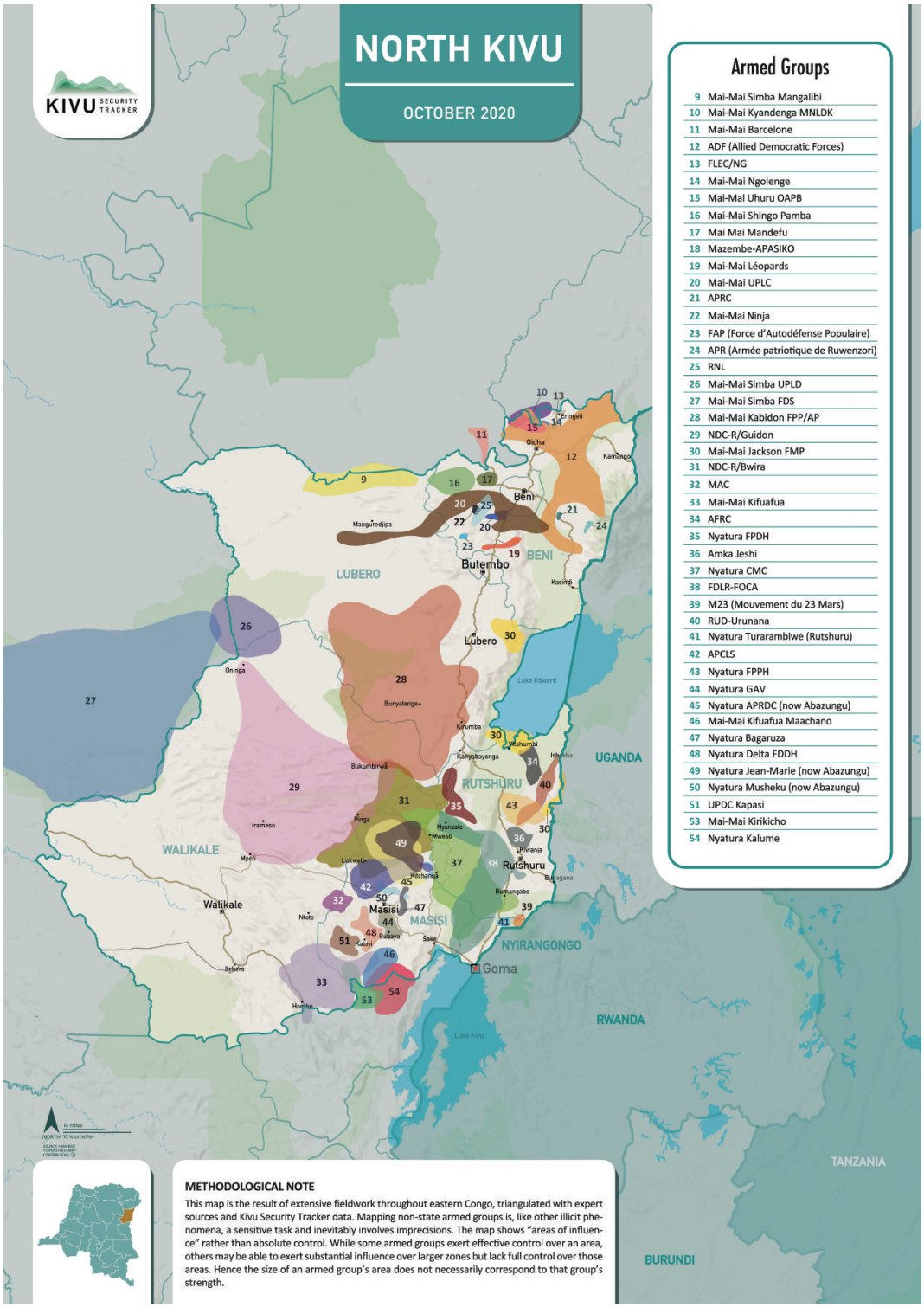
4) Collaborer avec toutes les forces légales susceptible de contribuer la promotion du mouvement au niveau national qu'international ;

5) Appliquer les règles du droit international humanitaire pendant nos opérations et l'appliquer aussi auprès des groupes armés amis ;

6) Préparer des éléments combattants révolutionnaires en réserve pour une armé congolaise forte capable d'assurer la protection de sa population et de leurs biens et de combattre dans le cas échéant toutes les antivaleurs de la mauvaise gouvernance ;

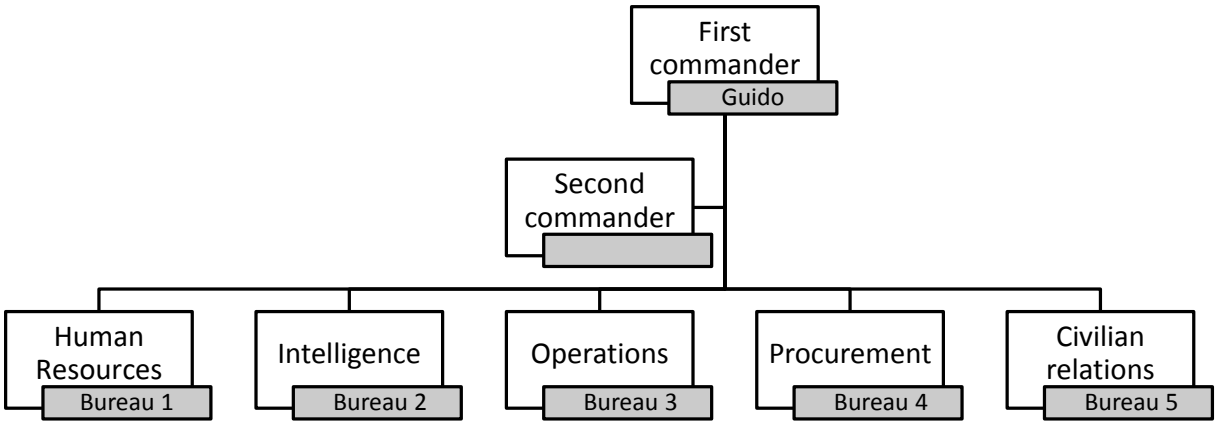
Source: NDC-R.

Figure A3: Nduma Defense of Congo-Rénové Territoral Control Compared to Other Groups



Source: KST (2021). At the time of the study and of writing, NDC-R/Guido and NDC-/Bwira are a unique group.

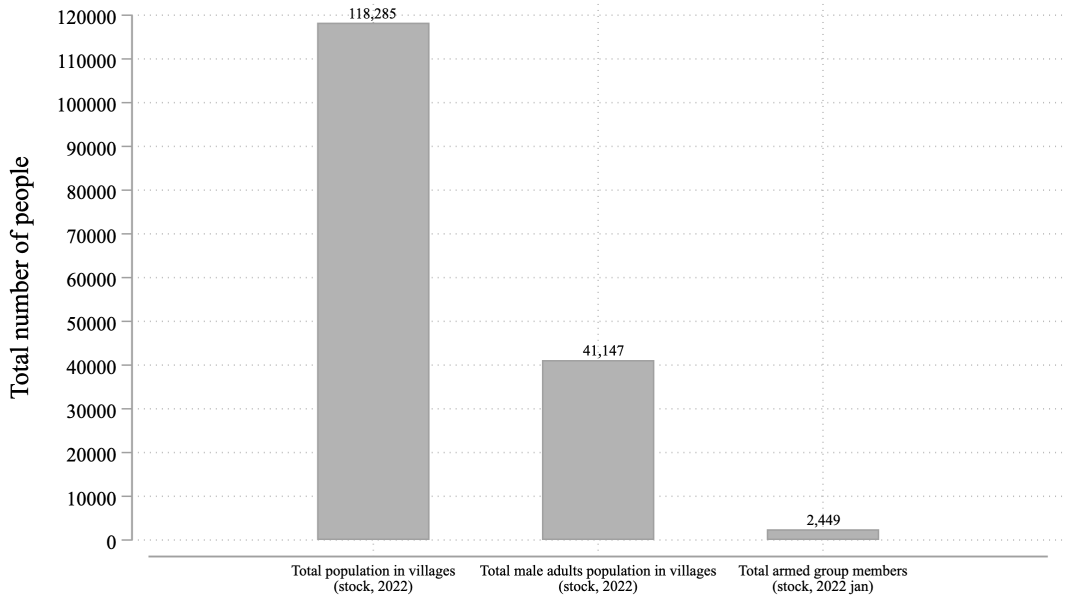
Figure A4: Nduma Defense of Congo-Rénové's Functional Organization



Source: Group's own documents, transcribed using computer by the authors for presentation

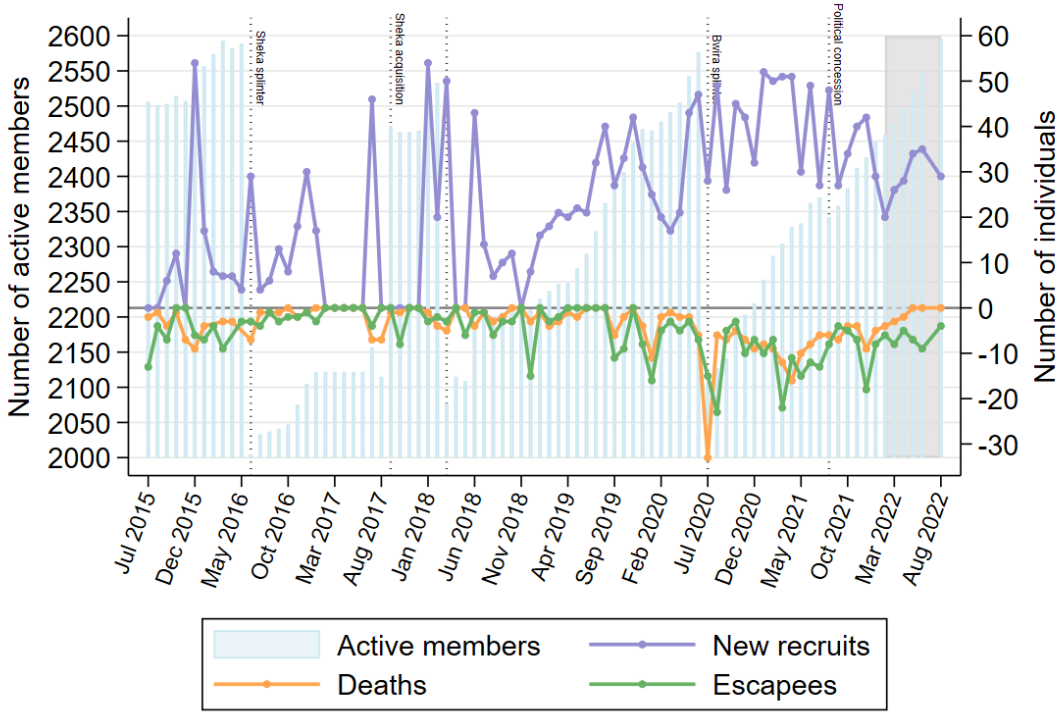
Figure A5: Nduma Defense of Congo-Rénové from its Administrative Records

Panel A. Fraction of the Population in the Catchment Area that is NDC-R member



Fraction who are in militia among male adults: 5%
Fraction who are in militia among total population: 2%

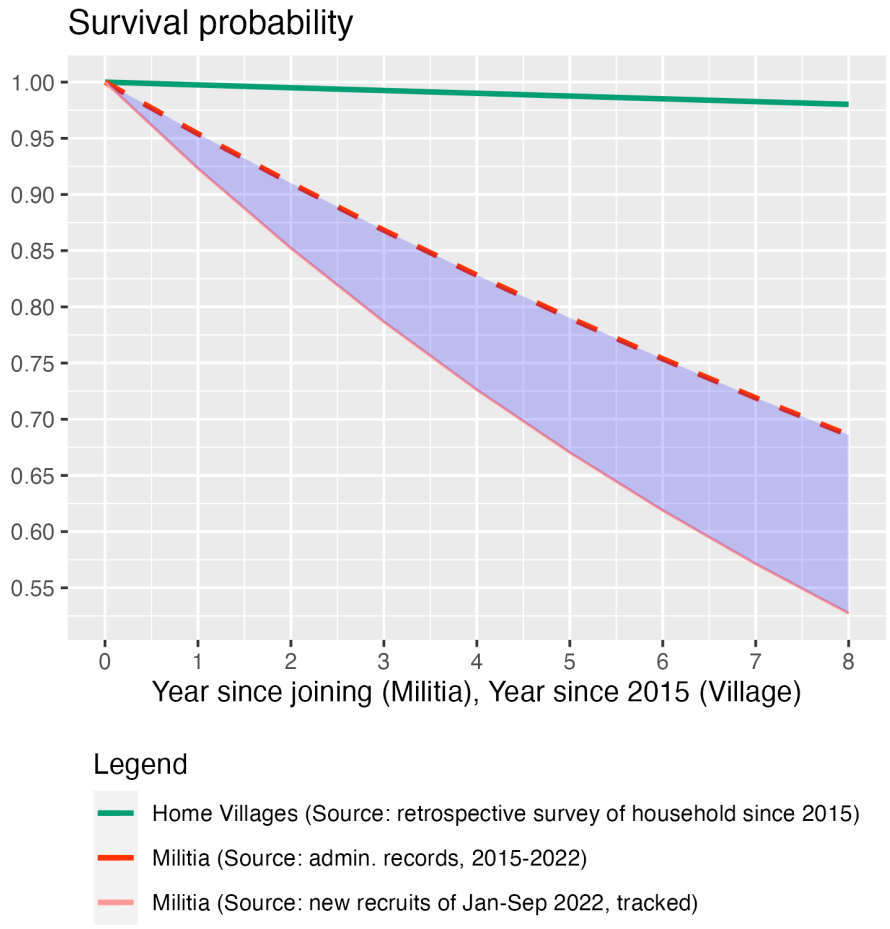
Panel B. Joiners and Stock of Combatants by Month



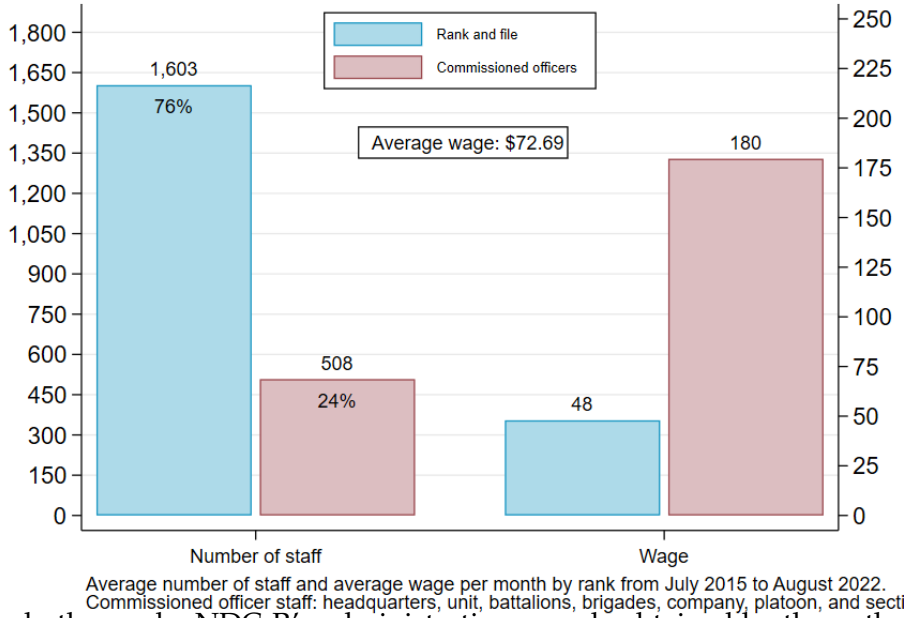
Number of individuals per month from July 2015 to August 2022

Sources. Panel A: NDC-R's administrative records obtained by the authors. Panel B: NDC-R's administrative records obtained by the authors, data from survey obtained in this paper among the civilians in the catchment area of the NDC-R.

Panel C. Survival Probability, Before and After Joining the NDC-R



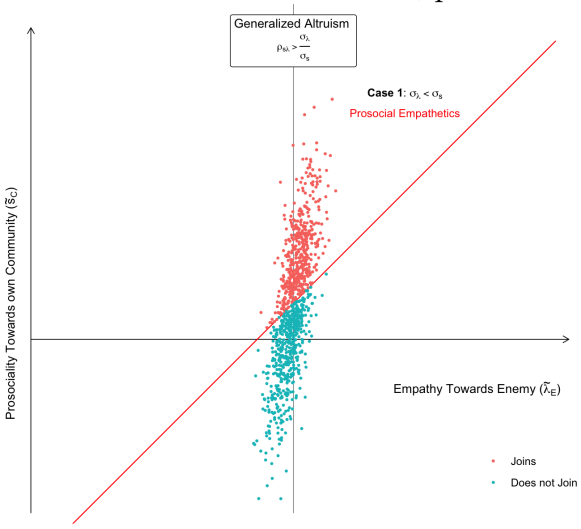
Panel D. Joiners and Stock of Combatants by Month



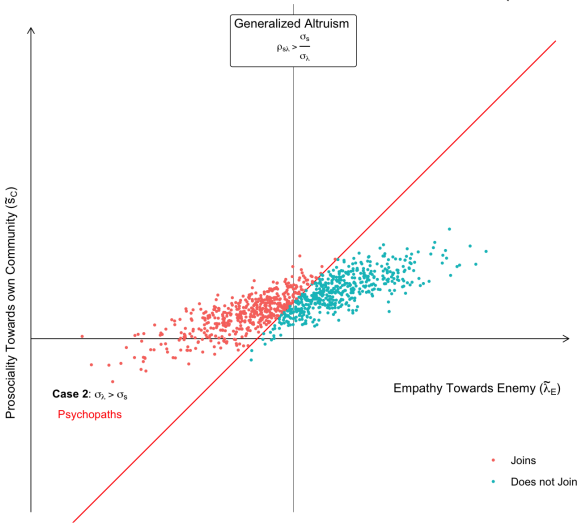
Sources. In both panels: NDC-R's administrative records obtained by the authors, data from survey obtained in this paper among the civilians in the catchment area of the NDC-R.

Figure A6: Organizing Framework: How (Unbundled) Altruism Governs Violence Participation

Panel A. Case I.a: Generalized Altruism, case a ("peaceful social movement")



Panel B. Case I.b: Generalized Altruism, case b ("mercenaries")



Panel C. Case II: Unbundled Altruism ("fanatics")

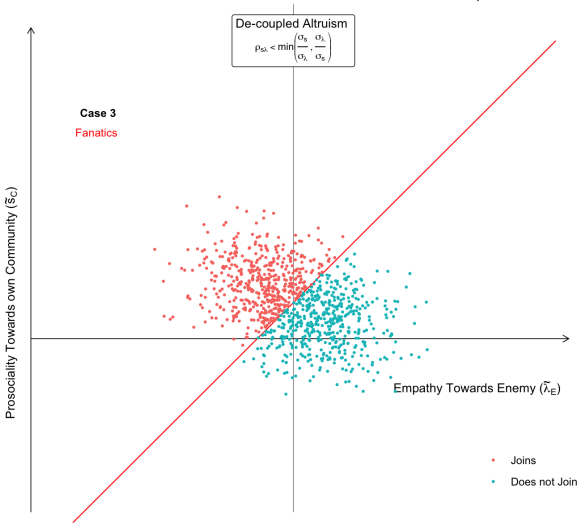
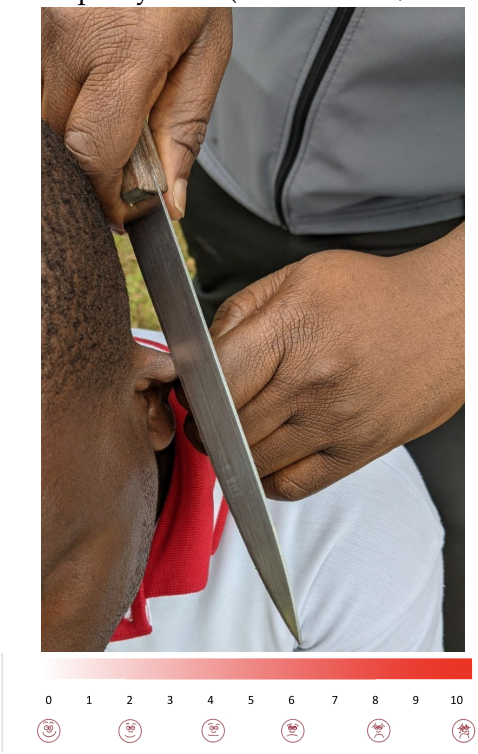
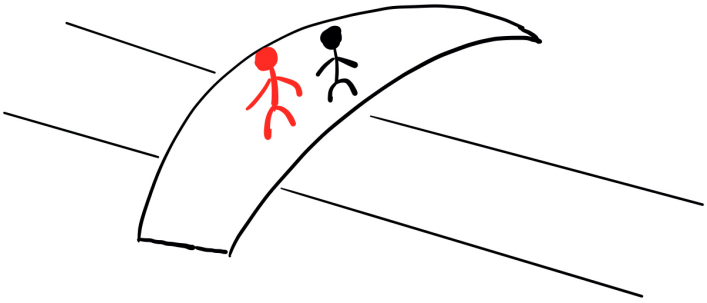


Figure A7: The Survey Tools of Disposition for Inter-Personal Harm and of Morality of Violence

Panel A. The Empathy Task (for Aversion/Taste of Harming)



Panel B. Video Game (Preferences of Harming)



Panel C. Moral Dilemma (Morality of Violence)

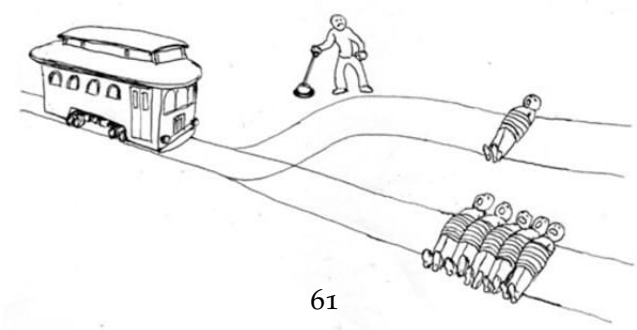
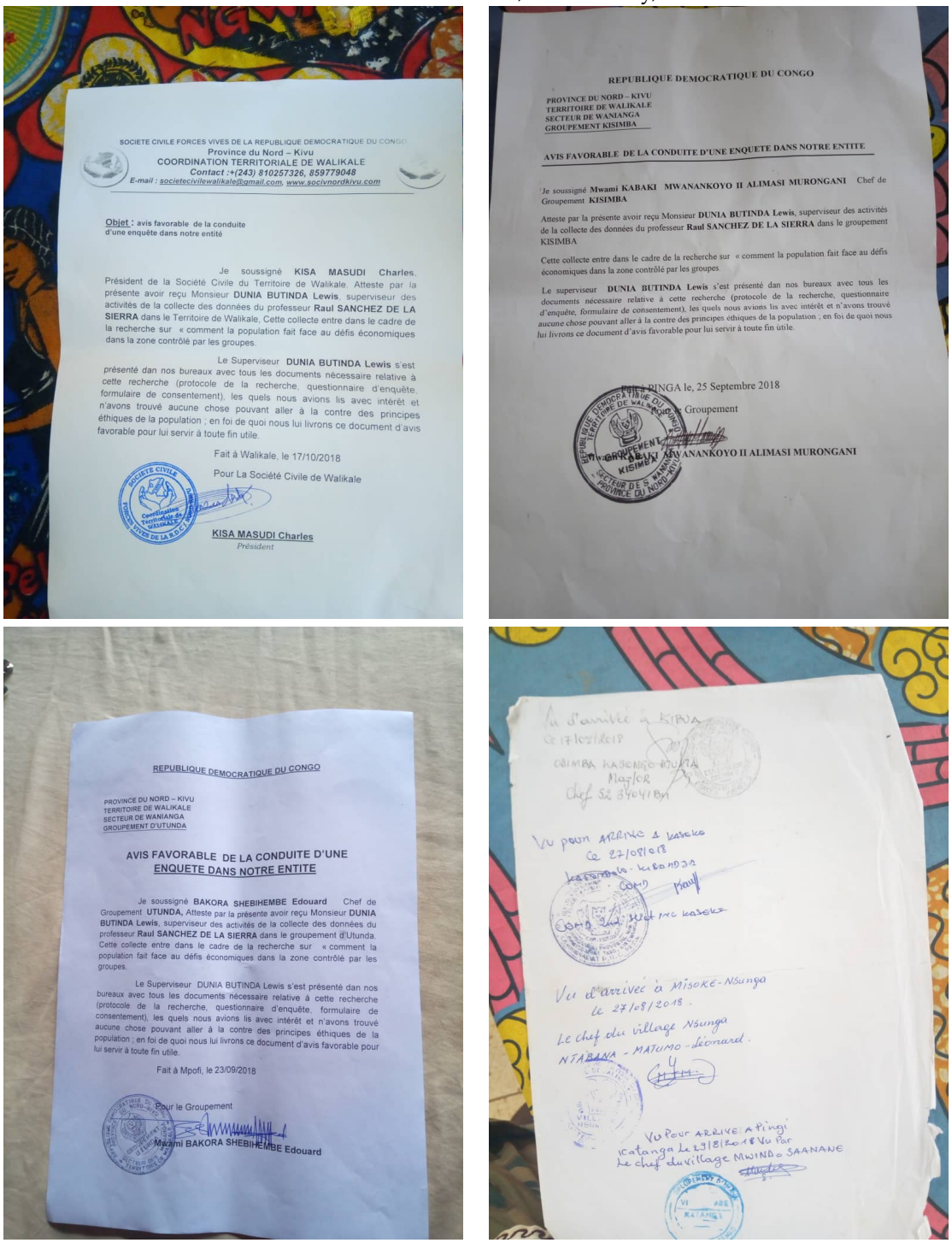


Figure A8: In-Country Authorizations to Conduct Research Among New Volunteers

Panel A. Authorizations from the Civilian, Customary, and State Authorities



SOCIETE CIVILE FORCES VIVES DE LA REPUBLIQUE DEMOCRATIQUE DU CONGO
Province du Nord – Kivu
COORDINATION TERRITORIALE DE WALIKALE
Contact :+(243) 810257326, 859779048
E-mail : societecivilewalikale@gmail.com, www.socivnordkivu.com

Objet : avis favorable de la conduite d'une enquête dans notre entité

Je soussigné KISA MASUDI Charles, Président de la Société Civile du Territoire de Walikale, Atteste par la présente avoir reçu Monsieur DUNIA BUTINDA Lewis, superviseur des activités de la collecte des données du professeur Raul SANCHEZ DE LA SIERRA dans le Territoire de Walikale, Cette collecte entre dans le cadre de la recherche sur « comment la population fait face au défis économiques dans la zone contrôlé par les groupes.

Le Superviseur DUNIA BUTINDA Lewis s'est présenté dan nos bureaux avec tous les documents nécessaire relative à cette recherche (protocole de la recherche, questionnaire d'enquête, formulaire de consentement), les quels nous avions lis avec intérêt et n'avons trouvé aucune chose pouvant aller à la contre des principes éthiques de la population ; en foi de quoi nous lui livrons ce document d'avis favorable pour lui servir à toute fin utile.

Fait à Walikale, le 17/10/2018

Pour La Société Civile de Walikale

KISA MASUDI Charles
Président

REPUBLIQUE DEMOCRATIQUE DU CONGO

PROVINCE DU NORD – KIVU
TERRITOIRE DE WALIKALE
SECTEUR DE WANIANGA
GROUPEMENT KISIMBA

AVIS FAVORABLE DE LA CONDUITE D'UNE ENQUETE DANS NOTRE ENTITE

Je soussigné Mwami KABAKI MWANANKOYO II ALIMASI MURONGANI Chef de Groupement KISIMBA

Atteste par la présente avoir reçu Monsieur DUNIA BUTINDA Lewis, superviseur des activités de la collecte des données du professeur Raul SANCHEZ DE LA SIERRA dans le groupement KISIMBA

Cette collecte entre dans le cadre de la recherche sur « comment la population fait face au défis économiques dans la zone contrôlé par les groupes

Le superviseur DUNIA BUTINDA Lewis s'est présenté dan nos bureaux avec tous les documents nécessaire relative à cette recherche (protocole de la recherche, questionnaire d'enquête, formulaire de consentement), les quels nous avions lis avec intérêt et n'avons trouvé aucune chose pouvant aller à la contre des principes éthiques de la population ; en foi de quoi nous lui livrons ce document d'avis favorable pour lui servir à toute fin utile.

Fait à PINGA le, 25 Septembre 2018

Le Groupement

Mwami KABAKI MWANANKOYO II ALIMASI MURONGANI

REPUBLIQUE DEMOCRATIQUE DU CONGO

PROVINCE DU NORD – KIVU
TERRITOIRE DE WALIKALE
SECTEUR DE WANIANGA
GROUPEMENT D'UTUNDA

AVIS FAVORABLE DE LA CONDUITE D'UNE ENQUETE DANS NOTRE ENTITE

Je soussigné BAKORA SHEBIHEMBE Edouard Chef de Groupement UTUNDA, Atteste par la présente avoir reçu Monsieur DUNIA BUTINDA Lewis, superviseur des activités de la collecte des données du professeur Raul SANCHEZ DE LA SIERRA dans le groupement d'Utunda. Cette collecte entre dans le cadre de la recherche sur « comment la population fait face au défis économiques dans la zone contrôlé par les groupes.

Le Superviseur DUNIA BUTINDA Lewis s'est présenté dan nos bureaux avec tous les documents nécessaire relative à cette recherche (protocole de la recherche, questionnaire d'enquête, formulaire de consentement), les quels nous avions lis avec intérêt et n'avons trouvé aucune chose pouvant aller à la contre des principes éthiques de la population ; en foi de quoi nous lui livrons ce document d'avis favorable pour lui servir à toute fin utile.

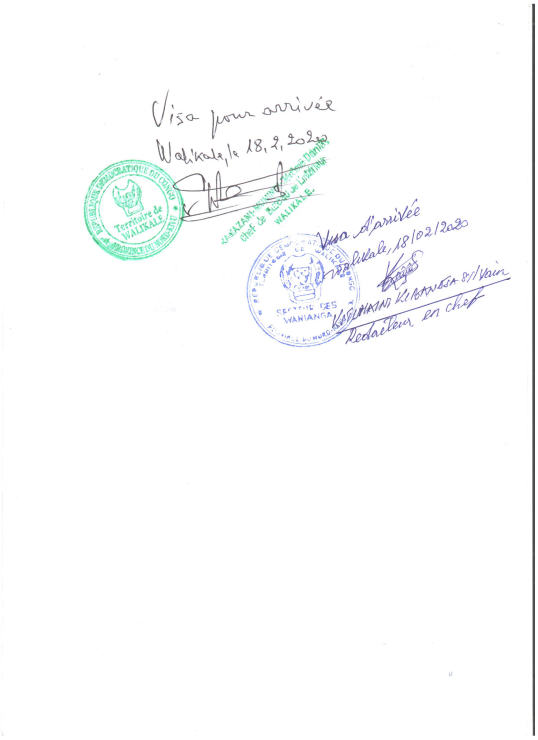
Fait à Mpofi, le 23/09/2018

Pour le Groupement

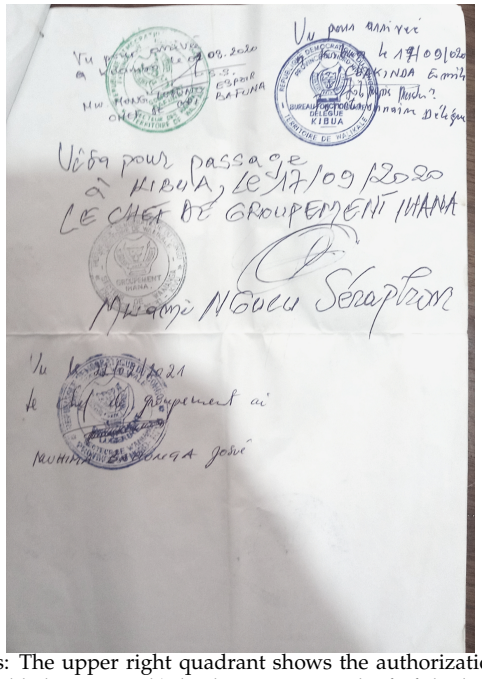
Mwami BAKORA SHEBIHEMBE Edouard

Notes: This figure presents local authorizations. Panel A shows the authorization by: a. the civil society of the district of Walikale (upper right quadrant); b. the chiefs of the affected groupement of villages (upper right and lower left quadrants). c the National army sections positioned in the area as well as individual village customary chiefs (lower right quadrant). The time of the study, the district of Walikale, along with the rest of the province, was run by the militiary upon presidential decree due to the security situation, hence the provincial military authority had taken over the civilian administration. The province of North Kivu was sub-divided into districts (Territoires, in French), one of which is the district of Walikale. The district of Walikale was in turn sub-divided into two sectors (Wanianga and Bakano), each headed by a chief of sector, representing the state authority. Each sector was divided into groupments (Groupement, in French), head headed by a customary King (Mwami, in Swahili).

Panel B. Additional Authorizations from the Civilian, Customary, and State Authorities

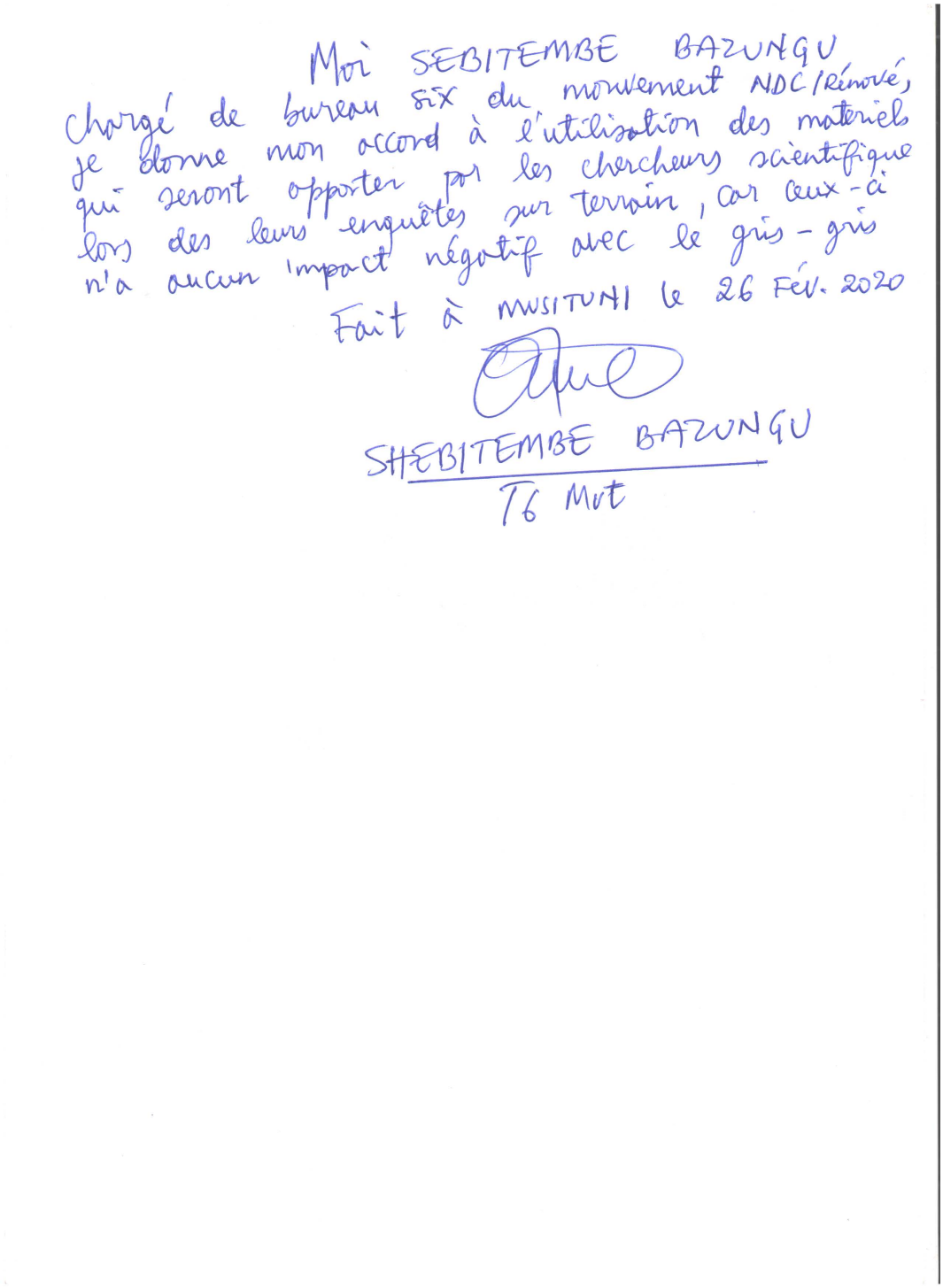






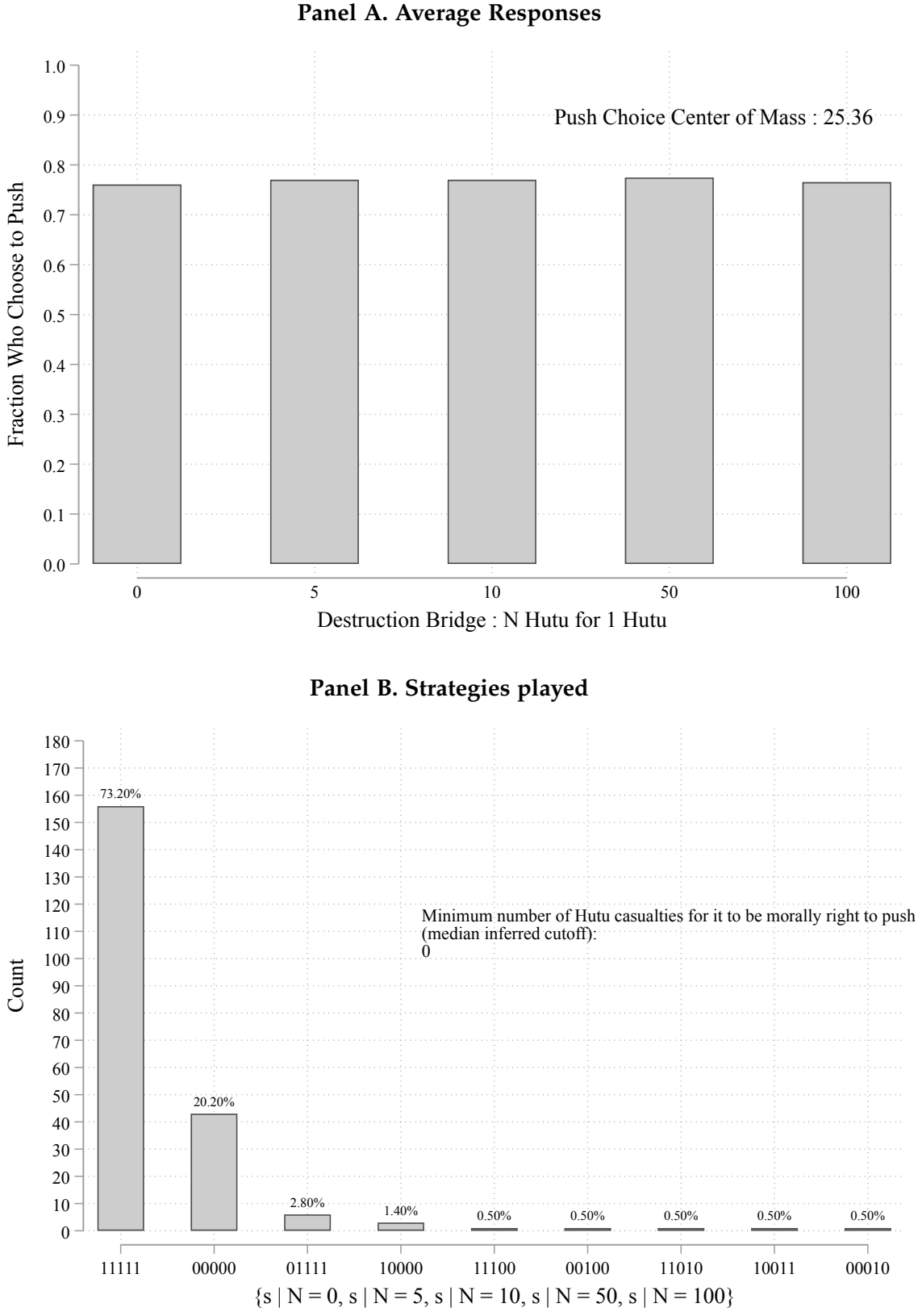
Notes: The upper right quadrant shows the authorizations by: a. the state administrator of the district of Walikale (Territoire de Walikale, in French); b. the customary chief of the kingdom of the Wanianga (Secteur des Wanianga, in French); the district of Walikale is sub-divided into two customary-run kingdoms, one of which is where the surveys took place (Wanianga). The upper right quadrant shows the authorizations by: a. the National Army's headquarters authority of the district of Walikale (Etat Major, in French); at the time of the study, the district of Walikale, along with the rest of the province, was run by the militiary upon presidential decree due to the security situation, hence the provincial military authority had taken over the civilian administration; b. by the national intelligence agency of the district of Walikale (Agence Nationale des Renseignements, in French); c. the civil society coordination of the district of Walikale (Societe Civile Forces Vives, in French); d. the Congolese National Police/Ministry of the Interior (Police Nationale Congolaise/Ministere de l'Interieur, in French). The lower left quadrant shows the authorizations of: a. the chief of the other sector (Bakano); b. the provincial government delegate in Kibua; c. the customary authority of the groupment of Ihana; d. the customary authority of the groupment of Luberiki.

Panel C. Authorization from NDC'R's mystico-religious affairs agency regarding use of bracelets measuring EDA (Bureau 6)


Moi SEBITEMBE BAZUNGU
Chargé de bureau six du mouvement NDC/Rénové,
je donne mon accord à l'utilisation des materiels
qui seront apporter par les chercheurs scientifique
lors des leurs enquêtes sur terrain, car ceux-ci
n'a aucun impact négatif avec le gris-gris
Fait à MWSITUNI le 26 Fév. 2020
SHEBITEMBE BAZUNGU
T6 Mut

Notes: This Panel shows ~~the authorizations by the NDC-R to conduct the research. The upper left and upper right quadrants present authorization to the study and its amendment, respectively. The lower left quadrant shows~~ the authorization by the bureau 6 commander (the T6) to use medical equipment on volunteers, such as the bracelets measuring emotional arousal; the bureau 6 is the agency in charge of mystico-religious affairs of the group, and the authorization stipulates that such equipment does not interfere with their practice. ~~The lower right quadrant shows the ultimate authorization to conduct the research, upon presenting the final surveys and protocols, obtained in April 2021.~~

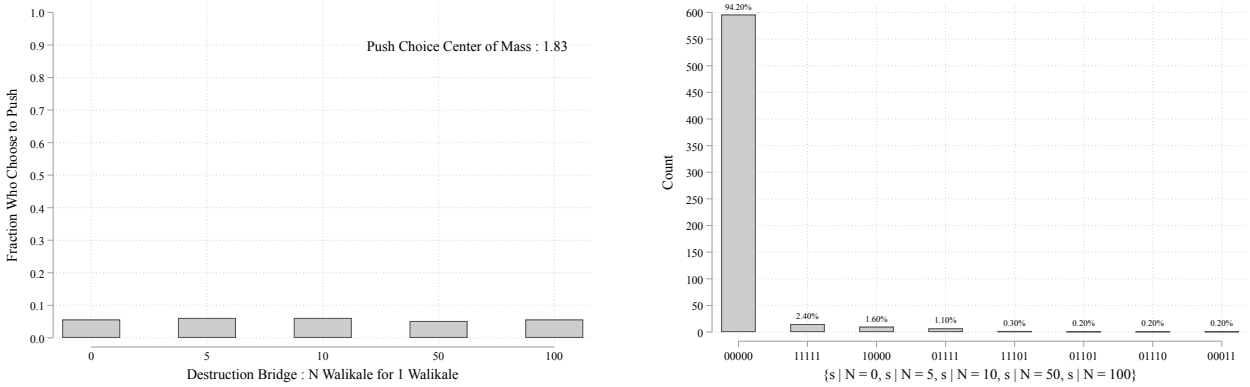
Figure A9: The Prevailing Moral Values: The Killing Dilemma with Hutu Casualties



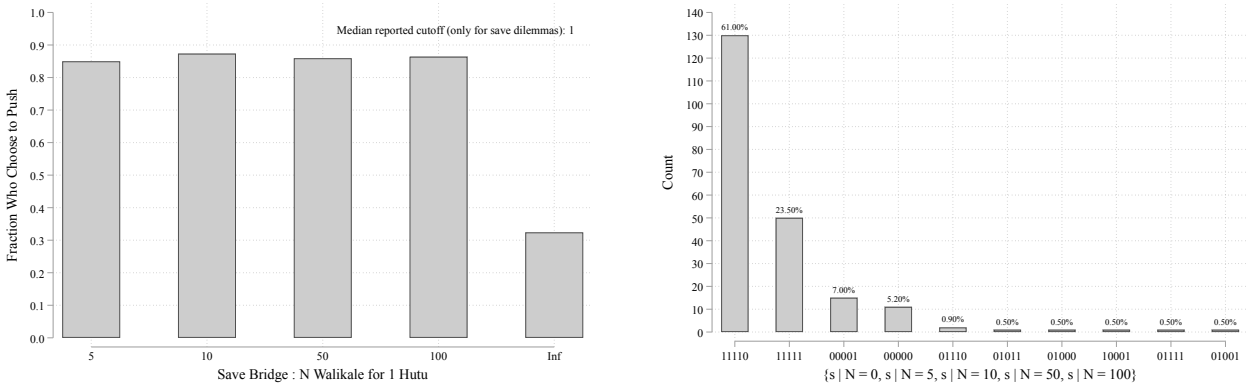
Notes. This figure presents it for the dilemma in which respondents are asked to ascertain whether pushing a Rwandan Hutu off the bridge is morally right, when it results in the death of $N \in \{0, 5, 10, 50, 100\}$ Rwandan Hutus. The term $s|N \in \{0,1\}$ denotes the choice of whether to push the Rwandan Hutu off the bridge ($s = 1$) or not join ($s = 0$) when pushing implies, in addition to his own death that N members of the Rwandan Hutu Community would also die (Panel B).

Figure A10: The Prevailing Moral Values: Robustness

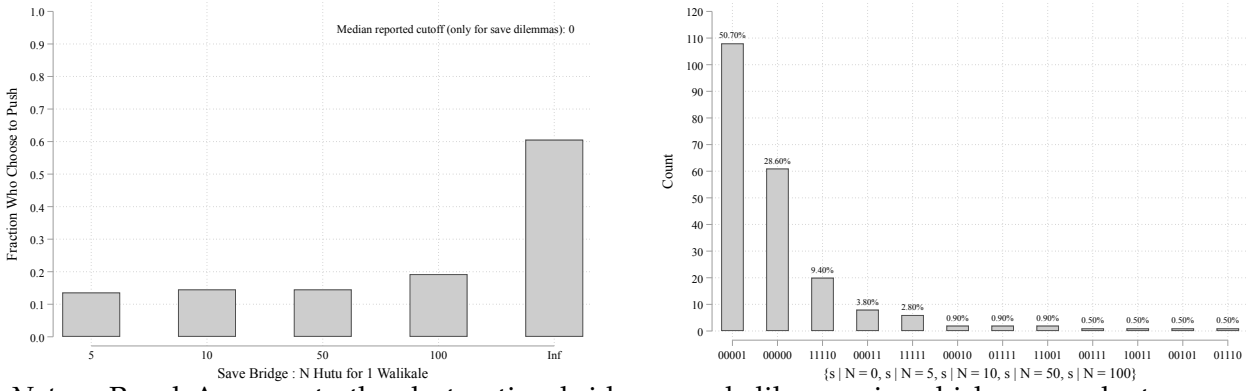
Panel A. Is Killing Own Community Members also Morally Right?



Panel B. Saving Bridge Dilemma: Saving N Walikale by Sacrificing 1 Hutu



Panel C. Saving Bridge Dilemma: Saving N Hutu by Sacrificing 1 Walikale



Notes. Panel A presents the destruction bridge moral dilemma in which respondents are asked to ascertain whether pushing a Walikale civilian off the bridge is morally right, when it results in the death of another $N \in \{0, 5, 10, 50, 100\}$ Walikale civilians. Panels B and C present the standard bridge moral dilemma in which not pushing results in the death of the people in the rail, but pushing saves them while killing the person pushed.

Figure A11: Anger Associated to the War

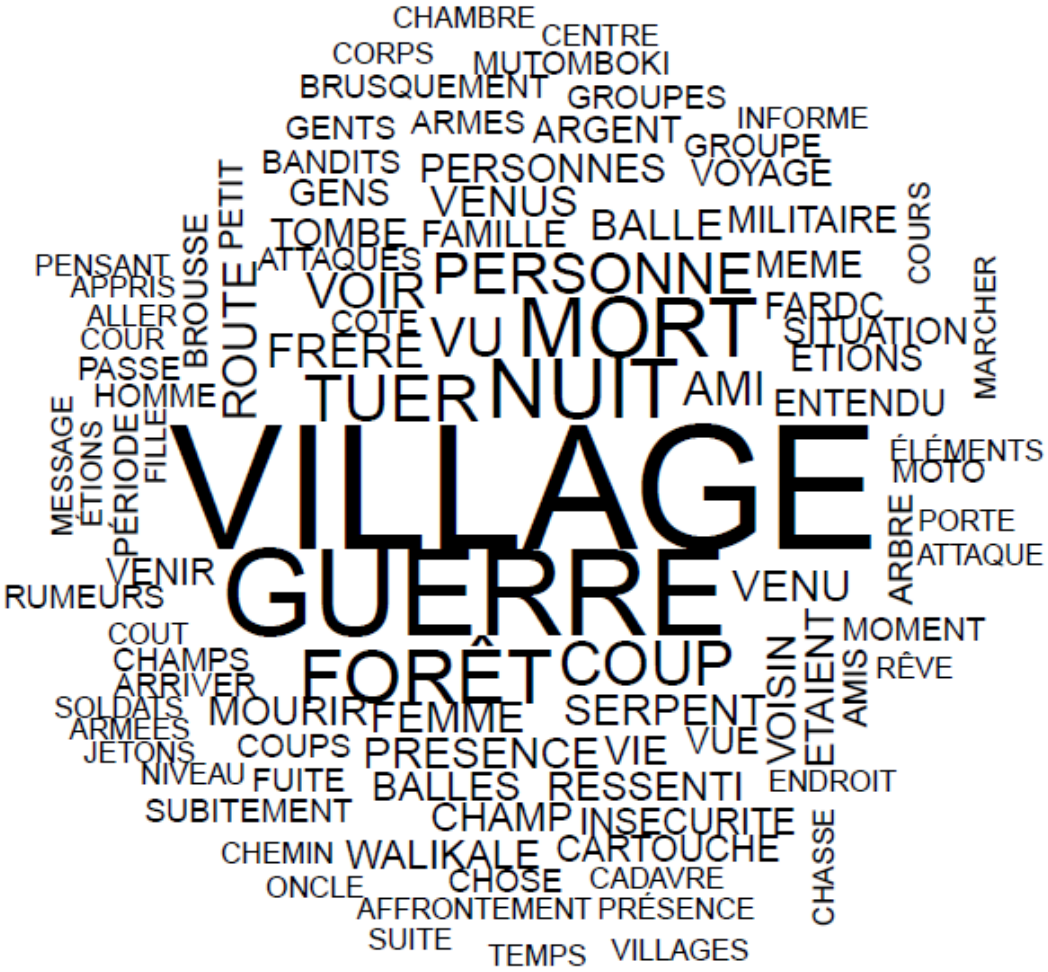
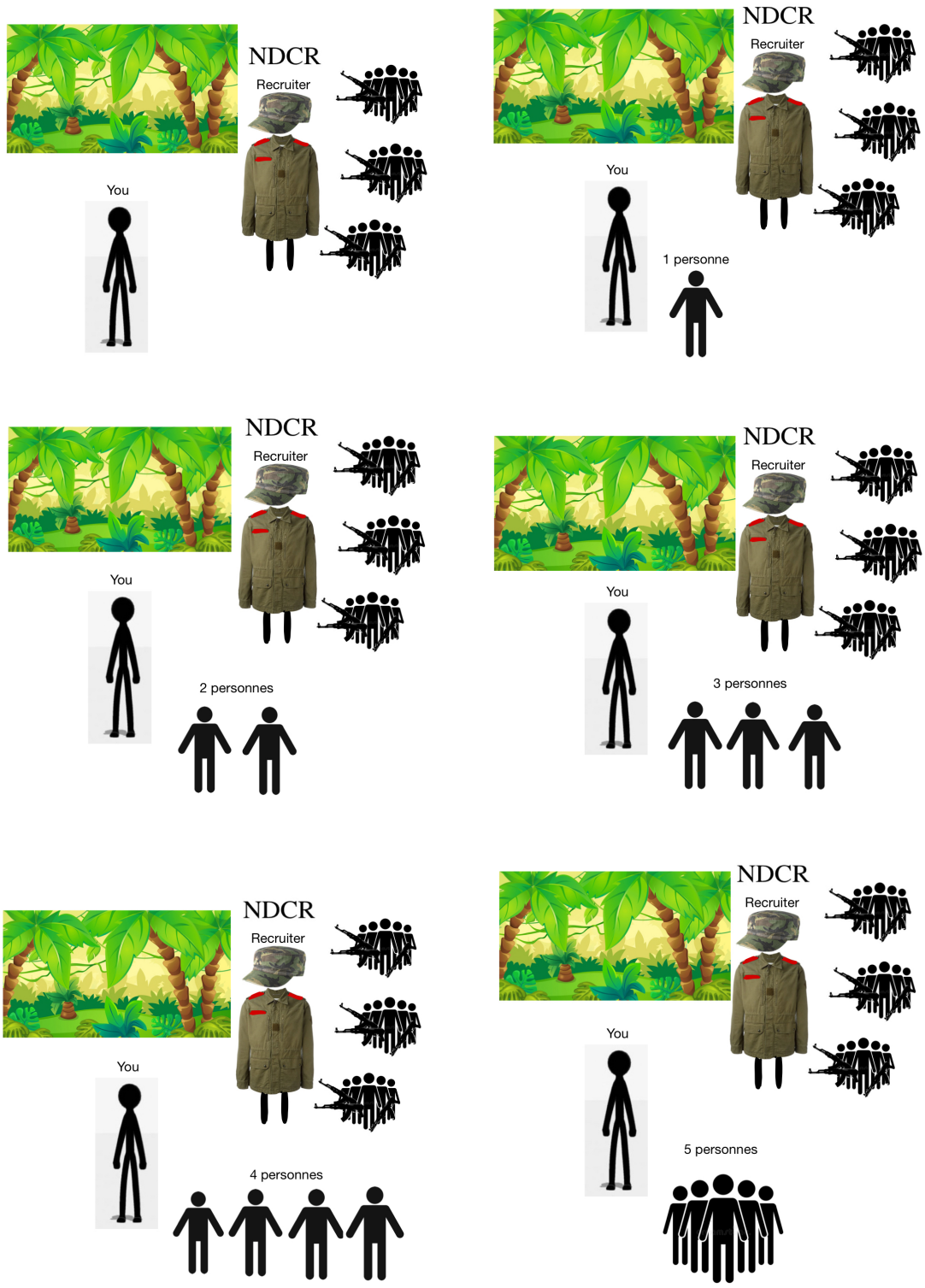


Figure A12: The Morality of Joining: Screenshots



Notes:

Table A2: Correlation Between Altruism and True Behavior—Raw correlations

	Mean Disposition		Regression coeff. Dep vars: Assessments of Joiner's Behavior						
	Baseline	Endline	Baseline	Harm of Own Civilians [0-10]. Report by			Enemy Harm [0-10]. Report by		
				Supervisor	Peer	Self	Supervisor	Peer	Self
Endline Altruism to Community									
Disposition for Inter-Personal Harm									
Empathetic Concern [0-10]	9.00	9.31***	0.015	-0.554***	-0.619***	-0.325***	1.066***	1.107***	1.024***
			(0.071)	(0.121)	(0.145)	(0.122)	(0.110)	(0.109)	(0.103)
Affective Empathy [0-10]	8.44	8.66*	-0.009	-0.472***	-0.345***	-0.436***	0.692***	0.832***	0.665***
			(0.075)	(0.100)	(0.128)	(0.097)	(0.099)	(0.100)	(0.092)
Cognitive Empathy [0-10]	8.36	8.70***	-0.108	-0.401***	-0.687***	-0.344***	0.676***	0.871***	0.653***
			(0.078)	(0.100)	(0.123)	(0.105)	(0.097)	(0.100)	(0.099)
Inclination (Taste) [0-10]	0.02	0.05	0.000	0.087	0.447	-0.299	0.020	-0.003	0.044
			(0.000)	(0.351)	(0.423)	(0.235)	(0.366)	(0.375)	(0.239)
Inclination (Preference) [0-10]	0.14	0.05	-0.015	0.000	0.000	0.052	0.000	0.000	-0.129
			(0.120)	(0.000)	(0.000)	(0.166)	(0.000)	(0.000)	(0.169)
Prosociality (Dictator Game)	2.97	1.01***	0.132	-0.066	-0.173**	-0.070	0.043	0.093	0.126**
			(0.115)	(0.059)	(0.081)	(0.058)	(0.062)	(0.072)	(0.059)
Endline Altruism to Enemy									
Disposition for Inter-Personal Harm									
Empathetic Concern [0-10]	2.85	1.86***	0.069	0.344***	0.410***	0.316***	-0.321***	-0.271***	-0.260***
			(0.091)	(0.042)	(0.050)	(0.043)	(0.045)	(0.047)	(0.046)
Affective Empathy [0-10]	3.42	1.93***	0.053	0.332***	0.427***	0.334***	-0.256***	-0.254***	-0.187***
			(0.094)	(0.039)	(0.050)	(0.038)	(0.044)	(0.048)	(0.044)
Cognitive Empathy [0-10]	4.96	6.99***	0.017	-0.066**	-0.069*	-0.037	0.098***	0.101***	0.071**
			(0.074)	(0.032)	(0.038)	(0.030)	(0.033)	(0.034)	(0.031)
Inclination (Taste) [0-10]	6.30	7.11*	-0.021	-0.161***	-0.220***	-0.137***	0.205***	0.219***	0.199***
			(0.074)	(0.025)	(0.028)	(0.024)	(0.024)	(0.023)	(0.023)
Inclination (Preference) [0-10]	8.29	8.61	0.069	-0.104***	-0.162***	-0.035	0.268***	0.335***	0.316***
			(0.074)	(0.039)	(0.043)	(0.036)	(0.037)	(0.032)	(0.029)
Prosociality (Dictator Game)	-2.31	-3.76***	-0.044	0.179***	0.132*	0.049	-0.434***	-0.337***	-0.365***
			(0.097)	(0.056)	(0.071)	(0.060)	(0.052)	(0.059)	(0.055)
Mean Dep. Var				1.57	1.78	1.48	8.90	9.01	9.28
Observations	211	209	209	209	209	204	209	209	204

Table A3: Altruism Unbundled: Prosociality and Empathy—All Targets Pooled, Undersatnding the Correlations

A. Empathy to X on Prosociality to X

	Empathy to X								
	(1) Concern	(2) Affective	(3) Cognitive	(4) Concern	(5) Affective	(6) Cognitive	(7) Concern	(8) Affective	(9) Cognitive
Prosociality to X	0.410***	0.371***	0.231***	1.335***	1.094***	0.669***	0.558***	0.481***	0.388***
	(0.034)	(0.038)	(0.034)	(0.038)	(0.035)	(0.038)	(0.051)	(0.048)	(0.057)
Target FE	✓	✓	✓				✓	✓	✓
Individual FE				✓	✓	✓	✓	✓	✓
Mean dep. var	68.11	68.26	77.19	68.11	68.26	77.19	68.11	68.26	77.19
R^2	0.59	0.44	0.23	0.68	0.65	0.49	0.75	0.70	0.54
Observations	1902	1902	1902	1902	1902	1902	1902	1902	1902

B. Sadism to X on Prosociality to X

	Sadism to X					
	(1) Preference	(2) Taste	(3) Preference	(4) Taste	(5) Preference	(6) Taste
Prosociality to X	-0.371***	-0.229***	-1.389***	-1.777***	-0.628***	-0.638***
	(0.040)	(0.059)	(0.053)	(0.046)	(0.070)	(0.076)
Target FE	✓	✓			✓	✓
Individual FE			✓	✓	✓	✓
Mean dep. var	19.29	30.05	19.29	30.05	19.29	30.05
R^2	0.47	0.54	0.59	0.64	0.67	0.73
Observations	1910	1914	1910	1914	1910	1914

Table A4: Past Victimization by the Hutus: The Facts

	Mean	Min	25th Percentile	Median	75th Percentile	Max
All Attacks	5.63	0.00	0.00	2.00	8.00	47.00
Enemy Attacks						
Murder	1.92	0.00	0.00	1.00	3.00	12.00
Sexual violence	0.67	0.00	0.00	0.00	1.00	10.00
Torture	1.67	0.00	0.00	0.00	2.00	12.00
looting	0.55	0.00	0.00	1.00	1.00	1.00
Non-enemy Attacks						
Murder	0.67	0.00	0.00	0.00	0.00	12.00
Sexual violence	0.15	0.00	0.00	0.00	0.00	8.00
Torture	0.57	0.00	0.00	0.00	0.00	12.00
looting	0.30	0.00	0.00	0.00	1.00	1.00

Notes.

Appendix B. Framework Extension: Prevailing Morality of Violence and the Correlation Between Prosociality and Empathy to the Enemy

Suppose that both prosociality towards the community and empathy towards the enemy are jointly governed by a common "goodness" parameter, $\tilde{\lambda} \in \mathbb{R}$, as stipulated in the following assumption:

Assumption B *Morality and the Correlation*

$$\tilde{\lambda}^E = \tilde{\lambda}\left(1 - \tilde{s}^C m\right)$$

and

$$\tilde{s}^C = \tilde{\lambda},$$

where $m \in \mathbb{R}$ is a scale parameter measuring the intensity of the prevailing moral values about the moral righteousness of killing enemies.

Applying Assumption B, it follows that:

$$\frac{cov(\tilde{s}^C, \tilde{\lambda}^E)}{\partial m} < 0$$

and

$$cov(\tilde{s}^C, \tilde{\lambda}^E) < 0 iff. m > 1$$

Table A5: Joiners vs. Non-Joiners: Controlling for Economic Variables

	Household (Mean)	Joiner - HH (Difference)	Villager (Mean)	Joiner - Vill (Difference)	Pooled (Mean)	Joiner - Pooled (Difference)
	(1)	(2)	(3)	(4)	(5)	(6)
A. Aversion/Inclination to Interpersonal Harm						
A.1. Aversion/Inclination to Interpersonal Harm on the Enemy						
Empathy to enemy (Index)	49.72	-8.11 **	45.39	-7.71 *	47.69	-7.82 **
Cognitive Empathy	68.57	-18.79 ***	61.01	-23.22 ***	65.01	-20.44 ***
Affective Empathy	43.59	-1.10	40.36	2.02	42.07	0.34
Empathic concern	37.01	-4.45	34.81	-1.94	35.98	-3.36
Sadism to enemy (Index)	60.92	4.47	69.02	-4.02	64.74	0.68
Taste (% enjoys pain of person)	51.46	2.36	55.43	-3.48	53.33	-0.33
Preference (% murders in video game)	70.39	6.58	82.61	-4.56	76.15	1.68
A.2. Aversion/Inclination to Interpersonal Harm on Community						
Empathy to community (Index)	85.53	-0.42	87.88	-1.55	86.64	-0.86
Cognitive Empathy	82.38	1.52	84.64	0.05	83.44	0.93
Affective Empathy	82.91	1.56	86.56	-0.19	84.63	0.80
Empathic concern	91.31	-4.35 ***	92.43	-4.51 ***	91.84	-4.32 ***
Sadism to community (Index)	0.49	0.28	1.09	0.62	0.77	0.33
Taste (% enjoys pain of person)	0.49	-0.46	0.00	0.34	0.26	-0.16
Preference (% murders in video game)	0.49	1.01	2.17	0.91	1.28	0.83
A.3. Psychopathy						
Primary Psychopathy	34.64	-0.51	34.09	-2.07 **	34.38	-1.30 *
Secondary Psychopathy	17.14	2.11 **	18.71	-2.38 **	17.88	-0.02
B. Prosociality						
B.1 Prosociality to community (index)	37.43	7.36 ***	44.46	1.47	40.74	4.71 ***
Income-sensitive prosociality measures	21.71	9.01 ***	31.66	-1.61	26.34	3.89 **
Hypothetical donation	26.31	3.25	32.72	-3.08	29.33	0.35
Dictator game (in %)	20.63	14.46 ***	33.80	1.74	26.85	8.68 ***
Index long sample measures	23.47	8.86 ***	33.26	-0.67	28.09	4.52 ***
Time dictator game working (in %)	18.52	9.83 ***	23.75	0.66	20.98	5.26 **
Time dictator game average (in %)	22.02	9.78 ***	29.72	2.26	25.60	6.19 **
Index short sample measures	20.04	9.77 ***	27.78	0.50	23.64	5.26 **
Income-robust prosociality measures	45.97	8.51 ***	50.79	3.54	48.21	6.12 ***
Declared prosociality	65.34	4.36 **	66.85	5.75 **	66.05	5.11 **
Socializing dictator game (in %)	21.80	7.88 **	26.48	-1.13	23.87	3.88
B.2 Prosociality to enemy (index)	6.33	1.57	8.57	-2.57	7.38	-0.32
Income-sensitive prosociality measures	3.76	0.77	6.19	-1.86	4.89	-0.52
Hypothetical donation	2.82	0.48	4.13	-3.17 *	3.44	-1.23
Dictator game (in %)	2.33	-0.05	4.02	-0.23	3.13	-0.12
Index long sample measures	2.57	0.21	4.08	-1.70	3.28	-0.67
Time dictator game working (in %)	4.78	1.98	6.88	-1.58	5.74	0.16
Time dictator game average (in %)	6.13	-0.16	8.15	-2.35	7.07	-1.19
Index short sample measures	5.48	0.85	7.55	-2.13	6.44	-0.62
Income-robust prosociality measures	10.48	1.82	12.82	-4.30	11.57	-1.19
Declared prosociality	13.83	4.29	17.55	-4.31	15.59	0.39
Socializing dictator game (in %)	5.98	1.14	6.77	-2.54	6.35	-0.66
Observations	184	385	206	407	407	591

Table A6: Joiners vs. Non-Joiners: Economic Differences

	Household (Mean)	Joiner - HH (Difference)	Villager (Mean)	Joiner - Vill (Difference)	Pooled (Mean)	Joiner - Pooled (Difference)
	(1)	(2)	(3)	(4)	(5)	(6)
A. Income and Work						
Income	28.86	-27.45 *** ***	27.46	-26.05 *** ***	28.20	-26.79 *** ***
Number of economic activities past 7 d	0.69	-0.58 *** ***	0.66	-0.55 *** ***	0.68	-0.57 *** ***
Looking for work past 7 d	0.26	0.10 ** **	0.24	0.12 * *	0.25	0.11 ** **
B. Education						
Years of education	10.49	-1.20 *** ***	10.05	-0.75 ** **	10.28	-0.99 *** ***
Secondary education	0.75	0.06	0.79	0.01	0.77	0.03
Observations	184	385	206	407	407	591

Proof By Assumption B:

$$cov(\tilde{s}^C, \tilde{\lambda}^E = cov(\tilde{\lambda}, \tilde{\lambda}(1 - m\tilde{s}^C,))$$

After minimal algebra, it follows that:

$$cov(\tilde{s}^C, \tilde{\lambda}^E = var(\tilde{\lambda}) \left(1 - m^2\right)$$

□

Table A7: Cognitive Processes that Rationalize Violence Against the Hutus: Moral Disengagement

	Mean of Random Civilians, of which			
		Bridge Dilemma		
	All	Kills	Does Not Kill	Difference
A. Moral Justification (Index)	**0.71**	**0.63**	**0.76**	**0.130*** **
It's OK to be violent on Hutus to protect the community.	0.60	0.53	0.64	0.112**
It's OK to beat a Hutu who hurt Walikaleans.	0.80	0.71	0.85	0.139***
It's OK to be violent to a Hutu to restore population honor.	0.71	0.64	0.75	0.113**
It's OK to lie to a Hutu to get you out of trouble.	0.73	0.63	0.78	0.155***
B. Euphemistic Language (Index)	**0.65**	**0.54**	**0.72**	**0.182*** **
Kicking is the only language Hutus understand.	0.64	0.53	0.71	0.184***
Being violent with a Hutu, it's just giving them a lesson.	0.70	0.59	0.77	0.180***
Taking Hutu's beer without paying them is just borrowing.	0.62	0.50	0.69	0.194***
It's not bad to hit Hutus a bit from time to time.	0.65	0.55	0.72	0.170***
C. Advantageous Comparison (Index)	**0.62**	**0.53**	**0.67**	**0.136*** **
Damaging Hutu's things is OK given they are abusing us.	0.63	0.56	0.68	0.126**
Stealing a bit of cash from a Hutu is ok given they steal too.	0.67	0.56	0.74	0.183***
Torturing a Hutu is OK, because killing it would be worse.	0.56	0.51	0.59	0.079
Given illegal things they do, stealing from a Hutu is OK.	0.60	0.50	0.66	0.157***
D. Diffusion of Responsibility (Index)	**0.56**	**0.49**	**0.61**	**0.115*** **
A combatant cannot be blamed if his platoon kills H.c.	0.59	0.51	0.64	0.125**
A commander abusing H.c. cannot be blamed when others also do it.	0.57	0.49	0.63	0.135***
If a group hurts a H.c., it's unfair to blame an individual for it.	0.53	0.49	0.55	0.057
It's unfair to blame a combatant if his platoon hurt Hutus.	0.56	0.47	0.61	0.142***
E. Displacement of Responsibility (Index)	**0.61**	**0.52**	**0.67**	**0.145*** **
If combatants have bad conditions, can't be blamed for abusing H.c.	0.61	0.50	0.68	0.177***
If combatants are not trained, they can't be blamed for misconduct.	0.54	0.48	0.58	0.097*
Combatants can't be blamed for hurting H.c. when peers do it too.	0.62	0.55	0.67	0.119**
Combatants can't be blamed for torturing H.c. if peers pressured.	0.67	0.55	0.74	0.189***
F. Distorting Consequences (Index)	**0.64**	**0.55**	**0.70**	**0.153*** **
It's ok to lie a Hutu because he doesn't really feel pain.	0.64	0.53	0.71	0.182***
Hutu doesn't mind to be teased or confused.	0.60	0.54	0.64	0.100*
Teasing a Hutu does not really hurt them.	0.66	0.59	0.71	0.122***
Hassling a Hutu does not hurt anyone.	0.66	0.53	0.74	0.208***
G. Dehumanization (Index)	**0.64**	**0.52**	**0.70**	**0.180*** **
Hutus deserve to be treated as animals.	0.60	0.48	0.67	0.193***
It is normal to treat someone who behaved like a rat.	0.71	0.63	0.76	0.127***
A Hutu does not deserve to be treated as a human.	0.56	0.45	0.63	0.184***
Hutu must be treated violently, they lack feelings, can't feel pain.	0.67	0.54	0.75	0.217***
H. Attribution of Blame (Index)	**0.68**	**0.60**	**0.73**	**0.120*** **
If combatants hurt Hutu, it's commander's fault.	0.60	0.56	0.63	0.070
If Hutu leaves property unsupervised, it's their fault if stolen.	0.66	0.60	0.71	0.110**
Hutu who have been abused usually did things to deserve it.	0.74	0.64	0.81	0.166***
Combatants can't be blamed, if commanders pressure.	0.71	0.62	0.76	0.134***
Moral Disengagement Index	**0.64**	**0.55**	**0.69**	**0.145*** **

Notes.

Table A8: Mechanisms II: Moral Convictions or Motivated Reasoning?

	Mean of Random Civilians, of which						
		Bridge Dilemma			Morally Right To		
	All	Kills	Does Not Kill	Difference	Join	Not Join	Difference
A. Moral Disengagement Index	**0.64**	**0.69**	**0.55**	**0.145*****	**0.68**	**0.57**	**0.108*****
Moral Justification Index	0.71	0.76	0.63	0.130***	0.73	0.66	0.077**
Advantageous Comparison Index	0.62	0.67	0.53	0.136***	0.65	0.56	0.082**
Attribution of Blame Index	0.68	0.73	0.60	0.120***	0.71	0.63	0.074**
Diffusion of Responsibility Index	0.56	0.61	0.49	0.115***	0.58	0.54	0.040
Displacement of Responsibility Index	0.61	0.67	0.52	0.145***	0.66	0.53	0.132***
Distorting Consequences Index	0.64	0.70	0.55	0.153***	0.70	0.54	0.155***
Dehumanization Index	0.64	0.70	0.52	0.180***	0.69	0.54	0.149***
Euphemistic Language Index	0.65	0.72	0.54	0.182***	0.71	0.55	0.158***
B. Altruism Towards Enemy							
Cognitive Empathy to Enemy	5.87	5.86	5.88	-0.020	5.27	6.94	-1.669***
Affective Empathy to Enemy	4.05	3.44	5.04	-1.598***	3.32	5.35	-2.025***
Empathetic Concern to Enemy	3.50	2.90	4.48	-1.585***	2.75	4.85	-2.099***
Sadistic Taste to Enemy	5.51	6.70	3.40	3.309***	6.35	3.82	2.535***
Sadistic Preference to Enemy	7.87	8.64	6.91	1.723***	8.39	7.24	1.157**
Prosociality to Enemy	0.90	0.56	1.47	-0.910***	0.83	1.06	-0.230
Prosociality to Enemy x Morally Right to Join	0.53	0.43	0.69	-0.257*	0.83	0.00	0.827***
B. Altruism Towards Community							
Cognitive Empathy to Community	8.42	8.45	8.38	0.068	8.39	8.48	-0.087
Affective Empathy to Community	8.65	8.73	8.54	0.188	8.63	8.70	-0.069
Empathetic Concern to Community	9.15	9.22	9.04	0.178	9.14	9.18	-0.038
Sadistic Taste to Community	0.00	0.00	0.00	0.000	0.00	0.00	0.000
Sadistic Preference to Community	2.78	3.79	1.23	2.553	2.92	2.63	0.288
Prosociality to Community	4.43	4.41	4.49	-0.081	4.36	4.59	-0.230
Prosociality to Community x Morally Right to Join	2.80	3.03	2.44	0.585*	4.36	0.00	4.358***
D. Emotions							
Anger Related to the War	0.23	0.29	0.13	0.158**	0.28	0.16	0.116*
Intensity of Anger	8.07	8.51	7.32	1.189***	8.16	7.90	0.261
Intensity of Anger Related to the War	2.10	2.69	1.07	1.619***	2.49	1.43	1.057*
E. Joining Status							
Is Joiner (Joiner and Villager Sample)	0.50	0.51	0.47	0.044	0.57	0.27	0.301***
Is Joiner (Joiner and Household Member Sample)	0.50	0.54	0.43	0.116**	0.65	0.19	0.460***