

Days of Rage: Introducing the NAVCO 3.0 Dataset

Erica Chenoweth¹

Jonathan Pinckney²

Orion Lewis³

Abstract

Although the empirical study of strategic nonviolent action has expanded in recent years, no current dataset provides detailed accounts of the day-to-day methods and tactics used by various nonviolent and violent actors seeking political change. We introduce the Nonviolent and Violent Campaigns and Outcomes (NAVCO) version 3.0 dataset, which assembles over 100,000 hand-coded observations of nonviolent and violent methods in 21 countries around the world between 1991 and 2012. Researchers can use these data and their associated coding framework to (1) replicate or challenge existing findings about nonviolent and violent action; (2) to test or uncover novel insights about the dynamics of violent and nonviolent action; and (3) recode existing protest events databases to capture specific variations in risk and disruption across event types. In particular, scholars can use these data to better understand which types of lower-level interactions between dissidents and regimes lead to large-scale mobilization; which sequences of nonviolent methods are most effective; and which types of spatial and participation diffusion yield the highest likelihood of success.

Key words: nonviolent action; civil resistance; protest; repression; diffusion; events data

Word count: 5,997

¹ Josef Korbel School of International Studies, University of Denver

² Department of Sociology and Political Science, Norwegian University of Science and Technology

³ Department of Political Science, Middlebury College

Introduction

Few recent instances of contentious politics were more powerful than the occupation of Tahrir Square, ending in the ouster of Hosni Mubarak in February 2011. Yet many did not acknowledge that the Tahrir Square sit-ins were the culmination of a years-long episode of contention—one that began in earnest with the Kefaya movement in 2005. Indeed, contentious activities from below became a routine part of Egyptian politics throughout the late 2000s.

Also less widely known was the wave of labor strikes concurrent with the Tahrir Square protests. Demonstrators across Egypt also blocked roads or engaged in other less dramatic forms of resistance. While less prominent in international images of the 2011 revolution, these forms of contention played a key role in the movement's success (International Crisis Group, 2011).

These examples demonstrate both the geographic dispersion of the January 25th revolution in Egypt and the diversity of its participants. They also represent crucial shifts in methods of contention—from concentrated acts of commission, like protests and the Tahrir Square occupation to dispersed acts of omission, like strikes and stay-aways. Scholars have long argued that such tactical innovation is a crucial factor in the success of civil resistance, defined as the nonviolent application of power by unarmed civilians (McAdam, 1982). In particular, shifts between methods of concentration and dispersion, or acts of commission and omission (Schock, 2005; Chenoweth & Stephan, 2011, 2014; Ackerman & Merriman, 2015) can have transformative potential.

The January 25 Revolution illustrates how civil resistance is far from momentary, monolithic, and modular. Instead, it often possesses a diverse range of tactics with differential dynamics. The *cumulative* effects of contentious practices over numerous years, the rapid *diffusion* of contention across participants and space, and the *sequencing of methods* to create disruption at a relatively low cost to movement participants are largely absent from the systematic empirical bases from which scholars study contentious politics.

Large-scale, aggregate datasets on nonviolent campaigns do not provide an appropriate unit of analysis with which to test such propositions or observe mechanisms during a campaign (e.g. Chenoweth & Lewis, 2013). Not only do aggregate datasets bias the sample toward contentious actions that have already “matured” into campaigns, they may misclassify some

campaigns as nonviolent or violent when observed tactics are mixed. Existing tactical-level studies are largely limited to single-country analyses (Wasow, 2015; Huet-Vaughn, 2015). And no such studies have evaluated the micro-dynamics of civil resistance within and across authoritarian regimes.

To address these scholarly and practical demands, we expanded the Nonviolent and Violent Campaigns and Outcomes (NAVCO) dataset to include event-level data. The NAVCO data project is a multi-level data collection effort that catalogues major nonviolent and violent resistance campaigns around the globe. The project produces aggregate-level data on resistance campaigns from 1900 to 2014 (NAVCO v1), annual data on campaign behavior from 1945 to 2014 (NAVCO v2), and events data on tactical selection during campaigns from 1991 to 2012 (NAVCO v3.0). NAVCO v3.0 is the first event-based dataset concentrated on the tactics and dynamics of oppositional methods categorized by theorized effects. It contains actions by both nonviolent and violent anti-government campaigns, as well as responses by governments, domestic non-aligned parties, and international actors. The data covers 21 countries from every region of the world from 1991 to 2012. For nonviolent methods, we categorize each method according to whether it is an act of commission or omission, or whether it is spatially concentrated or dispersed, allowing scholars to evaluate the distinct and cumulative effects of these different types of tactics.

In the remainder of this article we outline the data's theoretical rationale, describe the data collection process, and present information on the variables. While we do not showcase the data's full potential, we illustrate some of its possible uses by replicating the relative effectiveness of nonviolent tactics in achieving government concessions, and presenting a novel finding on the relationship between protest size and the likelihood of repression. We conclude with suggestions for future research.

Theoretical priors

The theory of nonviolent action rests on several key assumptions that inform the framing and construction of these data. Following seminal scholars such as Gene Sharp (1973), we assume that various conflict actors have agency—they can weigh, choose, and implement strategic choices with varying degrees of constraint and opportunity. Except possibly in the most

extreme circumstances, nonviolent options are available to most actors, and can often be used as a functional equivalent to violence (Sharp, 1973). Indeed, as some authors have shown (Kaplan, 2017), nonviolent resistance options are often available even in extreme settings such as armed conflict.

Second, we assume that conflict actors employ diverse, culturally and historically embedded repertoires of contention that almost always include both violent and nonviolent tactics (Tarrow, 1998; Tilly, 2006). Individual tactics and the repertoire as a whole change over time as a result of diverse stimuli, including political incentives, strategic learning, and contingent accidents of history. Specific tactics can have both individual and interactive effects. Thus, while looking at individual elements of the contentious repertoire in isolation can be fruitful, conflict chronologies are much more accurate when they provide information about violent and nonviolent methods by various actors, in interaction with one another and in their embedded political and cultural contexts.

Third, we assume that no conflict actor is monolithic. States are comprised of many different constituent organizations, which all employ people with varying degrees of loyalty to the incumbent. Broader societies are also diverse, with innumerable latent claims and expressed interests. And social movements are also often quite heterogeneous, often involving complex coalitions of actors.

Fourth, we assume that words and actions are both important and need not be considered in isolation. We also assume that protests do not represent the entire universe of nonviolent actions; other kinds of nonviolent action such as strikes, boycotts, or sit-ins have different political, economic, and social impacts and risks. We are agnostic about whether these different nonviolent actions operate in the same way everywhere in the world. We are also agnostic as to whether these actions operate in the same way within the same conflict at different periods in time.

Comparison to existing datasets

There are several existing datasets that NAVCO v3.0 complements. NAVCO v1.0 contains global aggregate, campaign-level information for violent and nonviolent campaigns with goals of regime change, anti-occupation, and secession from 1900 to 2006 (Chenoweth,

2008). NAVCO v2.0 disaggregates the individual campaign into campaign years, with variables coded for each annual time span (Chenoweth & Lewis, 2013). The forthcoming NAVCO v2.1 dataset adds additional campaign-years to the NAVCO v2.0 data, as well as expanding the number of variables (Chenoweth & Shay, 2016).

But NAVCO v1 and 2 are limited to campaigns with maximalist goals of regime change, anti-occupation, or secession. Yet understanding how reformist movements emerge, as well as the patterns of state responses to them, may be crucial in deepening our understanding of how maximalist campaigns emerge.⁴ Minorities at Risk-Organizational Behavior (MAROB) (Asal, Pate & Wilkenfeld, 2008), which covers religious and ethnic minority organizations from 1980 to 2004, is similarly limited to organizations representing claims of ethnic- or religious minorities in MENA and Eastern Europe. NAVCO v3.0 contains data on all events—regardless of the type of claim—in the five major world regions.

NAVCO v1 and 2 rely on an *a priori* decision as to whether particular campaigns were primarily violent or nonviolent. While NAVCO v2.0 and 2.1 allow for higher gradation in these distinctions by providing information on the existence of radical flanks, the definitions of these categories is still fairly rigid by the year. Diverging from prior coding decisions, NAVCO v3.0 is agnostic about whether campaigns are observably nonviolent or violent. Instead, our event-based coding allows for a mix of nonviolent and violent actions. Moreover, we do not assume the existence of campaigns just because we observe events. Instead we provide a method of identifying campaigns inductively by focusing on substantive links across events, discussed in more detail below—a key departure from existing campaign- or organizational-level datasets.

Of course, there are other event datasets with which scholars could analyze the cumulative effects of different nonviolent actions. Examples include the European Protest and Coercion Dataset (EPCD) (Francisco, 2000) and the Integrated Crisis Early Warning System (ICEWS) data (O'Brien, et al., 2010). However, as with MAROB, the EPCD is regionally-bound

⁴ The Swarthmore Global Nonviolent Action Database contains data on non-maximalist campaigns, but the dataset has no systematic inclusion criteria, nor are the data disaggregated below the campaign level.

and does not feature data that one could use to categorize the methods of contention themselves. Although it is a global dataset, ICEWS has a high degree of error and misclassification, suggesting that it is not necessarily a reliable source for testing micro-level questions.

NAVCO v3.0 divides its unit of analysis into the event-day rather than the event. This provides maximum reliability when researchers are interested in producing analysis at a daily level. In contrast, datasets with the event as a unit of analysis may lead to misspecification when attempting to convert events into daily patterns. For example, events described as continuous in multi-day event datasets may occur only once a week or once a month.

NAVCO v3.0 is also distinct from other event data in its inclusion of government behavior. For example, if a protest was the subject of government repression we not only code the level of repression as a variable relating to the protest itself, but also code each of the specific repressive actions taken by the government.

The Social Conflict in Africa Dataset (SCAD) is, perhaps, the most comparable among existing datasets (Salehyan, et al., 2012). It includes events, geolocations, and a broad range of claims in Africa and Latin America from 1990 to 2013. Crucially, however, NAVCO v3.0 extends beyond SCAD by allowing researchers to investigate various actors' rhetorical strategies during anti-government campaigns. One can see not only a protest and police arrests in response, but also whether the opposition made threats before the protest, whether the government escalated its police alert status, and whether international human rights groups condemned government repression. Table I includes a snapshot comparison of these data sources.

[Table I here]

The case selection process

Because of the intensive data collection and coding process, our resources allowed us to fully code only 21 countries. These countries are not globally representative; we selected them primarily based on intrinsic interest, that is to say, based on the judgments of the research team about the potential for data from a particular country to expand our knowledge of major

questions related to violent and nonviolent resistance.⁵ However, we wanted to include geographic range and diversity so we collected data from Asia, the Middle East, Africa, Europe, and Latin America. The data in NAVCO v3.0 therefore represent countries with very diverse political systems, economic conditions, and histories of contention. The sample includes at least one country from all major world regions.

However, the current sample has two particular sources of bias. First, we oversampled from the Middle East and Africa due to interest in these cases among the principal investigators as well as a tendency for such regions to be popularly represented as inherently violent.⁶

Second, because our aim was to understand, in part, what kinds of events precede massive popular uprisings, we focused on countries with such uprisings at some point during the sample time period. The data does include some countries with no major campaigns—Estonia and Tanzania, for instance—but overall our sample may be more “eventful” than the global average during this time period.⁷ In the online appendix, we examine some of the potential consequences of these two sources of bias by comparing the NAVCO v2.0 campaign-years that are represented in NAVCO v3.0 with those that are not.

[Table II here]

The coding process

After developing the codebook for the project (see attached), we hired a team of research assistants (RAs) among students at the University of Denver and Middlebury University. These RAs collected the NAVCO v3.0 data using Agence France Press (AFP) newswires downloaded from the Lexis-Nexis Academic database. We chose AFP because of its global coverage,

⁵ We describe the rationales for each country and reasons why some countries have only partial data available in the online appendix.

⁶ A further benefit of this overrepresentation is the possibility of data cross-validation with SCAD and/or combination of the data with MAROB.

⁷ Estonia and Tanzania have only 168 and 408 event-days recorded in NAVCO v 3.0, a much lower total than the average in our other completed countries, which average 5,210 event-days from 1991 to 2012.

consistent reporting standards, and lack of space constraints. As opposed to a traditional newspaper, AFP collects numerous individual stories each day and can thus provide coverage of typically underreported events, a major challenge in observing nonviolent action (Day, Pinckney & Chenoweth, 2015). Although ideally a dataset like this would feature local and native-language source materials, resource constraints prevented us from adopting this approach.

Individual RAs or small teams of two or three RAs coded all events in a country for the entire period of study. Research assistants downloaded all hits for a standard search string (see the codebook). They then read all the reports returned from this search string, coding every event mentioned in the discovered articles.

A key advantage of the human coding process, relative to machine coding, is that through this process RAs developed country-specific expertise that aided in their interpretation of events. RAs were also able to see the connections between distinct events reported in separate articles.

This country-specific knowledge was key for developing the necessary expertise to determine the correct population of events to include in the dataset. This is because a key criterion for inclusion was whether a particular event by governments or transnational actors was politically relevant to actions by a non-state, anti-government campaign. These actions did not need to have any particular degree of significant impact on the actions of anti-government campaigns; rather they were included if they were reported as a relevant to the claims or behaviors of an anti-government campaign.

Moreover, human coders are necessary to observe some events' symbolic nature, like the placing of teddy bears in a Ukrainian fountain as a protest of the Russian annexation of Crimea. Although automated coding may someday successfully capture and accurately classify such events, such technology does not yet exist.

For situations of RA uncertainty, we included an *arbitration* variable, and a *notes* field. RAs indicated with the arbitration variable whether they had questions regarding the event's inclusion or specific variable coding. RAs included specific inquiries, as well as explanations for potentially questionable coding, in the notes field. All of the lines marked for arbitration were later reviewed by the project manager, who, after reviewing the source articles and surrounding code, made a determination about keeping or deleting the ambiguous lines.

In addition, throughout the data collection process we replicated several random samples from different RAs to ensure that error rates were relatively low and ensure inter-rater reliability. This process showed an error rate that improved significantly as the replication revealed ambiguities in our initial codebook and helped us develop standard operating practices and improved training methods.⁸

Several key practices improved our inter-rater reliability over time. First, RAs went through common training. Second, we kept a wiki site with answers to RAs' frequently-asked-questions. Third, RAs worked in a common lab space, facilitating consultation. Fourth, each RA was subject to a two-month probationary period, after which the project manager reviewed all of their completed work.

RAs coded observations either chronologically or reverse chronologically, introducing the possibility that stories coded at either end of our period of examination could be recorded less reliably than in the middle years. However, coders returned to the initial codings to make updates based on lessons learned as the coding proceeded. Moreover, at the completion of RA coding, the project manager checked and corrected every observation by hand, adding additional consistency and reliability.

Variables

NAVCO v3.0's core structure is based on the CAMEO code developed by the KEDS project (Schrodt, 2012). For each event-day, RAs collected information on the *actors* engaging in *verbs* directed at *targets*. Each of these could be coded with up to three levels of specificity.⁹ We also include a brief textual description of the actors.

We limit our primary CAMEO codes to a small set (listed in the codebook) in order to simplify comparison and analysis. Secondary and tertiary codes are somewhat more flexible. Researchers can use them more appropriately for more specific questions where broad comparability is not a central concern.

⁸ We report quantitative measures of inter-rater reliability in the online appendix.

⁹ We provide an illustrative line of core NAVCO v3.0 code in the online appendix.

NAVCO v3.0 also includes several contextual variables: the *date*, a brief *description*, the *title* of the event's source article, the geographic *scope*, the specific *localities*, and which *type* of actor performed the action. This is a categorical variable with 5 possible values, government, anti-government campaign, international actors, domestic non-aligned actors, and local government actors.

Anti-government campaign actions include several auxiliary variables. We have a 7-category measure of campaign *goals*. The categories include three maximalist goals (regime change, anti-occupation, and secession), three reformist goals (major institutional transformation, policy change, and greater autonomy), and a residual category for events with unknown goals.

We also have a measure of *tactical choice*, indicating whether an event was primarily violent, nonviolent, or mixed. If an event was primarily nonviolent we code which of the *categories of nonviolent action* (Sharp, 1973) it falls into: protest and persuasion, noncooperation, and nonviolent intervention. We divide "protest and persuasion" into two categories, one for verbal attempts at persuasion and one for protest actions, and include a fifth category for political engagement.

For tactics of noncooperation we code a variable for which of the *categories of noncooperation* it falls into: social noncooperation, economic noncooperation-strikes, economic noncooperation-boycotts, or political noncooperation. As mentioned above, we also code whether the event is an *act of commission or omission* and whether it is a tactic of *concentration or dispersion*.

If the source articles contain any information on responses to particular actions, we code the *state posture*. This is a seven-level ordinal variable ranging from full concessions to all stated demands to physical repression with the intent to kill, following Dugan and Chenoweth (2012). Contingent on data availability we also include measures of *fatalities*, *injuries*, and *number of participants*. Finally we include measures of *property damage* and *economic impact*, both in an ordinal measure of scope of property damage and an open text field to summarize economic impact.

We instructed RAs to follow a conservative approach when coding all variables. For example, RAs only coded number of participants if the source specifically stated participation

numbers. In many cases sources reported ambiguous or multiple numbers of participants. In these cases RAs were instructed to put down the entire potential range of participants (Day, Pinckney, & Chenoweth, 2015).

This approach means that for several of our variables the amount of missing data is quite high. However, it ensures greater reliability of the data.

Summary statistics and patterns of events

We present below some summary statistics from NAVCO v3.0, dividing the data by country, actor, verb, and tactical choice. As the table shows, each country is well-represented, with Estonia having the smallest number of event-days (n=168) and Syria the largest (n=11,694). The majority of actions are either by governments or anti-government campaigns.

The largest individual category of actions is CAMEO verb 14, which captures various forms of anti-government protest such as public rallies, labor strikes, hunger strikes, and riots. The data contain many both violent and nonviolent actions (60.5% and 37.7% of the observations, respectively), with a smaller number of “mixed” actions (constituting 1.8% of the observations).

[Table III here]

As mentioned previously, NAVCO v3.0 does not impose categorization rules to aggregate individual actions into campaigns. Instead, we leave the decision on defining campaigns to the user. However, we estimate that under any number of aggregation criteria NAVCO v3.0 contains hundreds (if not thousands) of violent and nonviolent campaigns.

For example, Pinckney (2016) identifies 228 distinct violent and nonviolent campaigns in just 14 of NAVCO v3.0’s 26 countries. Pinckney follows an inclusion rule that approximates the definition of “campaign” common in the civil resistance literature (Chenoweth & Stephan, 2011): a sequence of events meaningfully linked through common actors or goals; consisting of at least three distinct physical events; and separated by less than a year (Pinckney, 2016: 78-79).

Putting the data to use

NAVCO v3.0's detail and variety of variables mean that the data has many potential uses. It can provide a new source of testing for existing arguments in the conflict, repression, and civil resistance literatures as well as be a resource for examining relationships never before examined. We illustrate this potential by examining how the data shed light on a few seminal arguments in the conflict and contentious politics literatures. We present summary results here, but full results and replication instructions are available in the online appendix.

First, we examine a core contention from the civil resistance literature that nonviolent action is more effective than violent action (Chenoweth & Stephan, 2011). One key measure of effectiveness is the ability of individual tactics to obtain government concessions (Cunningham, 2016). The cross-tabulation below shows that, while concessions are rare for both types of contention (occurring in only 673 event-days, or 1.5% of the 43,704 event-days), the overwhelming majority of government concessions in our sample came in response to nonviolent actions (612 of the 673 cases, or 91%).¹⁰ This difference in average level of concessions is highly significant, with a $p < 0.0000$ in a one-tailed t-test.

[Table IV here]

For most of the countries in NAVCO v3.0 coverage is sufficient to provide statistical power to test relationships between factors such as participation, tactical choice, repression, and concessions not just at the cross-national level but within countries as well. These disaggregated tests may reveal variation that can enhance the broader, more abstract lessons gained from pooled cross-national analysis.

To illustrate, we test the relationship between two factors central in contentious politics: mobilization and repression. We operationalize mobilization by the logged observed

¹⁰ The coding of the concessions variable sought to limit un-replicable individual coder interpretation. RAs coded state posture as indicating government concessions for a particular event-day if sources explicitly reported that the concession was made in response to that event. If it seemed clear from context that a concession took place in response to a particular action but the source did not explicitly make the connection coders still coded concessions as taking place but recorded their rationale in the notes field.

participation in protest events (verb code = 14) and operationalize repression as a one if the state posture variable is a five or higher and zero if otherwise. To account for possible non-linear effects we logged the participation variable and also included squared and cubed transformations of logged participation. In addition, we controlled for whether the protest was completely nonviolent or had at least some violent elements.

In a logistic model of repression with the full pooled sample, all three transformations of the participation variable are highly significant.¹¹ Thus, the model predicts an s-shaped relationship between the size of protests and repression, depicted in Figure 1 below.¹² As protest size increases the likelihood of repression is low. It then increases sharply, declines again for mid-sized protests, and then increases again at extremely high levels of participation. The solid line depicts the predicted likelihood of repression for nonviolent protests and the dashed line depicts the predicted likelihood of repression for violent protests. Although the shape of the pattern is similar for both, at all levels of participation the relative likelihood of repression for nonviolent protests is lower.¹³

[Figure 1 here]

How does the picture change when separating the observations by country? We analyzed the model for all 26 countries. For five countries all three transformations of participation remained significant at a $p < 0.05$ level. For country models where at least one transformation was not statistically significant we removed variables from the model using stepwise deletion until all remaining variables were statistically significant.

Dividing the data in this way reveals both areas of continuity and important differences (see Figure 2). In almost every case nonviolent protests are less likely to be repressed than

¹¹ Coefficients and standard errors reported in the online appendix.

¹² With logged participation along the x axis and the model's predicted likelihood of repression on the y axis.

¹³ We also ran a version of the model dropping the nonviolence variable to check for robustness. All three coefficients remain statistically significant and the shape of the relationship is fundamentally the same. We report coefficients and standard errors in the online appendix.

violent protests of a similar size (as shown by the lower position of the predicted probability curve). The one exception is in Uzbekistan, where nonviolent protests are slightly more likely to be repressed than violent protests. However, the number of observations from Uzbekistan is quite low ($n = 27$), so the findings are very unstable.

While the effect of nonviolent protest was similar across cases, the participation's effect differs widely. Eight countries have no significant relationship between any of the transformations of participation and repression and thus are not depicted. The 18 remaining countries have significant predicted relationships ranging from s-curves similar to the pooled sample to simple positive or negative linear relationships, to U-curves and inverted U-curves.

Because this analysis is intended primarily to illustrate the potential uses of the data we do not offer a specific causal argument to explain these varying relationships. In future research we hope to examine this participation-repression relationship with models better designed to address simultaneity bias. However, this test highlights the benefit from both a common cross-national format for ease of comparison and detailed country-specific data to confirm whether cross-national trends hold at the country level. We should note, too, that the data contain more detailed information about each event's location including, where possible, city and region. This allows researchers ample opportunity for subnational analysis.

[Figure 2 here]

Conclusion

These illustrations are primarily replications and expansions of previous work. However, these data also lend themselves to more novel analysis—in particular, the cumulative impacts of different methods over time; the effects of different sequences and combinations of concentrated and dispersed methods as well as acts of commission and omission; and the impacts of the spatial and participatory diffusion of campaigns. We intend to take on each of these important questions in future research.

Moreover, a number of research questions currently prominent in the study of civil resistance can be addressed using the existing NAVCO v3.0 data. For instance, Chenoweth and Stephan (2011) argue that participation is one of the key advantages that nonviolent resistance

possesses relative to violent resistance. How does participation in nonviolent action actually emerge and grow? What tactics of nonviolent action tend to have higher levels of participation, greater mobilization potential, and thus perhaps a greater impact on success? NAVCO v3.0 also provides a fine-grained look at the preconditions for major revolutionary nonviolent campaigns. Are there regular patterns of contentious politics that precede major outpourings? What kinds of actions by governments and anti-government campaigns are the best predictors of future instability? Moreover, NAVCO v3.0 contains extensive information about oppositional actors and their profiles. This information could be used to test prominent theories from political economy on mass mobilization and regime change, which often makes certain assumptions about who the participants in anti-regime mobilization are (e.g. elites vs. the middle class vs. the poor).

Data collection and coding for NAVCO v3.0 is time- and resource-intensive. In the future, we hope to encourage other researchers to adopt the inclusion and coding criteria, to expand the representativeness of the cases and increase the number of observations available for study. Improved automated collection methods may reduce the amount of money and labor required to collect the data in the future. Currently, however, NAVCO v3.0 makes a number of contributions to the study of civil resistance as well as broader scholarship on political contention. It is the first data source to systematically incorporate detailed characteristics of nonviolent action, as proposed and theorized by leading scholars and activists, into its structure. Second, it is a detailed and finely-tuned event-day dataset, incorporating a much wider range of strategic actions than other human-coded datasets while still providing far superior levels of accuracy than current machine-coded datasets. Third, its many different kinds of information make it useful for research in many different fields.

The greatest strengths of NAVCO v3.0 are its broad inclusion criteria, its wide variety of actors and actions, and its detailed categorization of contentious events, which make it an ideal source for rigorously analyzing the complex dynamics of violent and nonviolent actions around the world.

Replication

The dataset, codebook, and do-files for this article can be downloaded from <http://www.prio.org/jpr/datasets>. All analysis was conducted using Stata 13. The NAVCO v3.0

data is also available for download at www.navcodata.com.

Acknowledgments

The authors gratefully acknowledge financial support from the International Center on Nonviolent Conflict, the Social Science Foundation at the University of Denver, the Allbritton Center for the Study of Public Life at Wesleyan University, and the Department of Political Science at Middlebury College. We also thank Jonathan Caverley and participants in a panel at the 2016 American Political Association Annual meeting for their helpful comments and feedback. We also thank our research assistants for their work on this project.

References

- Ackerman, Peter, & Hardy Merriman (2015) "The checklist to end tyranny." In: Mathew Burrows & Maria J. Stephan (eds.) *Is Authoritarianism Staging a Comeback?* Washington, DC: The Atlantic Council.
- Asal, Victor; Amy Pate & Jonathan Wilkenfeld (2008). "Minorities at risk organizational behavior data and codebook version 9/2008." *Minorities at Risk Project* (<http://www.mar.umd.edu/data.asp>).
- Chenoweth, Erica (2008) The Nonviolent and Violent Campaigns and Outcomes (NAVCO) Dataset, version 1.0. Josef Korbel School of International Studies, University of Denver (https://www.du.edu/korbel/sie/research/chenow_navco_data.html).
- Chenoweth, Erica, & Christopher Shay (2017) "Updating nonviolent campaigns: Introducing NAVCO 2.1." *Presented at 2017 APSA Annual Meeting*.
- Chenoweth, Erica, & Maria Stephan (2011) *Why Civil Resistance Works: The Strategic Logic of Nonviolent Resistance*. New York: Columbia University Press.
- Chenoweth, Erica, & Orion Lewis (2013) "Unpacking Nonviolent Campaigns: Introducing the NAVCO 2.0 Dataset." *Journal of Peace Research* 50, no. 3: 415-23.
- Cunningham, Kathleen Gallagher (2016). "The efficacy of nonviolence in self-determination disputes." *Working Paper* (http://www.kathleengallaghercunningham.com/uploads/4/5/5/8/45589607/efficacy_of_nv_cunningham_may_10_2016.pdf).

- Day, Joel; Jonathan Pinckney, & Erica Chenoweth (2015) "Collecting data on nonviolent action." *Journal of Peace Research* 52(1): 129-33.
- Francisco, Ronald A (2000) *Codebook for European Protest and Coercion Data, 1980-1995*. Lawrence, KS: University of Kansas (<http://web.ku.edu/~ronfrand/data/codebook.doc>).
- Huet-Vaughn, Emiliano (2015) "Quiet riot: The causal effect of protest violence." *Working Paper* (<https://www.aeaweb.org/conference/2016/retrieve.php?pdfid=1372>).
- International Crisis Group (2011) "Popular protest in North Africa and the Middle East (I): Egypt victorious?" *International Crisis Group* (http://www.ecoi.net/file_upload/1226_1298906986_101-popular-protest-in-north-africa-and-the-middle-east-i-egypt-victorious.pdf).
- Kaplan, Oliver. *Resisting War: How Communities Protect Themselves*. New York: Cambridge University Press, 2017.
- McAdam, Doug (1982) *Political Process and the Development of Black Insurgency, 1930-1970*. Chicago, IL: University of Chicago Press.
- Pinckney, Jonathan (2016) *Making or Breaking Nonviolent Discipline in Civil Resistance Struggles*. Washington, DC: International Center on Nonviolent Conflict Press.
- Salehyan, Idean, et al (2012). "Social conflict in Africa: A new database." *International Interactions* 4: 503-11.
- Schock, Kurt (2005) *Unarmed Insurrections: People Power Movements in Nondemocracies*. Minneapolis, MN: University of Minnesota Press.
- Schrodt, Philip A (2012) "CAMEO: Conflict and mediation event observations event and actor codebook." *Parus Analytical Systems, LLC* (<http://eventdata.parusanalytics.com/cameo.dir/CAMEO.Manual.1.1b3.pdf>).
- Sharp, Gene (1973) *The Politics of Nonviolent Action*. Boston, MA: Porter Sargent Publishers.
- Tarrow, Sidney (1998) *Power in Movement: Social Movements and Contentious Politics*. New York: Cambridge University Press.
- Tilly, Charles (2006) *Regimes and Repertoires*. Chicago, IL: University of Chicago Press.
- Wasow, Omar (2015) "Nonviolence, violence and voting: The effects of the 1960s Black protest movements on white attitudes and voting behavior." *Working Paper* (http://www.omarwasow.com/Protests_on_Voting.pdf).

ERICA CHENOWETH, b. 1980, PhD in political science (University of Colorado, 2007); Professor, Josef Korbel School of International Studies, University of Denver (2015-); Research Fellow, One Earth Future Foundation (2015-).

JONATHAN PINCKNEY, b. 1986, PhD in international studies (University of Denver, 2018); Post-Doctoral Research Fellow, Norwegian University of Technology and Science (2018-).

ORION LEWIS, b. 1978, PhD in political science (University of Colorado, 2009); Assistant Professor, Middlebury College(2017-); Visiting Assistant Professor, Middlebury College (2012-17); Post-Doctoral Research Associate, Wesleyan University (2009-12)

Table I. Comparison of major conflict data sets

Dataset	Unit of Analysis	Included Characteristics	Population	Include Speech Events?
MAROB	Organization-Year	Actor, Action Types, Goals, Structure, Repression, Outcomes,	1980-2004 (MENA, Eastern Europe)	No
SCAD	Event	Actor, Action Type, Target, Issues Location, Size, Repression	1990-2015 (Africa, Latin America)	No
EPCD	Event-Day	Actor, Action Type, Target, Issues, Location, Repression	1980-1995 (Europe)	No
ICEWS	Event-Day	Actor, Action Type, Target, Location	1979-2015 (Global)	Yes
NAVCO 1.0	Campaign	Actor, Target, Size, Goals, Violence/Nonviolence, Outcome, Repression	1900-2006 (Global)	No
NAVCO 2.0	Campaign-Year	Actor, Target, Size, Goals, Violence/Nonviolence, Outcome, Repression,	1946-2006 (Global)	No
NAVCO 3.0	Event-Day	Actor, Action Type, Target, Size, Location, Violence/Nonviolence, Types of Nonviolence, Repression, Outcomes	1991-2012 (21 Full Countries, 5 Partial)	Yes

Table II. Countries covered in NAVCO v 3.0, 1991 to 2012

Asia	Middle East & North Africa	Africa	Europe	Americas
Uzbekistan	Bahrain	Madagascar	Estonia	Mexico
Pakistan	Egypt	Kenya	Ukraine	United States*
China*	Libya	Sierra Leone		
India*	Yemen	South Sudan		
South Korea*	Tunisia	Sudan		
	Jordan	Tanzania		
	Turkey			
	Syria			
	Iraq*			
	Morocco			
	Algeria			

*Indicates partial years collected. Years of coverage include: 1991-1992, 2010-2012 (China); 1991-1996, 2012 (South Korea); 1991, 2011-2012 (India); 1999-2000, 2009-2012 (Iraq); and 2007-2011 (United States).

Table III. Event-day summary statistics

Country	n	Verb	n	Actor Category	n	Tactical Choice	n
United States (partial)	1,775	General Statement	11,741	Government	39,778	Violent	16,795
Mexico	3,838	Appeal	7,368	Anti-Government	43,799	Nonviolent	26,981
Estonia	168	Express Cooperative Intention	4,245	International	27,528	Mixed	808
Ukraine	3,752	Negotiate	9,351	Non-Aligned	1,192	Total	44,584
Sierra Leone	4,869	Political Cooperation	3,506	Local Government	84		
Kenya	5,342	Material Cooperation	707	Total	112,381		
Tanzania	408	Provide Aid	574				
Madagascar	1,432	Yield	2,657				
Morocco	2,894	Investigate	1,144				
Algeria	5,723	Demand	2,875				
Tunisia	1,684	Disapprove	13,551				
Libya	3,544	Reject	1,828				
Sudan	8,058	Threaten	3,193				
South Sudan	589	Protest	15,170				
Turkey	8,161	Exhibit Force Posture	1,852				
Iraq (partial)	8,138	Reduce Relations	1,374				
Egypt	8,617	Coerce	9,366				
Syria	11,694	Repress & Abuse	10,468				
Jordan	1,529	Violent Combat	11,109				
Yemen	8,617	Mass Violence	10				
Bahrain	2,106	Defect	292				
Uzbekistan	1,065	Total	112,381				
China (partial)	3,206						
South Korea (partial)	963						
India (partial)	3,347						
Pakistan	10,862						
Total	112,381						

Table IV. Cross-tabulation of event type and direct concessions

	Violent	Nonviolent
No Concessions	17,343	25,688
Concessions	61	612
Total Events	17,404	26,300

Figure 1. Participation and repression (pooled)

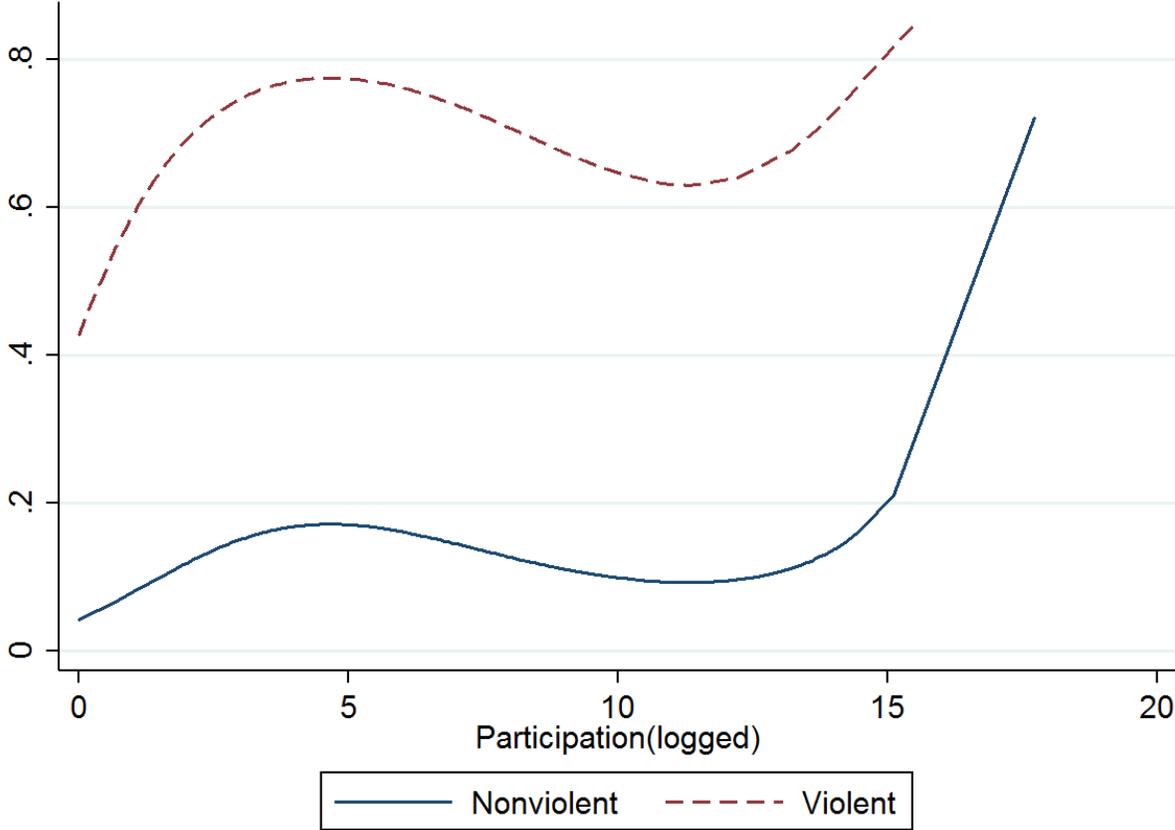


Figure 2. Participation and repression (country-specific)

