Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

Tracking the SDGs: A methodological note on measuring deaths caused by collective violence

Anke Hoeffler, Frederike Kaiser, Birke Pfeifle, and Flora Risse

Anke Hoeffler is Professor of Development Policy, Department of Political Science and Public Administration, University of Konstanz, Germany. She can be reached at anke.hoeffler@uni-konstanz.de. Frederike Kaiser, MA Political Science and Public Administration (University of Konstanz), works at GIZ within the Corporate Unit for Evaluation. She can be reached at kaifrederike@gmail.com. Birke Pfeifle, MA student Political Science and Public Administration at the University of Konstanz. She can be reached at birke.pfeifle@uni-konstanz.de. Flora Risse, MA Political Science and Public Administration (University of Konstanz), is a consultant for strategic modernization of public administrations. She can be reached at flora.risse@posteo.de.

Abstract

As part of recording the progress toward promoting peaceful societies as envisioned in the Sustainable Development Goal (SDG) 16, it is important to provide accurate estimates of violence-related deaths (SDG 16.1). These estimations face a number of methodological challenges, resulting in rather conservative estimates in the social sciences. In this article, we discuss SDG indicator 16.1.2 on conflict-related deaths, proposing its enlargement to cover different forms of collective violence. Various types of collective violence, their definition, measurement, and methods to combine them without double counting are reviewed. Comparing the Georeferenced Events Dataset (GED) to the Global Terrorism Database (GTD) shows that events of armed conflict and terrorism overlap to a certain degree. Our argument is that merging data from different event databases can provide a more accurate account of collective violence. We augment the GED data on organized armed conflict with data on terrorism—as a result, our estimates of the numbers of collective violence-related deaths are indeed significantly higher than suggested by GED (one of the most widely used databases in the social sciences).

In 2000, the United Nations set out an aspirational agenda, the Millennium Development Goals (MDGs). The goals ranged from reducing poverty and hunger, to achieving universal primary education, and to combating diseases such as HIV/AIDS and malaria. Fifteen years later global poverty had been more than halved and the MDGs were judged to have produced the most successful anti-poverty movement in history (United Nations, 2015). However, across the world progress had been uneven and many challenges to human development remained. This inspired the new ambitious 2030 Agenda for Sustainable Development, consisting of seventeen Sustainable Development Goals (SDGs). Unlike the MDGs, the SDGs include a target to promote peaceful societies, aiming to reduce all forms of violence and related deaths everywhere (SDG target 16.1). In order to gauge progress, two important indicators are the number of intentional homicides (16.1.1) and conflict-related deaths (16.1.2). In this article we discuss why this distinction results in an undercounting of violent deaths. Further, we make suggestions on how to address this problem by measuring "collective violence" instead of only conflict-related deaths. We start with the observation that conflict-related death figures provided by social scientists tend to be conservative and argue that merging data from different event databases can provide a more accurate account of collective violence-related deaths. In the following section, we elaborate on the definition of conflict-related deaths followed by an introduction of commonly used event datasets. Next, recently developed methods of merging these data are discussed and are then applied to present regional and

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

global estimates of collective violence-related deaths. The last section provides conclusions and discusses avenues for further research.

Measuring progress on the Sustainable Development Goal 16 currently faces a number of methodological challenges in the estimation of violence-related deaths. We propose enlarging on the current estimates to cover different forms of collective violence. Careful merging of data from different established event databases can provide a more accurate account. As a result, estimates of the numbers of collective violence-related deaths are significantly higher than suggested by the Georeferenced Events Dataset (GED).

Definitions and data

To track the development and the achievements on SDG 16.1, indicator 16.1.2 aims to capture conflict-related deaths.[1] Here the U.N. understands conflict as the "protracted armed confrontations occurring between governmental armed forces and of one or more armed groups". Two types of conflict-related deaths are considered: First, direct deaths resulting from force; and second, indirect deaths resulting from restricted access to essential goods and services, such as food and medical care, due to the conflict. However, only data sources to measure direct deaths have so far been identified and thus this article will only consider direct deaths. It is important to underline that the U.N. definition of conflict mentioned above excludes violence by an organized group that targets civilians and therefore does not include terrorism. Instead, deaths as a result of terrorist activities are included in SDG 16.1.1, because the U.N. measure of intentional homicides is based on the International Classification of Crime for Statistical Purposes (ICCS).[2] Thus, terrorism deaths should in principle be accounted for if the U.N. homicide statistics are used for tracking progress toward the SDGs. However, the United Nations relies on member states to report homicides, but these reports are difficult to compare, e.g., some countries appear to include deaths from terrorism, while others do not. U.N. homicide numbers are in some cases even lower than the deaths from terrorism, confirming that terrorism deaths are not consistently included in the Criminal Justice data on homicides provided by the member states.[3] It is also of interest to note that, in general, U.N. homicide numbers are lower than the World Health Organization (WHO) estimates of deaths from interpersonal violence, suggesting potential underreporting of deaths from homicide by the U.N. All of this suggests that there are gaps in the definition and data collection efforts by the U.N., which may provide an inaccurate picture of the progress toward the SDGs.

To improve the reporting on the progress of SDG 16 we suggest that SDG 16.1.2 should not exclusively capture the rather restrictive concept of *conflict-related deaths*, but be enlarged to *collective violence*. Even though there is no commonly accepted definition of *collective violence*, we base our following analysis on the definition of the WHO—restricting the concept of *collective violence* to "the instrumental use of violence by people who identify themselves as members of a group against another group or set of individuals, in order to achieve political, economic or social objectives" (WHO, 2002) and, thus, include conflict-related deaths as well as deaths due to terrorism.

In the remainder of this section, we turn to social science data projects that define and collect data on the different forms of collective violence in a systematic manner. We will discuss the available data on armed conflicts between states, within states, between groups and on organized groups that target civilians, including terrorism.

The Uppsala Conflict Data Program (UCDP) provides detailed information on organized armed conflict. UCDP is a large ongoing data collection effort that has become the most commonly used global dataset for research in the social sciences. Within UCDP the Georeferenced Event Dataset (GED) provides detailed information that is also easy to merge with other data. Here an event is defined as "An incident where armed force was used by an organized actor against another organized actor, or against civilians, resulting in at least 1 direct death at a specific location and a

1 https://unstats.un.org/sdgs/metadata/?Text=&Goal=16&Target= accessed 25 May 2022.

2 https://www.unodc.org/unodc/en/data-and-analysis/statistics/iccs.html accessed 25 May 2022.

3 This issue is also mentioned in the methodological annex of UNODC's Global Study on Homicide (UNODC, 2019).

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

specific date" (Högbladh, 2021: 4). When at least one organized actor is the state, UCDP refers to these conflicts as **state-based armed conflicts** (making up the majority of conflicts). Conflicts between armed groups that do not include the state, e.g., conflicts between ethnic or religious groups, are categorized as **non-state conflicts**. When armed groups, including the state, kill civilians this is referred to as **one-sided violence**.[4] More generally armed conflicts are contested incompatibilities, causing a minimum of 25 deaths per year. The main sources of information are global newswire reporting (e.g., Reuters News, Agence France Presse, and Xinhua) but the UCDP team also consults local media as well as reports by intergovernmental and non-governmental organizations. Each conflict is assigned a conflict identifier and the GED records details on each conflict event, including the names of the opposing sides, the location and time of the violent event and a count of how many people were killed. Since the compilation of the death counts from news reports requires some judgement on the reliability of the sources, the GED offers estimates of the highest, lowest, and most reliable ("best") estimates.

The three categories of organized violence, state-based, non-state based and one-sided violence, are exclusive and by design there is no overlap. Thus, adding all of the conflict-related deaths provides information on how many people died as a result of direct violence in organized conflicts during a specific period in a particular region. A comparison with alternative sources suggests that the GED conflict death data are conservative. For the year 2015 the WHO estimates that about 186,400 people died as a result of collective violence WHO (2022). This compares to an estimate of about 147,200 from GED based on their "high" death counts (see Table A3 in the Appendix). Country comparisons also suggest that GED numbers are much lower than the estimates from public health studies. As an example, take the careful study of the armed conflict in South Sudan conducted by a team of epidemiologists at the London School of Hygiene and Tropical Medicine. Checchi *et al*. (2018) estimate that about 190,000 people died as the result of direct violence between 2014 and 2018, compared to only 12,200 according to the GED. These differences between social science and public health estimates are due to variations in definitions as well as collection methods.

Fatality data collected by social scientists, such as UCDP, rely on media reports as major information sources, and it is likely that these data suffer from a downward bias. Even though media-reported information on fatalities is easily accessible in large quantities, it is often not complete. Such databases miss events which are not reported, mostly due to deliberate selection or inaccessibility of information on certain incidents. One of the few studies to systematically investigate this under-reporting bias is Weidmann (2016). He compares detailed military data on violent events with GED entries from Afghanistan and his results clearly indicate that events in areas with poor mobile phone coverage are less likely to be reported by GED. Furthermore, incidents with high numbers of casualties among coalition soldiers in accessible places increased the reporting probability. Thus, statistical analysis aiming to explain collective violence measured by such event datasets may risk biased results if the systematic measurement error is associated with the independent variable. More importantly for our work, merely tracking the progress of SDG 16 based on one of such event databases might lead to biased reporting.

Given these known shortcomings of the UCDP data, should we use public health data to estimate collective violence-related deaths? Although the WHO provides global and regional data on deaths due to collective violence, these data are not available by country-year over a longer period. There is also little information on how these estimates are derived as the WHO does not provide the estimation method. Alternatively, one could try and collect country or conflict specific studies that have estimated the number of excess deaths. Although this is an active area of research,[5] there is no agreed upon methodology and in many studies it remains unclear how many of the excess deaths are due to the direct impact of violence and how many died due to indirect factors such as hunger and disease.

4 Pettersson et al. (2021); Sundberg and Melander (2013).
5 For example: Burnham *et al*. (2006); Coghlan *et al*. (2006); Crawford (2015); IPPNW (2015); Obermeyer *et al*. (2008).

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

Some of the public health studies have also been criticized for overstating the number of victims.[6] These methodological issues make it impossible to add data from different country studies.

Thus, in contrast to the public health studies, UCDP provides data that have been collected by using the same method across all countries, but we acknowledge that this data collection effort suffers from downward bias. One way to address this shortcoming could be to augment the UCDP data collection effort with information from other global event databases covering death estimates due to collective violence. One such data collection is the Global Terrorism Database (GTD), a widely used database for the study of terrorism. For the purpose of the GTD, terrorism is defined as "the threatened or actual use of illegal force and violence by a non-state actor to attain a political, economic, religious, or social goal through fear, coercion, or intimidation".[7] The violence must be intentional, but if it is exclusively used to pursue financial gain it is excluded from the database. In principle, we should be able to add the deaths due to terrorism to the UCDP numbers of conflict deaths because for an event to be included in the GTD "the action must be outside the context of legitimate warfare activities".[8] In Figure 1 we present the data from the two different data sources (2000–2018), the dotted line shows the number of deaths due to organized violence (GED) and the solid line the number of terrorism deaths (GTD). Until 2011 both counts were relatively low, but then increased until 2014. The increase in organized conflict deaths is due to the war in Syria and the increase in terrorism deaths is driven by the events in Iraq. Moreover, both data series are characterized by a very skewed global distribution of the number of deaths. Five countries, Syria, Iraq, Afghanistan, Nigeria, and Ukraine, account for almost 79 percent of all the GED deaths in 2014. The terrorism numbers are dominated by Iraq, Nigeria, Afghanistan, Syria, and Pakistan—these five countries account for almost 74 percent of all global terrorism deaths. Since 2014 both data series, GED and GTD, have been decreasing. Since these trends are similar, it raises the suspicion that the two data series may not measure entirely separate phenomena.

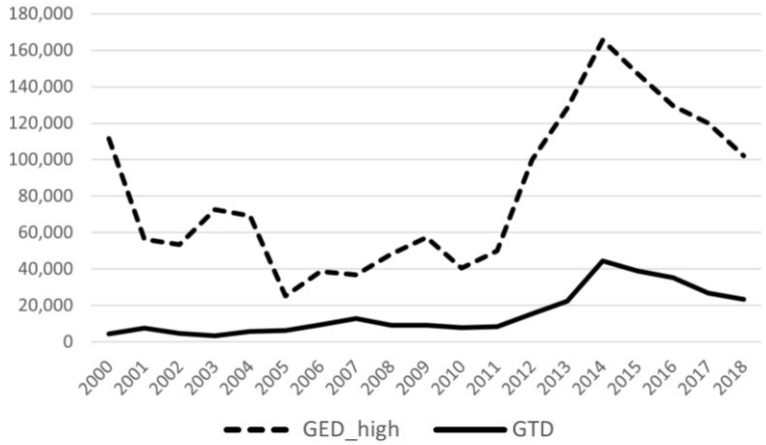

Figure 1: Global number of deaths reported by GED and GTD
Note: For the GED data, the highest reliable estimate of deaths ("high") is used. Zero GED deaths for all countries in "the West" are assumed (for further discussion see the "Methods" section).

Another look at the data also supports the suspicion that the UCDP and the GTD concepts may not be mutually exclusive. All of the top terrorism countries during 2000–2018 are countries that also experienced large-scale armed conflicts during the period: Iraq, Afghanistan, Nigeria, Syria, Pakistan, Somalia, and Yemen. Scholars of civil war have long noted that terrorism is a common tactic in armed conflicts, for example Fortna (2015) suggests that almost one quarter of all insurgency groups use high casualty terrorist tactics in civil war. In this sense, terrorism is not understood as an ideology but as a tactical choice.[9] In contrast, other scholars dispute that terrorism can be defined as a distinct phenomenon as many state and non-state organizations frequently use terrorism alongside other tactics.[10]

6 For further discussion see Johnson *et al*. (2008) and Spagat *et al*. (2009).
7 University of Maryland (2019: 2).
8 University of Maryland (2019: 11).
9 Kis-Katos *et al*. (2014).
10 Tilly (2004).

© EPS Publishing, 2022.
All rights reserved
For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

Civil war and terrorism are difficult to distinguish, particularly in their early stages. At the start of an armed conflict small groups operate in a clandestine fashion, similar to terrorist cells. Thus, terrorism can also be described as a proto-civil war.[11] These theoretical problems in distinguishing insurgencies from terrorism highlight the difficulties in developing separate measures for deaths resulting from organized conflicts and those from terrorism.[12]

For theoretical and data reasons we cannot improve on the counts of collective violence deaths by simply adding the GTD terrorism deaths, because it would result in overcounting victims. In a theoretical contribution, Sambanis argues that although many insurgents use terrorist tactics in civil war, it may be useful to distinguish terrorism *outside* of civil war from terrorism *within* civil war. He uses the expression "pure terrorism" when he refers to terrorism outside of civil war. Based on this concept, we want to augment the GED estimates by the "pure terrorism" counts. We therefore have to identify the deaths that are listed in the GED as well as in the GTD, but then add terrorism victims that are only listed in the GTD to the GED to obtain an estimate of collective violence deaths. We now turn to the discussion of two different methods that enable us to identify "pure terrorism" and thus avoid double counting.

Methods

Having established that the definitions of organized conflict and terrorism are not mutually exclusive and that some events are included in both the GED and the GTD datasets, we investigate how we can identify the overlap of conflict and terrorism events and the associated deaths. Currently, there are two efforts to systematically compare the two databases. In this section, we describe both approaches and how we can use them for filtering out duplicate entries from conflict and terrorist event databases.

The first method we want to introduce is called the *Matching Event Data by Location, Time, and Type* (MELTT) developed by Donnay *et al*. (2019) for integrating data from different violent event datasets. MELTT's aim is to identify entries in those conflict event datasets that probably refer to exactly the same event through iterative pairwise comparison. The main challenge with this comparison is that the same event may be coded differently depending on the source and differing internal coding practices. Hence, one has to allow for some variation in the measurements.

Donnay *et al*.'s protocol enables researchers to apply this *imperfect matching* technique in a systematic manner. The first parameter users choose is the spatio-temporal window in which entries refer to the same event. Choices can range from same day or preceding/following day(s) and as for location from zero to one or more kilometers apart. This is necessary since event databases do not always record time and place measurements with precision. However, if different entries do refer to the same event, it is likely that they occur within a narrow spatio-temporal window. In a second step, the protocol compares other attributes to distinguish unique from matching events within these windows—such as the type of the events or the actors. If users want to allow for "fuzzier" matches, they can choose from a taxonomy with multiple levels. Unlike the spatio-temporal decisions, the mapping of equivalent categories can be very labor intensive depending on which level of detail the researchers set for their taxonomy. If there are several potential matches, the algorithm decides on the one which is the most similar.

For the present study, Donnay *et al*. kindly provided us with an integrated dataset (the combined data where duplicate events have been filtered out). Specifically, in the case of events co-occurring in both databases, the GTD event was filtered out and the GED event retained in the integrated dataset. Therefore, GTD events remaining in the integrated data can be interpreted as terrorism occurring outside of armed conflict, or "pure terrorism". Data are available for all African countries from 1997 until 2016 and we allowed for 5km spatial and one day temporal "fuzziness". Furthermore, for this integration, the taxonomies created by Donnay *et al*., were on the type of event and the actors involved, as well as the degree of geo-precision. Thus, for all the African countries for the years from 2000

11 Sambanis, (2008).
12 See Hoeffler (2022) for more discussion on the interrelationship between armed conflict and terrorism.

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

to 2016 we have estimates based on the application of MELTT. To estimate global data for 2000-2018, we extrapolated the information on Africa onto the entire world. From the integrated data, we were able to calculate each of the included country-year's ratios of "pure terrorism" to total GTD deaths and we used these ratios to create estimates for the missing country-years. To obtain estimates for the years 2017–2018 for Africa, we applied each African country's ratio from 2016 to the GTD data for those two years.[13] For the rest of the world, we extrapolated in a rather crude manner. We simply applied the mean ratio of "pure terrorism" to total GTD deaths to all of the countries outside the region that had an ongoing conflict according to the GED.

The second data matching effort is the *Terrorism in Armed Conflict* (TAC) project, which approaches the integration of GED and GTD from a different angle. With the motivation to find out whether rebel organizations use terrorist tactics, Fortna *et al*. (2022) created the TAC database, which matches perpetrators of GTD events with rebel organizations listed by the GED. They do not only consider perfect matches, but they systematically tackle the issue of varying precision regarding the perpetrators of terrorism events. In a large coding effort, they looked in detail at over 9,000 GTD events possibly linked to a UCDP-listed rebel group. It is a unique feature of TAC to allow researchers to include groups that are fractions, umbrellas, or affiliates of the UCDP rebel organizations as well as generic descriptors and unknown links. Accordingly, the TAC project provides different matching levels which users can choose from. For our estimates we included all GTD events where (1) the perpetrators are *connected* in some way with a UCDP rebel group or (2) where generic descriptors are used. For example, the GED lists the Kurdistan Workers' Party (PKK) as one conflict side while events in GTD list groups like Kurdish separatists, Kurdish rebels, and Kurdish militants. We want to match these groups although they are not named as the PKK, since we consider events perpetrated by groups connected to UCDP rebel organizations as part of the armed conflict. In the TAC parlance, we applied level E (Fortna *et al*., 2022: 220) for the classification. However, we do not include events that list "unknown perpetrators" (such as gunmen or individuals, listed as level F) from events that took place in a country during a time period where it could possibly be linked to locally operating rebel organizations. The justification is that we consider that "pure terrorism" can occur even in countries currently in armed conflict.

Using the TAC methodology, the overlap between GED and GTD is defined as events being perpetrated by the same actors. Thus, we removed GTD events perpetrated by a rebel organization listed by the UCDP, or a group connected to one of those, in order to create a measure for "pure terrorism". One of the advantages of TAC is that coverage is global, however it only covers years up to 2013. Therefore, to extrapolate our measure of "pure terrorism" fatalities to more years, we applied each country's ratio of fatalities in GTD which are linked to a rebel organization from 2013 to all the country-years thereafter.

How good are these methods in recognizing duplicates of violent events and the associated deaths? The main issue with the MELTT methodology arises when the encoding of what actually is the same event differs too much between the two databases. In these cases, the algorithm will not recognize events as duplicates. The TAC methodology, on the other hand, relies on identifying actors and does not rely on exact information on time and place of an event. Using TAC will result in identifying more duplicate events, because MELTT requires information on the time and place of an event. If these are stated imprecisely, MELTT will not recognize these as duplicate events. However, some events may be erroneously identified as duplicates by TAC. If the violence was committed by an affiliated actor listed in the GED, the associated events from GTD will be filtered out even though the *event* might not have been contained in GED. Filtering them out would therefore result in losing this event and its associated fatalities. To summarize, with MELTT one can be more certain that what is filtered out are actually the same events. Using TAC,

[13] To apply Donnay *et al*.'s taxonomies to the latest data on Africa, we would have to check whether they still fit the data and possibly adapt and extend them. While we were unable to do so within the scope of this article, it is worthwhile to tackle this in future research.

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

it could happen that events committed by a UCDP (related) actor are removed even though they have not been a duplicate but only been contained in the GTD.

Neither the MELTT nor the TAC project provides a matched or integrated dataset with global coverage for the years 2000–2018. MELTT covers only Africa until 2016, whereas TAC covers all countries but only until 2013. Extending estimates of collective violence deaths to achieve global and up-to-date coverage is problematic with the currently available data and taxonomies. So far, we have focused on the countries that experience armed conflicts as well as terrorism. For the many countries that have no ongoing armed conflict it is straightforward to just use the GTD death estimates. For some countries, the use of the GED as a basis may be problematic. Take the example of 9/11. Consistent with public perception, all four events that occurred on this day have been categorized as terrorism by GTD. However, GED classifies the attacks against the World Trade Center as one-sided violence, while the plane crashes in Pennsylvania and Virginia are considered as state-based conflict. The rationale behind this classification is that the GED categorizes events based on the (intended) targets of the attack. Hence, the attacks on the Pentagon and the White House indicates that "the state" was targeted—consequently the GED classifies these events as "state-based armed conflict". Thus, the United States is listed as a conflict country for 2021 in the GED.

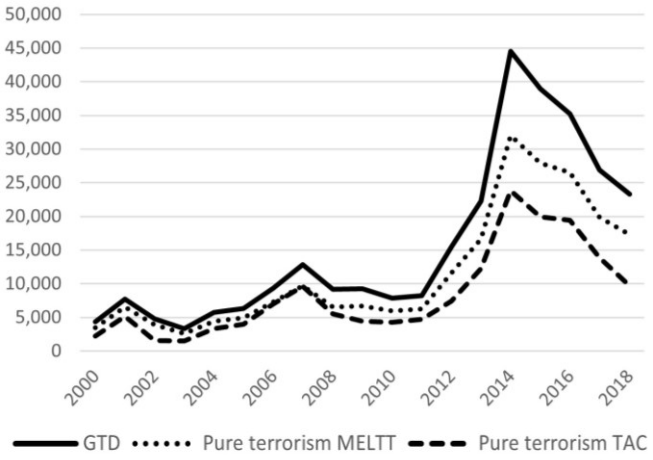

Figure 2: Global estimates of deaths due to "pure terrorism"

Given that some of these categorizations are contested, we decided to assume that no country in "the West" was a conflict country and use the GTD to estimate collective violence for these countries. Apart from fitting in with the common understanding of the type of collective violence in "the West", it has the added advantage that the GTD lists many more events. Since there is no minimum death threshold for events to be listed in the GTD the death toll in "the West" is higher than in GED and addresses somewhat the downward bias in the GED data.

To summarize, for countries with no organized conflicts we use the GTD to estimate deaths from collective violence. For all countries in "the West", we assumed that they were not experiencing organized conflict and use the GTD to estimate deaths from collective violence. For all other countries, i.e., those that experienced organized conflict as well as terrorism, we apply information from TAC and MELTT to estimate conflict-related deaths and "pure terrorism". The sum of conflict-related deaths and "pure terrorism" can be interpreted as a measure of deaths resulting from collective violence. Our current estimates are quite crude—to apply these two methods in deriving global estimates for 2000–2018, we either must assume that the world is like Africa (because MELTT only covers Africa) or that the world is still like it was in 2013 (since TAC only covers 2000–2013).

Estimates

Applying the two estimation methods, MELTT and TAC, we start our analysis by deriving two estimates of "pure terrorism" (terrorism outside of armed conflict). Figure 2 provides three time series: GTD; "pure terrorism" from our application of MELTT; and a "pure terrorism" estimate based on TAC. By construction our "pure terrorism" estimates are always lower than GTD death counts, and the MELTT estimates of "pure terrorism" are always between the GTD and TAC estimates. In the early 2000s it is difficult to distinguish the lines, i.e., with our methods we identify only very few events that are in the GED as well as in the GTD databases. This changes over time and around 2011 there

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

is a considerable difference between GTD and the "pure terrorism" estimates. All the terrorism estimates peak in 2014 and for this year the difference is very pronounced. Our global MELTT estimate is about 28 percent lower than the GTD count. The country with the largest difference is Nigeria, as about 45 percent of all Nigerian deaths in the GTD are also in the GED. The total difference is almost 3,500 deaths. The global TAC estimate is even lower. It is about 46 percent lower than the GTD counts, again there is a particularly large discrepancy for Nigeria, where about 79 percent of the deaths are included in the GED and the GTD.[14]

Table 1 presents an overview of the data used and our estimates of deaths due to "pure terrorism" and collective violence. In column 1 we present the sum of all deaths listed in the GED for 2000–2018 by region (see Appendix tables A1 and A2 for regional classification). The last row provides the global sum. According to the GED over 1.5 million people died as a result of direct violence in organized conflicts during this period. This is equivalent to almost the number of inhabitants of Philadelphia or the entire population of Equatorial Guinea. The Middle East and North Africa (MENA) was the most violent region, accounting for more than one third of all global deaths, followed by Sub-Saharan Africa (SSA) with almost half a million deaths. The terrorism figures as per the GTD are presented in column 2, the global total for this period is just under 300,000. Terrorism was also most prevalent in the Middle East and North Africa, accounting for about 42 percent of all terrorism deaths in the GTD. The region least suffering from terrorism is Latin America and the Caribbean, followed by "the West" with a death count of about 4,400. While we assume zero for the GED figure for "the West" (due to GED drawbacks described in the Methods section), it is of note that GTD number is higher than the "raw" GED figure of 3,653. Column 3 lists the "pure terrorism" estimated using the MELTT method and Column 4 the "pure terrorism" estimates based on the TAC method. As discussed above, the total MELTT estimates are higher than the TAC numbers.

Table 1: Total fatality estimates for the years 2000–2018

Interventions	*GED (1)*	*GTD (2)*	*Pure terror MELTT (3)*	*Pure terror TAC (4)*	*Collective Violence MELTT (5)*	*Collective Violence TAC (6)*
The West	0	4,371	4,371	4,371	4,371	4,371
Eastern Europe	31,159	6,089	4,725	4,502	35,884	35,661
Latin America and the Caribbean	81,621	3,672	2,882	1,802	84,503	83,423
Asia	387,072	95,278	73,359	43,223	460,431	430,295
North Africa and the Middle East	585,080	123,078	95,180	86,805	680,260	671,885
Sub-Saharan Africa	466,930	63,109	39,460	18,936	506,390	485,866
World	1,551,862	295,597	219,977	159,639	1,771,839	1,711,501

Notes: The GED estimates for "the West" are 3,653. As stated in the Methods section we assume zero GED deaths for "the West". Columns 5 and 6 provide estimates for collective violence, column 5 is the sum of columns 1 and 3, column 6 is the sum of columns 1 and 4.

[14] Here we refer to the 2013 data since the TAC project does not span 2014.

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

In column 5 and 6 we present estimates of collective violence deaths, where the estimates use MELTT and TAC estimates, respectively. We estimate that the number of people who died as a result of armed conflict and terrorism to be between 1.71 and 1.77 million people. Note that, given the magnitude of these numbers, our assumptions regarding the difference of death estimates between the GED and the GTD for "the West", an overwhelmingly peaceful and secure region, makes very little difference for the total estimates of collective violence.

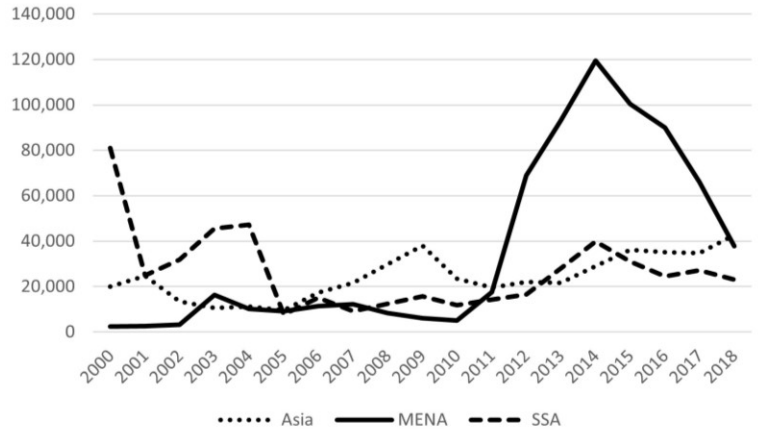

Figure 3: Integrated regional annual fatality estimates due to collective violence based on TAC

In Figure 3 we investigate the time series of our collective violence death estimates for the three most violent regions: the Middle East and North Africa, Asia, and Sub-Saharan Africa. Here we use the MELTT estimates (the TAC estimates are qualitatively similar). In the early 2000s Sub-Saharan Africa had relatively high death per annum counts, with over 80,000 deaths per year. These numbers have declined to about half in 2014 and have further declined toward the end of the period. For Asia, the millennium started with relatively low numbers of about 20,000,however, by 2018 this had doubled. Up until 2011 the Middle East and North Africa had mostly lower death counts than the two other regions, however, in 2014 collective violence killed almost 120,000 people. These figures had come down by 2018 but the Middle East and North Africa remains a very violent region.

Conclusion

To assess progress for SDG 16.1, the United Nations has suggested the measure of conflict-related deaths. We argue that the focus on conflict as state based armed conflicts, or wars, results in an undercounting of violent deaths. Instead, we suggest also considering deaths from other forms of collective violence, such as one-sided violence and terrorism. This type of collective violence is not mentioned in the U.N. targets but exploiting existing data sources could help to provide a more accurate number of deaths caused by armed organized groups.

In this article we discuss the available data sources and suggest that the commonly used social science data provided by UCDP suffer from underreporting bias. Alternative public health data tend to provide higher death counts, but the lack of a common methodological approach make it impossible to add up counts from different countries. The WHO provides data on victims of collective violence, but they are not provided for every year and there is a lack of information on the model on which the estimates are based. It is instructive to compare the numbers for 2015, because we have GED, GTD, and WHO data, as well as estimates from MELTT; additionally, the last year (2013) of the TAC project can still serve as a useful benchmark. For 2015 the WHO estimates about 186,400 collective violence deaths for this year. This is considerably higher than the high estimates from the GED, at about 147,200 deaths. We suggest augmenting the GED with information from other global event databases recording fatalities of collective violence and use the terrorism deaths from the GTD. However, augmenting does not simply entail adding terrorism deaths to conflict deaths, since armed conflict and terrorism are difficult to distinguish both theoretically and in data collection. We use two recently developed methods to identify deaths from "pure terrorism", i.e., deaths that occurred due to terrorism outside of organized armed conflicts. Donnay *et al*. (2019) suggests a

© EPS Publishing, 2022.
All rights reserved
For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

comparison of the individual events—resulting in an estimate of about 175,000 deaths due to organized conflict and terrorism worldwide in 2015. The method by Fortna *et al*. (2022) compares the violence committed by insurgent and terrorist groups—this provides an estimate of about 167,000 deaths. Note that although both estimates are higher by design than the armed conflict death counts (GED), they are still lower than the public health counts (WHO), thus our estimates fall between a plausible upper and lower bound.

As discussed, our estimates have to rely on a number of crude assumptions, and we see our study as a first suggestion of how deaths from collective violence may be quantified. One possible extension of our work is to enlarge the actor taxonomy of the MELTT protocol to cover countries outside of Africa. A further option is to extend the existing TAC project beyond 2013. A third extension is to consider a combination of the two methods, MELTT and TAC, by deriving improved estimates through developing an actor taxonomy usable for MELTT from the TAC. Such methodological advancements have a number of implications for research. Distinguishing "pure terrorism" events from terrorism within organized armed conflict will improve our understanding of terrorism itself, a concept difficult to define and measure. The identification of actors present in both data collections (GED and GTD), will also enable further research of the use of terrorist tactics in armed conflicts. Similarly, the sensible integration of these two databases will benefit violence-research, particularly the research of phenomena that are not fully covered by either database, like the targeting of civilians.

In addition, there are a number of closely related questions that open new avenues for research. Collective violence not only kills but also maims.[15] However, there is currently no systematic effort to estimate the number of injuries due to collective violence. An investigation of the number of injured due to organized violence and terrorism would help us to capture the burden of collective violence more fully. In addition, organized violence not only kills people directly through the use of force, but also through malnutrition and disease. There are a number of efforts in the public health literature to estimate the excess death rates due to organized conflict.[16] Global estimates suggest that about 1.8 additional people die due to malnutrition and disease per one direct GED death, most of them are children under the age of five.[17] Based on these recent studies, collaborations between social scientists and public health experts appear promising in establishing more defensible estimates of the human cost of organized violence by including the deaths and injuries from direct violence plus the health implications for the conflict affected populations. These estimates would help the research community and the United Nations to better assess whether we are making progress on the 2030 Sustainable Development Agenda.

Acknowledgements

This research was funded by the Alexander von Humboldt Foundation. We would like to thank Karsten Donnay for providing data and his support with MELTT.

References

Bendavid, Eran, Ties Boerma, Nadia Akseer, Ana Langer, Espoir Bwenge Malembaka, Emelda A Okiro, Paul H Wise, Sam Heft-Neal, Robert E Black, Zulfiqar A Bhutta, et al. 2021. “The Effects of Armed Conflict on the Health of Women and Children”. *The Lancet* 397(10273), pp.522–532. https://doi.org/10.1016/S0140-6736(21)00131-8

Burnham, Gilbert, Riyadh Lafta, Shannon Doocy, and Les Roberts. 2006. “Mortality after the 2003 Invasion of Iraq: A Cross-Sectional Cluster Sample Survey”. *The Lancet 368*(9545), pp.1421–1428. https://doi.org/10.1016/S0140-6736(06)69491-9

15 Ghobarah et al. (2003).
16 For example, see Bendavid *et al*. (2021) and Wise *et al*. (2021).
17 Jawad *et al*. (2020).

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

Checchi, Francesco, Adrienne Testa, Abdihamid Warsame, Le Quach, and Rachel Burns. 2018. *Estimates of Crisis-Attributable Mortality in South Sudan: A Statistical Analysis*. Report. Department of Infectious Disease Epidemiology, Faculty of Epidemiology, Population Health, London School of Hygiene, and Tropical Medicine.

Coghlan, Benjamin, Richard J Brennan, Pascal Ngoy, David Dofara, Brad Otto, Mark Clements, and Tony Stewart. 2006. "Mortality in the Democratic Republic of Congo: A Nationwide Survey". *The Lancet 367*(9504), pp.44–51. https://doi.org/10.1016/S0140-6736(06)67923-3

Crawford, Neta C. 2015. *War-related Death, Injury, and Displacement in Afghanistan and Pakistan 2001-2014*. Watson Institute for International Studies, Brown University.

Donnay, Karsten, Eric T Dunford, Erin C McGrath, David Backer, and David E Cunningham. 2019. "Integrating Conflict Event Data". *Journal of Conflict Resolution, 63.5*, pp.1337–1364. https://doi.org/10.1177/0022002718777050

Fortna, Virginia P. 2015. "Do Terrorists Win? Rebels' Use of Terrorism and Civil War Outcomes". *International Organization,* 69(3), pp.519–556. https://doi.org/10.1017/S0020818315000089

Fortna, Virginia P, Nicholas J Lotito, and Michael A Rubin. 2022. "Terrorism in Armed Conflict: new data attributing terrorism to rebel organizations". *Conflict Management and Peace Science 39*(2), pp.214–236. https://doi.org/10.1177/0738894220972996

Ghobarah, Hazem Adam, Paul Huth, and Bruce Russett. 2003. "Civil Wars Kill and Maim People—Long after the Shooting Stops". *American Political Science Review, 97*(2), pp.189–202. https://doi.org/10.1017/S0003055403000613

Hoeffler, Anke. 2022. "Counting Every Body - Determining the Global Burden of Terrorism and Armed Conflict". *The Handbook of the Economics of Terrorism*. Ed. by Atin Basuchoudhary and Günther Schulze. Cambridge: Cambridge University Press. Chap. 11.

Högbladh, Stina. 2021. *UCDP GED Codebook version 21.1*.

International Physicians for the Prevention of Nuclear War. 2015. *Body Count - Casualty Figures after 10 Years of the "War on Terror!" Iraq, Afghanistan, Pakistan*. https://www.ippnw.de/frieden/konflikte-kriege/ body-count.html (Accessed 25 May 2022).

Jawad, Mohammed, Thomas Hone, Eszter P Vamos, Paul Roderick, Richard Sullivan, and Christopher Millett. 2020. "Estimating Indirect Mortality Impacts of Armed conflict in Civilian Populations: Panel Regression Analyses of 193 Countries, 1990–2017". *BMC medicine, 18*(1), pp.1–11. https://doi.org/10.1186/s12916-020-01708-5

Johnson, Neil F, Michael Spagat, Sean Gourley, Jukka-Pekka Onnela, and Gesine Reinert. 2008. "Bias in Epidemiological Studies of Conflict Mortality". *Journal of Peace Research, 45*(5), pp.653–663. https://doi.org/10.1177/0022343308094325

Kis-Katos, Krisztina, Helge Liebert, and Günther G Schulze. 2014. "On the Heterogeneity of Terror". *European Economic Review, 68*, pp.116–136. https://doi.org/10.1016/j.euroecorev.2014.02.009

Obermeyer, Ziad, Christopher JL Murray, and Emmanuela Gakidou. 2008. "Fifty Years of Violent War Deaths from Vietnam to Bosnia: Analysis of Data from the World Health Survey Programme". *Bmj* 336.7659, pp.1482–1486. https://doi.org/10.1136/bmj.a137

Pettersson, Therése, Shawn Davies, Amber Deniz, Garoun Engström, Nanar Hawach, Stina Högbladh, Margareta Sollenberg, and Magnus Öberg. 2021. "Organized Violence 1989–2020, with a Special Emphasis on Syria". *Journal of Peace Research, 58*(4), pp.809–825. https://doi.org/10.1177/00223433211026126

Sambanis, Nicholas. 2008. "Terrorism and Civil War". *Terrorism, Economic Development, and Political Openness*. Ed. by Philip Keefer and Norman Loayza. Cambridge University Press, Cambridge and New York, pp.174–206. https://doi.org/10.1017/CBO9780511754388.007

Spagat, Michael, Andrew Mack, Tara Cooper, and Joakim Kreutz. 2009. "Estimating War Deaths: An Arena of Contestation". *Journal of Conflict Resolution, 53*(6), pp.934–950.

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

https://doi.org/10.1177/0022002709346253
Sundberg, Ralph and Erik Melander. 2013. "Introducing the UCDP Georeferenced Event Dataset". *Journal of Peace Research*, *50.4*, pp.523–532. https://doi.org/10.1177/0022343313484347
Tilly, Charles. 2004. "Terror, Terrorism, Terrorists". *Sociological Theory, 22*(1), pp.5–13. https://doi.org/10.1111/j.1467-9558.2004.00200.x
United Nations. 2015. *The Millennium Development Goals Report 2015*.
United Nations Office on Drugs and Crime. 2019. *Global Study on Homicide 2019: Trends, Contexts, Data*. UNODC.
University of Maryland. 2019. *Global Terrorism Database (GTD) - Codebook: Inclusion Criteria and Variables*. https://www.start.umd.edu/gtd (Accessed 01 May 2022.).
Weidmann, Nils B. 2016. "A Closer Look at Reporting Bias in Conflict Event Data". *American Journal of Political Science*, *60*(1), pp.206–218. https://doi.org/10.1111/ajps.12196
Wise, P.H., Shiel, A., Southard, N., Bendavid, E., Welsh, J., Stedman, S., Fazal, T., Felbab-Brown, V., Polatty, D., Waldman, R.J. and Spiegel, P.B., 2021. The political and security dimensions of the humanitarian health response to violent conflict. The Lancet, *397*(10273), pp.511-521. https://doi.org/10.1016/S0140-6736(21)00130-6
[WHO] World Health Organization. 2002. *World Report on Violence and Health: Summary*. http://apps.who.int/iris/bitstream/handle/10665/42512/9241545623_eng.pdf;jsessionid=28D0A9E77643FF24BC0AB34E8DA0C522?sequence=1 (Accessed 26 May 2022).
[WHO] World Health Organization. 2022. *Global Health Estimates: Leading Causes of Death*. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates/ghe-leading-causes-of-death (Accessed 01 May 2022).

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

Appendix

Table A1: Countries and regions included in the analysis (part 1)

Middle East and North Africa	*Eastern Europe*	*West*
Cyprus	Poland	United States of America
Morocco	Hungary	Canada
Algeria	Czech Republic	United Kingdom
Tunisia	Slovakia	Ireland
Libya	Albania	Netherlands
Iran	Montenegro	Belgium
Turkey	Macedonia	Luxembourg
Iraq	Croatia	France
Egypt	Serbia	Monaco
Syria	Bosnia and Herzegovina	Liechtenstein
Lebanon	Kosovo	Switzerland
Jordan	Slovenia	Spain
Israel	Bulgaria	Andorra
Saudi Arabia	Moldova	Portugal
Yemen	Romania	Germany
Kuwait	Russia	Austria
Bahrain	Estonia	Italy
Qatar	Latvia	San Marino
United Arab Emirates	Lithuania	Malta
Oman	Ukraine	Greece
Palestine	Belarus	Finland
	Armenia	Sweden
	Georgia	Norway
	Azerbaijan	Denmark
	Turkmenistan	Iceland
	Tajikistan	Australia
	Kyrgyzstan	Greenland
	Uzbekistan	Saint Pierre and Miquelon
	Kazakhstan	Holy See
		New Zealand

© EPS Publishing, 2022.
All rights reserved
For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

Table A2: Countries and regions included in the analysis (part 2)

Sub-Sahara Africa	*Latin America and the Caribbean*	*Asia*
Cape Verde	Colombia	Afghanistan
Sao Tome and Principe	Venezuela	China
Guinea-Bissau	Guyana	Mongolia
Equatorial Guinea	Suriname	Taiwan
Gambia	Ecuador	North Korea
Mali	Peru	South Korea
Senegal	Brazil	Japan
Benin	Bolivia	India
Mauritania	Paraguay	Bhutan
Niger	Chile	Pakistan
Ivory Coast	Argentina	Bangladesh
Guinea	Uruguay	Myanmar
Burkina Faso	Bahamas	Sri Lanka
Liberia	Cuba	Maldives
Sierra Leone	Haiti	Nepal
Ghana	Dominican Republic	Thailand
Togo	Jamaica	Cambodia
Cameroon	Trinidad and Tobago	Laos
Nigeria	Barbados	Vietnam
Gabon	Dominica	Malaysia
Central African Republic	Grenada	Singapore
Chad	St. Lucia	Brunei
Republic of the Congo	St. Vincent and the Grenadines	Philippines
Democratic Republic of the Congo	Antigua	Indonesia
Uganda	St. Kitts and Nevis	East Timor
Kenya	Mexico	French Polynesia
Tanzania	Belize	Guam
Burundi	Guatemala	New Caledonia
Rwanda	French Guiana	Papua New Guinea
Somalia	Guadeloupe	Niue
Djibouti	Martinique	Vanuatu
Ethiopia	Montserrat	Solomon Islands
Eritrea	Honduras	Kiribati
Angola	Puerto Rico	Tuvalu
Mozambique	Turks and Caicos Islands	Fiji
Zambia	United States Virgin Islands	Tonga
Zimbabwe	El Salvador	Nauru
Malawi	Nicaragua	Marshall Islands
South Africa	Costa Rica	Palau
Namibia	Panama	Micronesia
Lesotho	Anguilla	Samoa
Botswana	Aruba	Hong Kong
Swaziland	Bermuda	China, Macao SAR
Madagascar	British Virgin Islands	Cook Islands
Comoros	Cayman Islands	
Mauritius		
Seychelles		
Sudan		
South Sudan		
Mayotte		
Reunion		

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org

Vol. 17, No. 2 (2022) | doi:10.15355/epsj.17.2.32

Table A3: Global sum of victims by year

Year	*GED high*	*GTD*	*Pure terror MELTT*	*Pure terror TAC*	*Collective Violence MELTT*	*Collective Violence TAC*	*Collective Violence WHO*
2000	111,477	4,370	3,468	2,214	114,945	113,691	123,834
2001	56,284	7,706	6,514	5,152	62,798	61,436	
2002	53,439	4,795	3,886	1,580	57,325	55,019	
2003	72,618	3,310	2,559	1,520	75,177	74,138	
2004	69,318	5,716	4,441	3,333	73,759	72,651	
2005	25,116	6,342	4,920	3,962	30,036	29,078	
2006	38,634	9,316	7,306	6,979	45,940	45,613	
2007	36,686	12,824	9,695	9,617	46,381	46,303	
2008	48,003	9,157	6,568	5,493	54,571	53,496	
2009	57,280	9,277	6,651	4,412	63,931	61,692	
2010	40,530	7,829	5,968	4,276	46,498	44,806	59,262
2011	49,912	8,246	6,223	4,676	56,135	54,588	
2012	100,088	15,494	11,569	7,397	111,657	107,485	
2013	128,281	22,280	16,612	12,250	144,893	140,531	
2014	165,544	44,524	31,998	23,800	197,542	189,344	
2015	147,022	38,993	27,936	19,933	174,958	166,955	186,375
2016	129,526	35,236	26,494	19,451	156,020	148,977	
2017	120,040	26,892	19,885	13,875	139,925	133,915	
2018	102,064	23,290	17,284	9,719	119,348	111,783	
Total	1,551,862	295,597	219,977	159,639	1,771,839	1,711,501	

© EPS Publishing, 2022. All rights reserved For permissions, email: EPSJManagingEditor@EPSJournal.org