

The Militarized Interstate Events (MIE) Codebook, 1816-2014

This document provides user information on the coding procedures for the Militarized Interstate Event (MIEs), 1816-2014. We include discussions of data formats, amended coding rules, and variable information for the dataset. Please report any errors in the data to the authors.

Date: July 17, 2023

Citation

We ask that users of this dataset cite the following article and note the version number of the data they are using in their study:

- D. M. Gibler and S. V. Miller, "The Militarized Interstate Events (MIE) Dataset, 1816-2014." *Conflict Management and Peace Science*. Forthcoming.

Rationale for the Militarized Interstate Event (MIE) Data

A Militarized Interstate Event (MIE) is a threat, display, or use of force by one state in the Correlates of War (CoW) international system against another state. We faithfully implement the coding rules described in [Jones, Bremer and Singer \(1996\)](#) and the changes to the coding rules made by CoW in the years since. Nevertheless, we also do make several coding rule changes to remove ambiguities in the rules and to provide additional information for researchers.

The changes we make and rationale for them are described well in the manuscript cited above. Of particular note, though, is that we have added an action code labeled "war battle" that describes sustained fights between opposing, organized armed forces. We no longer code "beginning a war" or "join ongoing war" as a militarized action; we only code threats, displays, or uses of force. We code fatality ranges for each event, discarding CoW's use of ordinal fatality scales and the fatality precision variable. This produces minimum and maximum estimated fatalities for each MIE, with no missing values, which can aggregate to meaningful ranges for confrontation-level, confrontation-participant-level, and dyad-year summaries.

The rest of this codebook describes how to access the data and the variables included. We also include several coding notes about wars in our data in the last section.

Data Files Associated with the Militarized Interstate Event (MIE) Data

We provide the following files for users of the MIE data set. We advertise the data of interest in the list below first as the more accessible comma-separated values file (.csv), but we also have the same data available in a Stata data file format (.dta) or an R serialized data frame (.rds). Please note that we are constantly revising these data; users should report the version number of the dataset used in any research. Future releases of the data will come with a text file summarizing changes to the data.

- `mie-[version].csv`: This is the directed-dyad-level MIE data with one case per event.

Militarized Interstate Event (MIE) Data Variables

The MIE data are provided in directed-dyad format. Side A in the event is the first code. The following variables are included in the Militarized Interstate Event (MIE) data, with notes about usage:

- `micnum`: The Militarized Interstate Confrontation (MIC) number for each case. If the event is part of a Militarized Interstate Dispute (MID) originally coded by CoW, we use that number. Note that we have also added numerous confrontation cases that are not in the CoWMID data as events or disputes. These cases begin with `micnum` values in the 9000's to demonstrate completely new cases.

- **eventnum**: This is the individual event number within the confrontation. Note that these numbers are meaningful. An event number value of “1” denotes the first threat, display, or use of force in that particular confrontation. The highest event number is the last *onset* of a threat, display, or use of force that confrontation—previous events may still last longer and cause the duration of the confrontation to extend. A large number of events have missing start and/or end dates. We researched the cases in which the missing-day event could be the first or last participation of the state in the confrontation and confirmed the temporal placement of the missing-day event using a meaningful event number. We have not yet sequentially ordered the missing-day events that could not have been first or last events, but we hope to add that feature soon.
- **ccode1**: This is the CoW country code for the state that issued a threat, display, or use of force targeting ccode2. In clashes, the lower number ccode is ccode1. Since a clash is coded when the initiating state for an attack cannot be determined, these cases are non-directional.
- **ccode2**: This is the CoW country code for the state that was targeted by a threat, display, or use of force. In clashes, the higher number ccode is ccode2.
- **stmon**: The start month of the event, with values ranging from 1 to 12. There are no missing month values; if the month of a militarized action could not be determined, the action was not coded.
- **stday**: The start day of the event, with values ranging from 1 to 31 and missing days reported as values of -9. We have done our best to appeal to the historical record to eliminate missing days in the data, but there are numerous cases where only the month of the militarized action is known.
- **styear**: The start year of the event, with values ranging from 1946 to 2014. There are no missing year values; if the year of a militarized action could not be determined, the action was not coded.
- **endmon**: The end month of the event, with values ranging from 1 to 12. There are no missing end month values.
- **endday**: The end day of the event, with values ranging from 1 to 31 and missing days reported as -9. Again, we have done our best to appeal to the historical record to eliminate missing days in the data, but there are numerous cases where only the end month of the militarized action is known.
- **endyear**: The end year of the event, with values ranging from 1946 to 2014.
- **sidea**: Side A in the confrontation is defined by whichever state initiated the first militarized action in the confrontation. All states coordinating with the original Side A state are also considered members of Side A. The variable is dichotomous, with a “1” indicating Side A and “0” indicating Side B. Reciprocated confrontations will have one or more actions by states with a Side A value of “0”.
- **action**: The type of action ccode1 targeted ccode2 with in the event. Possible values [with *hostlev* in brackets] include the following: 0 No militarized action [1] 1 Threat to use force [2], 2 Threat to blockade [2], 3 Threat to occupy territory [2], 4 Threat to declare war [2], 5 Threat to use CBR weapons [2], 6 Threat to join war, 7 Show of force [3], 8 Alert [3],

9 Nuclear alert [3], 10 Mobilization [3], 11 Fortify border [3], 12 Border violation [3], 13 Blockade [4], 14 Occupation of territory [4], 15 Seizure [4], 16 Attack [4], 17 Clash [4], 19 Use of CBR weapons [4], and 22 War Battle [5]. Note that this scale is actually not ordinal. See above for a discussion of how we changed how wars are coded, and see the last section for additional notes on wars and battles.

- `hostlev`: The hostility level of the event. Possible values include the following: 1 No militarized action, 2 Threat to use force, 3 Display of force, 4 Use of force, and 5 War.
- `fatalmin1`: This is an estimate of the minimum number of military battle-deaths for `ccode1` in the event.
- `fatalmax1`: This is an estimate of the maximum number of military battle-deaths for `ccode1` in the event.
- `fatalmin2`: This is an estimate of the minimum number of military battle-deaths for `ccode2` in the event.
- `fatalmax2`: This is an estimate of the maximum number of military battle-deaths for `ccode2` in the event.
- `version`: The MIE version number. We are constantly reviewing and updating our datasets, and we ask users of our data to always report which version number of the data they are using in their research.

General Notes

Below, we clarify various characteristics of the dataset that may sometimes be misunderstood, leading to misuse of the data.

- Each MIE is dyadic and directed, with the action describing what `ccode1` initiated against `ccode2`. There are two exceptions, however. The original CoW coders considered clashes (action code of 17) to be the same as attacks (action code of 16) except the initiator of the event could not be determined by the coder. Similarly, though war battles are often planned and initiated by one force in the dyad, which force actually fires the first shot in the event is not recorded. Therefore, all war battles (action code of 22) are also non-directed in the data.
- We include the hostility level of each action in the data. However, there is ample heterogeneity within each hostility level value. For example, threatening force and threatening war imply much different levels of hostility.¹ More nuanced variability rests with the use of force hostility level (a code of 4). We have pointed out elsewhere that declarations of war (an action code of 18) are political events, not militarized, and followed by no subsequent militarized events for the majority of events. These are not uses of force, which is why we have separated these cases into a different dataset.
- There is also heterogeneity within the same action value. For example, many attacks are countered effectively, with outstanding force. The attacker may often lose the event. Other

¹Even threatening war has a substantial level of heterogeneity as some autocrats have threatened war even though most observers doubted the sincerity of the threat or the ability of the state to carry out the threat.

times, an attack is not responded to—submarine attacks, for example—and lopsided fatalities occur.

- The action numbering is not ordinal. In retrospect, some threats to use force contained much greater potential for escalated conflict than attacks and clashes. Threats and blockades during, for example, the Cuban Missile Crisis had the potential to escalate to global war; meanwhile, we have coded clashes between two forces over goats and mules that have crossed their border.
- Relatedly, selecting on fatalities is not a good way of differentiating the conflicts with some likelihood of escalating. The Cuban Missile Crisis, a confrontation without fatalities, is a good example of a case that would be excluded from an analysis of conflict escalation/severity based on this decision.
- We do not include a confrontation-level reciprocation variable. It is possible to generate one based on the events data though we caution that doing this would result in a variable with some strange properties. Reciprocation is defined as positive when Side B in a confrontation initiates a militarized event. However, clashes (and war battles) are, by definition, events in which we as coders could not ascertain which side fired first and are numbered by the lower code as the initiator of the event. Clashes and war battles are of course “reciprocated” events, but the coding cannot denote this. For example, the La Plata War of 1851-1852 between Brazil and Argentina includes three non-battle events initiated by Brazil and six war battles. Since the ccode of Brazil is 140 and Argentina’s ccode is 160, the entire war is coded as not reciprocated due to coding rules. This is just another reason why users should be extremely careful when inferring conflict information from a reciprocation variable like one might get from CoWMID.

Notes on War Battles

- War battles (action code of 22) most often, but not always take place during what various scholars have classified as wars. We remain agnostic on war definitions but implicitly include the CoW international war coding with our MICNames data (CoW wars include the CoW war title and no parentheses in the data). Again, war battles are defined as sustained fights between opposing, organized armed forces, and these can occur in confrontations that have fewer than 1,000 battle deaths. Our treatment of war battles allows researchers to easily use our data in their research design, regardless of their definition of war.
- We based the war battles data on [Min’s 2021 Interstate War Battle Dataset \(IWBD\)](#), with some notable changes. First, we changed dates (mostly start and end days) in approximately 10-15% of the cases. We code dates according to when active conflict began and ended. Second, the IWBD includes both campaigns and individual battles within those campaigns. For example, the Battle of the Somme and the Battle of Albert, which was part of the larger battle, are included in the IWBD. Whenever possible we coded fatalities for the individual battles and include only those in the MIE dataset. Third, we added approximately 100 battles that were not in the IWBD. We used the IWBD battle names for all cases from that data, and we added conventional battle names for the events we added.
- For multilateral battles, we did our best to ascertain direct dyadic fatalities. Thus, for example, if Romanian forces and German forces were fighting the same battle against the

Soviets during World War II, we consulted available sources to determine where in the battle the Romanians and Germans were and Soviet fatalities in those area of the battle. Since that data is often not available, or two different forces fought together in unified commands, we relied on overall kill rates in the battle to determine dyadic fatalities. In some cases we augmented the kill rate data with the balance of forces in the battle, inferring dyadic fatalities from those data.

- Some battles included actors that were not CoW system members, such as Montenegro fighting in World War I. We ignore these dyadic events in this dataset since the MIE data is interstate. This sometimes leads to lower fatality levels recorded in these battles in the MIE data when compared to sources like [Clodfelter \(2008\)](#). We also exclude non-state actors and their actions for the same reason. Thus, partisan guerrilla warfare in Yugoslavia, Albania, and Greece during World War II is also not included. We are heavily involved in another project to code the actions of non-state and non-system members in all confrontations. When completed, that data can be easily merged with existing confrontation, participation, and events data.
- Some of our uncertainty concerning overall confrontation fatalities results from not having good information for specific battles. This implies that, though we have good information over how many fatalities were incurred by specific states during the entire war (this is especially true for English-speaking countries), the battle fatality uncertainty induces wider fatality ranges as the data is aggregated to the confrontation level. We have done our best to narrow these ranges whenever possible.
- We introduce a war battle named “General Conflict” to describe situations in which substantial fatalities were incurred over a relatively short period of time, but those fatalities cannot be attributed to any particular named battle. This describes the trench warfare found in World War I and the constant fighting found in the Korean War and especially the Vietnam War.
- The independence wars of the Baltic states often had concurrent, overlapping battles that were significant to different independence wars. We did not want duplicate events in the data, so we attributed battles based on location. For example, a battle that occurred in modern Latvia that also influenced Lithuanian independence was attributed to Latvia because of its location. We still coded active participants for each battle, so that same battle may include Lithuanians and Latvians fighting together against Soviet forces.
- Battles include ship fatalities if the battle was principally a naval battle. Isolated shipping fatalities are coded as attacks (e.g. U-boat attacks). We were able to use various Honour Rolls and other sources to code shipping fatalities with a great deal of precision in most cases.
- Both world wars included substantial numbers of troops seconded to other commands. In most cases we were able to determine the country of origin of those troops and attribute those fatalities to origin countries. For countries that were not yet considered state system members, like India in World War II, troop fatalities were attributed to the commanding country (Britain in the case of India).

References

- Clodfelter, Micheal. 2008. *Warfare and armed conflicts: a statistical encyclopedia of casualty and other figures, 1494-2007*. McFarland.
- Gibler, Douglas M. and Steven V. Miller. Forthcoming. "The Militarized Interstate Events (MIE) Dataset, 1816-2014." *Conflict Management and Peace Science* .
- Jones, Daniel M, Stuart A Bremer and J David Singer. 1996. "Militarized interstate disputes, 1816–1992: Rationale, coding rules, and empirical patterns." *Conflict Management and Peace Science* 15(2):163–213.
- Min, Eric. 2021. "Interstate War Battle dataset (1823–2003)." *Journal of Peace Research* 58(2):294–303.