

Quantifying intelligence cooperation: The United States International Intelligence Behavior (USIIB) dataset

Ersel Aydinli and Musa Tuzuner
Journal of Peace Research 2011 48: 673
DOI: 10.1177/0022343311412808

The online version of this article can be found at:
<http://jpr.sagepub.com/content/48/5/673>

Published by:



<http://www.sagepublications.com>

On behalf of:

Peace Research Institute Oslo



Journal of Peace Research Replication Data

Additional services and information for *Journal of Peace Research* can be found at:

Email Alerts: <http://jpr.sagepub.com/cgi/alerts>

Subscriptions: <http://jpr.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

>> [Version of Record](#) - Sep 22, 2011

[What is This?](#)

Quantifying intelligence cooperation: The United States International Intelligence Behavior (USIIB) dataset

Journal of Peace Research

48(5) 673–682

© The Author(s) 2011

Reprints and permission:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0022343311412808

jpr.sagepub.com



Ersel Aydinli

Department of International Relations, Bilkent University

Musa Tuzuner

Intelligence Studies Research Center, Turkish National Police Academy

Abstract

This report describes the creation of the United States International Intelligence Behavior dataset (USIIB). The USIIB represents the first collection of event data specifically intended for exploring in a quantifiable manner the international intelligence cooperation behaviors of the United States. A total of 293,615 events are recorded in the USIIB, covering the years 2000–09. The report first provides a detailed description of the steps involved in building such a dataset, including the development of search terms, the use of a machine coding program (TABARI – Text Analysis by Augmenting Replacement Instructions) to extract data from wire news releases, and the extension of an existing coding scheme (CAMEO) to include intelligence behaviors. Following a discussion of issues related to the reliability and validity of event datasets in general and the USIIB in particular, the report then includes suggestions and examples for how the data in the current USIIB dataset may be used in order to add to our understandings of patterns and anomalies in international intelligence cooperation behavior. As a specific example, it offers results from an empirical test exploring variation in intelligence cooperation behaviors among democracies and non-democracies, asking specifically whether the United States has been more likely in the early 21st century to cooperate on intelligence matters with democratic states, and finding this not to have been the case. Finally, it aims to provide a guide for others who would like to extend this dataset to explore intelligence cooperation activity of other countries or regions.

Keywords

event dataset, intelligence cooperation, security

Introduction

Shortcomings in intelligence cooperation both intra- and internationally have been identified as perhaps the most significant impediment to a successful countering of transnational security challenges in a post-9/11 environment (an argument made most powerfully in the *9/11 Commission Report*, 2002). While problems at the level of practice suggest the need for focused research to gain deeper understandings of intelligence cooperation behaviors, intelligence studies remain largely dispersed, limited, and under-theorized (Svendsen, 2009; see for a

recent exception Aldrich, 2009). Methodologically, the intelligence literature is arguably also limited by a lack of quantitative research. A brief survey of recent literature on intelligence cooperation reveals a predominance of case studies (e.g. Munton, 2009; Hussain, 2009; Matei, 2009; Dumbrell, 2009) and practitioners' analyses of challenges to or innovations within intelligence

Corresponding author:

ersel@bilkent.edu.tr

cooperation (e.g. Lefebvre, 2003; Strachan-Morris, 2009; Lander, 2004).

These limitations are partly due to the difficulties involved in studying a naturally secretive area of activity, but also because intelligence studies are not yet widely considered as an independent discipline or area of research within international relations. A starting point in addressing these various gaps in the literature and in producing a more systematic research agenda is, therefore, to have a comprehensive dataset that could be used to quantifiably examine different aspects of intelligence behavior by various countries. The following report describes the construction of such a dataset, based on event data from between the years 2000 and 2009, and focusing on behaviors initiated by the United States: the United States International Intelligence Behavior dataset (USIIB). The report goes on to provide guidelines for how the current dataset might be used and suggestions for how it might be extended.

Event data research

Event data research has a long history in several academic disciplines,¹ and has been used quite extensively in international relations research. The approach first saw a rapid growth in the 1960s and 1970s, with pioneering projects such as the World Event/Interaction Survey (WEIS, McClelland, 1978) and the creation of the Conflict and Peace Data Bank (COPDAB, Azar, 1980). Both of these efforts aggregated individual events by converting them into some kind of measure of cooperation or conflict. While early event data studies made use of human coding, which was time-consuming and expensive, this problem was alleviated in the 1990s when machine coding computer programs were developed. The earliest such software program was the Kansas Event Data System (KEDS) (Schrodt, Davis & Weddle, 1994), which in turn helped spawn TABARI (Text Analysis by Augmenting Replacement Instructions) and VRA (Virtual Research Associates), all of which provide relatively easy ways of extracting data from international wire news sources by coding the lead sentences in press releases – or in the case of VRA, coding the first few sentences (see <http://vranet.com/> for information

on VRA, and <http://gking.harvard.edu/data> for replication data of King & Lowe's [2003] study using VRA).

Past major event data projects on political or security-related issues have covered a wide range of events, from those at the global level (e.g. COPDAB, WEIS, IDEA – Integrated Data for Event Analysis and PANDA – the Protocol for the Analysis of Nonviolent Direct Action, Bond et al, 2003; Jenkins & Bond, 2001) to those focused on particular regions (e.g. the Intranational Political Interactions project and the European Protest and Coercion dataset; see also Howell & Barnes, 1993). Other datasets have been created to cover more focused types of events, such as global terrorism (ITERATE – International Terrorism: Attributes of Terrorist Events and TKB – Terrorism Knowledge Base), and domestic or regional terrorism (e.g. TWEED – Terrorism in Western Europe, Engene, 2007; and PCSTERROR – Project Civil Strife-Terror, Shellman, 2008). Prior to the USIIB project, no one has attempted to use event data to explore intelligence-related interactions or, more specifically, behaviors of intelligence cooperation.

Creating the USIIB

To extract data on US intelligence cooperation behavior from wire news releases and create the USIIB dataset, we used TABARI, the open source machine coding program developed by Schrodt (2006) as a successor to KEDS.

Before beginning to create the dataset, we first had to select our text source. Sources used in past datasets have included the *New York Times* (WEIS), Reuters (IDEA and KEDS, Schrodt, Davis & Weddle, 1994; Schrodt & Gerner, 1994), and Agence France Presse (AFP) (CAMEO – Conflict and Mediation Event Observation, Gerner et al, 2002; and more recent updates of KEDS, Schrodt, 2006). In the case of the USIIB, we discovered that including multiple sources was not an easy prospect. Wire news sources use different systems for dissemination, posing problems not only for the compiling of data from those sources but synchronizing those data for analysis purposes. An early attempt revealed that doing so would require a technical redesigning of the tool programs used to make the data readable by the TABARI program. For example, the tool program *NewNexisFormat.pl* (Perl) is designed to work with LexisNexis, which covers AFP, whereas Reuters is accessed through Factiva, which has its own search rules and

¹ A long tradition of such research exists in history and sociology, with the latter including a particularly rich literature on collective action and social movements (reviewed in Franzosi, 1987; Koopmans & Rucht, 1999; Olzak, 1989; Rucht, Koopmans & Neidhardt, 1999).

Table I. Comparing Reuters and AFP

	<i>Reuters</i>	<i>AFP</i>
'CIA'	827	952
'Senate Select intelligence'	0	1
'Intelligence czar'	0	4
'US national intelligence'	0	20
'Defense intelligence agency'	11	16

requires a different tool program. Direct contact with the creator of TABARI revealed that making the necessary program changes to allow for machine coding of multiple sources was a major undertaking that would require professional technical assistance, something that was not feasible for this initial project.²

The decision was therefore made to use AFP for two reasons. First, AFP can be considered to have better coverage of international intelligence events than a single US source such as the *New York Times* – even when the events being analyzed are focused on interaction behaviors initiated by the USA. A comparison of the stories covered by AFP and by a few major US newspapers over a limited period revealed that for the latter there seems to be limited news value attributed to, and therefore less likelihood of reporting on, small scale examples of intelligence cooperation events, for example the visit of an individual US intelligence representative to a single country; but from the perspective of an AFP reporter in that particular country, the event is more likely to seem newsworthy and thus get reported.³ Second, AFP is considered superior to Reuters for political news, while Reuters is considered preferable for looking at business news (Schrodt, 2006). A small investigation comparing the intelligence-related leads produced by Reuters and AFP between 1 January 2009 and 31 December 2009 revealed that AFP consistently produced more leads than Reuters (see Table I).

The following sections describe the basic steps necessary in creating a dataset, including the developing of search terms for finding relevant machine-readable text, the downloading of wire news leads, and the machine coding of these leads via prescribed coding schemes.

² Philip Schrodt, personal communication, 5–8 January 2007. We are planning to update and expand the USIIB dataset in several ways. The planned future dataset (Küresel İlişkiler Veri ve Analiz Merkezi/Data and Analysis Center for Global Affairs, or KIVAM) will prioritize the hiring of professionals to make possible the inclusion of multiple sources.

³ Based on an actual example of a visit by an FBI officer to Venezuela, which was reported in AFP but not in leading US newspapers.

The final section of the article reports on an application demonstration of how the autocoded events can be aggregated to explore a particular issue.

Developing search terms

The first phase in the development of the dataset was the selecting of appropriate search terms that would enable TABARI to pick out appropriate wire news leads revealing instances of international intelligence behaviors. Our starting point for this stage was to create a list of names of potential actors conducting or involved in US intelligence cooperation. By referring to official US government websites, we devised a list of more than 20 names and abbreviations of US intelligence community members and their supervisory organizations in both the executive (e.g. National Security Council, the President's Intelligence Oversight Board) and legislative (e.g. the Senate Select Committee on Intelligence (SSCI), House Permanent Select Committee on Intelligence (HPSCI)) branches.

Identifying these basic names allowed us to conduct a search of AFP-English news wires on LexisNexis, using search terms in the following format:

CIA OR 'Central Intelligence Agency' OR FBI OR 'Federal Bureau of Investigation' OR DIA OR 'Defense Intelligence Agency' . . .

By reading through the resulting AFP news leads, the original list of actors was extended; for example, unofficial names ('Intelligence Czar') were added to official actors ('United States Director of National Intelligence'), and sub-units were revealed, for example the National Clandestine Service, which coordinates human intelligence between the CIA and other agencies. Additional readings of each actor's webpage revealed further terms, abbreviations, names of sub-units, and phrases associated with these offices.

By repeating the search process, this time with an extended list of terms for each actor, and again reading the resulting news leads, still further search terms were generated. Essentially, this stage of the dataset development process resembles certain qualitative data analysis methodologies, in which the researcher engages repeatedly with the data to become deeply familiar with them. By examining the AFP leads and websites repeatedly, we generated extensive lists of terms related to each potential actor and thus increased the number of news leads produced. For example, the search terms developed to refer to the CIA included the following:

CIA OR 'Central Intelligence Agency' OR D/CIA OR 'Director of the Central Intelligence Agency' OR DCIA OR DD/CIA OR 'Deputy Director of the Central Intelligence Agency' OR AD/MS OR 'Associate Director for Military Support' OR ADCI/MS OR 'Associate Director of Central Intelligence for Military Support' OR NCS OR 'National Clandestine Service' OR D/NCS OR DD/NCS/CIA OR 'Intelligence Community Affairs' OR 'National Intelligence Council' OR 'National Intelligence Officers' OR 'National Intelligence Estimates' OR ...

Downloading AFP wire news leads

The next phase in the development of a dataset becomes less exploratory and more technical, as all the AFP leads must be downloaded into a folder and reformatted to be readable by TABARI. TABARI's 'nexispider.pl' program was used to do this for the years 2000–06, and NewNexisFormat.pl for the updated addition of the years 2007–09. For example, all interactions led by the United States Senate Select Intelligence Committee were listed in the USIIB with file names beginning with USIC: 'USIC(01)SSCI.000122-0601108'.⁴ Between the dates of 2 January 2000 and 31 December 2009, a total of 369,138 AFP leads were downloaded into 790 separate files.⁵

Coding event categories

TABARI and its predecessor KEDS have been used primarily to code global interactive events using the WEIS scale, which includes 22 event categories grouped according to their degree of cooperativeness. WEIS was later modified by the creators of the TABARI coding program, to create CAMEO (Gerner et al., 2002; Schrodt, 2006). Primary among the changes it incorporated, CAMEO reduced certain categories in WEIS that could not be differentiated

by machine coding (e.g. promise/agree or warn/threaten), and then greatly expanded other categories. For example, under the original main heading of 'cooperative' behaviors, one option is 'engage in diplomatic cooperation'. This category was divided into eight sub-categories, including 'grant diplomatic recognition', 'apologize', and 'sign formal agreement'.

Updating the CAMEO dictionaries

The TABARI computer program works by recognizing sentence patterns (subject-verb-object) in the short news summaries. To do this, relevant information must be compiled in actor, verb, noun, and adjective 'dictionaries', which enable TABARI to do a 'sparse parsing' of the news lead sentences and thus identify sources (US intelligence actors), targets (191 independent states are listed in the USIIB), and verb patterns of cooperative and conflictual events.

The most critical part of the creation of the USIIB dataset, therefore, was the effort made to revise the CAMEO dictionaries in order to make them appropriate for investigating intelligence cooperation behavior in particular. The original CAMEO actor, verb, noun, and adjective dictionaries were created for different purposes and therefore did not include, for example, an actor dictionary with specifically coded US intelligence actors. New verb patterns were also necessary in order to capture intelligence interaction between US intelligence actors and foreign states. To make these modifications, we examined a random selection of 22,125 AFP leads from over the full years to be covered and, ultimately, added 1,102 nouns to the noun dictionary (e.g. INTELLIGENCE_GATHERING), 162 adjectives to the adjective dictionary (e.g. CONFIDENTIAL), 1,744 verb patterns to the verb dictionary (e.g. SHARING {INFORMATION|INTELLIGENCE}), and 2,225 actor names to the actor dictionary (e.g. SENATE_INTELLIGENCE-COMMITTEE [USACONSCI]).

Autocoding the AFP wire news leads

After developing and revising the dictionaries, the AFP wire news leads were ready to be autocoded by TABARI. During this process, 293,615 US international intelligence events were generated from a total of 369,138 AFP leads. Duplicate events were eliminated using the 'One_A_day_Filter' program included with the KEDS/TABARI software (KEDS, 2003, 2005, 2008), resulting in a total of 157,030 events, which were then ready for aggregation. An example of a news lead before and after TABARI sparse parsing and coding is as follows:

⁴ Example of an AFP lead about the US Senate Select Intelligence Committee, formatted by nexispider.pl:010828 AFPN-0001-01Pakistani President Pervez Musharraf Tuesday met Bob Graham, chairman of the US Senate select committee on intelligence, to discuss regional security and Afghanistan, the foreign ministry said.

⁵ Downloading is not as automatic as one might intuitively expect, as each search term must be entered manually – a process full of complications and changing conditions. For example, early searches within LexisNexis allowed multiple search terms for a particular actor to be entered simultaneously, whereas in the updating of the USIIB (2007–09), multiple search terms were not allowed, requiring numerous searches for each actor. The total process for the years 2000–06 took approximately 100 hours, and for 2007–09 around 45 hours.

Before TABARI coding:

Date: 19 Oct 01 Record: AFPN-0002-01

US President George W. Bush and his Chinese counterpart Jiang Zemin agreed Friday the two countries would work together in the war against terrorism through intelligence cooperation and blocking funds to terror groups.

After coding:

011019 USAPRS CHNGOV 030 (Express intent to cooperate) HIS CHINESE COUNTERPART JIANG ZEMIN AND US PRESIDENT GEORGE W. BUSH AGREED WORK TOGETHER

Validity and reliability of the USIIB dataset

A basic starting question when considering the design or use of a dataset such as this is whether event data, as measured through wire news leads from open news sources, can be considered a valid measurement of political actors' international behaviors. A number of studies have questioned any use of data ultimately stemming from newspapers, given that newspapers may be subject to both 'selection bias' (choosing particular topics to report or not to report) and 'description bias' (reporting events erroneously) (Earl et al., 2004; McCarthy, McPhail & Smith, 1996; Ortiz et al., 2005; Wilkes & Ricard, 2007). On the other hand, there have been past validity tests showing that international wire news leads can be a valid global textual source for event data studies, such as Howell & Barnes (1993) and Schrodtt & Gerner's (1994) examination of the KEDS coded event data for Middle East actors and the United States. Clearly these past studies cannot attest to the validity of other datasets like the USIIB, but they, alongside the sheer numbers of studies that have used and continue to rely on newspaper or wire leads for data collection, can be seen as providing basic background support for the initial assertion that wire news leads in principle constitute a legitimate source of data in studies of international political behavior.

The question becomes obviously much more complex when we consider the use of such data for studying intelligence behavior, which is by nature extremely secretive and may therefore go undetected by (or be denied to) public media sources. There is no denying the seriousness of this concern, and therefore future plans for the development of the USIIB include the conducting of a study comparing the data produced

by the public media sourced USIIB with data derived from commissions of inquiry reports and/or subsequently released formerly confidential documents. Perhaps most exciting is the prospect of using recently released data by WikiLeaks to run such a check of the USIIB. At present, however, the USIIB should be considered as a starting point for quantitative research on intelligence behavior, and researchers would be advised to acknowledge this issue when making claims based on findings that emerge from the USIIB.

A further question of validity arises with respect to the sources used in creating the dataset – a question which has, over the years, troubled researchers using events data (Azar, 1980; Davies & McDaniel, 1994; Howell, 1983; Huxtable & Pevehouse, 1996; Vincent, 1983; Woolley, 2000). As described earlier, the USIIB relies on AFP wire leads as its single source. Although intuitively the use of more than one news source would seem to be desirable, many major past datasets (e.g. WEIS, IDEA, PANDA) have relied on single sources, an obvious exception being the human-coded COPDAB, which used a combination of over 70 sources. Interestingly, studies have shown that a single textual source can generate as many events as multiple sources (e.g. Howell's early comparison of WEIS and COPDAB, 1983), and the numerous well-known event datasets relying on single sources make it impossible to immediately rule out their validity, though the researcher must be aware of the potential limitations of any event dataset and of its particular source. It should be noted again that future plans for the USIIB include extension of technical capacities to allow for multiple sources.

With respect to the reliability of the data in the USIIB dataset, two factors can be considered: the reliability of machine coding overall, and that of the coding scheme used for this project in particular. For the former, various studies have failed to find any significant difference in reliability between human coded and machine coded data (Hillard, Purpura & Wilkerson, 2008; King & Lowe, 2003; Schrodtt & Gerner, 1994). Indeed, machine coding has been found to be more reliable, since machines do not get tired or affected by political or cultural biases (Schrodtt, 2006); machine coding avoids problems of inconsistency when applying coding rules in different occasions (Leng, 1993); and machine coding eliminates variance between coders (Gerner et al., 1994). While it is perhaps over-optimistic to argue, as some have, that machine coding 'guarantees the reliability of the system' (Gerner et al., 2002: 2), it does seem true that machine coding eliminates the natural irregularities

that stem from having multiple human coders and instead places emphasis on the dictionaries themselves. In this case, maximum consistency in the development of the USIIB dictionaries was achieved by insuring that the process was conducted by a single, well-trained individual.

Applying the USIIB dataset

The resulting USIIB dataset, as it currently stands, holds the potential for addressing in a quantitative manner a number of important questions in the field of international relations and questions about intelligence cooperation behaviors, in particular. With minor adjustments to the dataset, still further issues may be considered.

In its current form, the USIIB dataset allows researchers to aggregate the data in terms of actor types, event types, and time frames. Thus, for example, the behaviors of separate members of the US intelligence community can be isolated, analyzed, and compared – what types of behavior does the CIA display (cooperative/conflictual)? How do these compare with the behaviors of the FBI? US intelligence actors can also be grouped and then compared in terms of intelligence cooperation behaviors – for example, those actors connected to the legislative branch as opposed to those in the executive branch: which are more active? By adjusting the ‘target actor’ from foreign nations to domestic players, questions can be addressed about *intra*-sector cooperation among or between certain US agencies. And by keeping the target actor as foreign countries or regions, researchers could explore whether the USA is engaging in intelligence cooperation with certain countries/regions more than others (see Appendix A for a brief example comparing two countries as ‘targets’ of US intelligence cooperation efforts), or with individual countries more than on a regional basis – in other words, is bilateral or multilateral cooperation more common?

Turning to the events themselves, analyses could consider which different types of cooperation or conflict prevail – for example, to what extent does the USA participate in *verbal* cooperation behaviors over *material* ones? Verbal *conflict* behaviors over material ones? Considering time frames also allows for interesting possibilities for analysis, for example, considering trends in actors’ behaviors or in different behavior types. The current dataset (2000–09) could be used to compare US intelligence cooperation behaviors pre- and post-9/11, behaviors before and after the start of the war in Iraq, or before and after the Intelligence Reform and

Terrorism Prevention Act of 2004. Researchers could consider differences between the Bush administration’s behaviors in its first term (with a majority Republican Congress) and second term (with a stronger Democratic Congress) and thus address questions about the role of the Congress in affecting intelligence cooperation behavior. Alternatively, a comparison could be made between Bush administration behaviors and those of the early Obama administration – is there a concrete reflection of the latter’s push for change in foreign policy, in terms of actual intelligence cooperation activity? Is there a rise in cooperative behaviors over conflictual ones?

Obviously, a further extension of the dataset can be made to expand the actor dictionary beyond the US intelligence community. Such an effort would allow for comparisons of practices between and among countries and regions, as well as for in-depth studies of different countries’ particular patterns of intelligence cooperation behaviors.

Are democracies more likely to cooperate with other democracies?

Following the first phase of the USIIB creation (2000–06), a sample demonstration was run exploring the question of cooperation among democracies. The conventional wisdom in intelligence research holds that successful cooperation relationships are those in which the gains each state secures from the intelligence sharing relationship are greater than the expected costs of defection by the other state. Inherent in determining the likelihood of defection by the other state is the need for a certain amount of trust between the two. The literature on intelligence cooperation has long emphasized, therefore, the issue of trust as essential in building up cooperative relationships (e.g. Lefebvre, 2003; Aldrich, 2004; Clough, 2004; Behnsahel, 2006) and gone on to presume that trust between countries is increased when they share common governing systems – specifically, both being democracies.

Turning to the literature exploring the connection between states’ governing systems and intelligence cooperation, a minority of works have suggested an increase in post-9/11 US cooperation with non-democratic countries (Rudner, 2004); however, the predominant view remains that the international intelligence behavior of democratic countries is significantly different from that of non-democratic countries (Andrew, 2004) and that the *non*-democratic characteristics of states can prevent the fully efficient functioning of intelligence cooperation (e.g. Shpiro, 2001; Sims, 2006; Tan & Ramakrishna, 2004).

Table II. Intelligence cooperation and regime type

Variables	B	<i>Robust</i> SE	<i>p</i> > <i>z</i>
<i>Regime type variable</i>			
Democracy (1–7 scale of political rights)	–0.175	0.024	0.000
ln alpha	0.037	0.064	
Alpha	1.037	0.069	
Wald	1502.3		0.000
Log pseudo Likelihood	–3879.33		
N	1316		

Negative binomial regression coefficients with robust standard errors. No multicollinearity is detected.

Statistically significant coefficients (at the $p > .05$ level) in two-tailed test. Iraq and Afghanistan are outliers and are excluded from the analysis.

Using the USIIB dataset, the question was therefore posed whether, in the 2000–06 period, the United States was more likely to cooperate on intelligence matters with democratic states. As seen in Table II, the results showed that in this period there was a negative relationship between the democracy rating variable of foreign nations (established using Freedom House measurements) and the likelihood of US engagement in intelligence cooperation with that state. In all the negative binomial regression models, democracy was found to be statistically significant, but the direction of the relationship was negative. In other words, the United States was more likely to engage in intelligence cooperation with *less* democratic states.

Various explanations may be offered for such a finding: for example, the possibility that cooperation with non-democracies is simply more newsworthy than the established and institutionalized cooperation between the USA and other democracies – though even if this were so, the results would still be interesting since they reveal an unexpected cooperation behavior. It is also possible, however, that this finding reflects the nature of the perceived predominant threat post-9/11 and the increased value placed on intelligence coming from countries of the Middle East, South Asia, or Central Asia, which in many cases have lower democratic ratings than the USA's more traditional intelligence cooperation partners. The USIIB data may be providing evidence, therefore, that intelligence cooperation decisionmaking is more dependent on pragmatic calculations of the threat than on preferences for like-minded countries with common governance practices.

Conclusion

This three-year project sought to create a dataset that would quantify the international intelligence behaviors

of the United States. While previously designed coding schemes have been prepared for categorizing conflictual/cooperative behaviors in IR in general (e.g. WEIS and CAMEO) and software exists that can produce event datasets ready for analysis with these coding schemes (e.g. KEDS and TABARI), the USIIB dataset represents the first attempt at adapting coding schemes to collect event data specifically focused on cooperative behaviors of intelligence actors. Through substantial revisions and additions to the actor, verb, noun, and adjective dictionaries of the CAMEO coding system, and the collecting of AFP news leads from the years 2000–09, the USSIB provides a dataset that can be used for quantifying responses to questions about intelligence cooperation actors and their behavioral patterns.

Transnational challenges include a variety of burgeoning global concerns, from weapon smuggling and human trafficking to environmental crises and the spread of disease. Because of the cross-border nature of such threats, the key to countering them is to develop broad intelligence cooperation among agencies and security establishments worldwide and within individual countries. Understanding the current practices of global intelligence actors and being able to explore these practices in a quantifiable manner will contribute towards defining and developing more effective practices. The USIIB dataset is a first step in allowing for such research to be conducted.

Appendix A

<i>US cooperative interaction with Afghanistan</i>		<i>US cooperative interaction with Algeria</i>	
<i>Years</i>	<i>Cooperation events</i>	<i>Years</i>	<i>Cooperation events</i>
2000	7	2000	7
2001	99	2001	9
2002	111	2002	6
2003	66	2003	9
2004	56	2004	4
2005	55	2005	4
2006	42	2006	8
2007	49	2007	12
2008	67	2008	6
2009	171	2009	4

Acknowledgements

The authors would like to thank the three anonymous reviewers who provided very valuable comments on earlier drafts of this article.

Replication Data

Replication data for this article are available at <http://www.prio.no/jpr/datasets>. The complete dataset and further information on the project are available on the KIVAM website at www.kivam.org but will require downloading additional software, such as KEDS_Count (<http://eventdata.psu.edu/software.dir/utilities.html>), to quantify and aggregate the data before conducting statistical analysis. Researchers with specific questions about the data may contact Musa Tuzuner (musatzuner@yahoo.com).

References

- Aldrich, Richard J (2004) Transatlantic intelligence and security cooperation. *International Affairs* 80(4): 731–753.
- Aldrich, Richard J (2009) US–European intelligence co-operation on counter-terrorism: Low politics and compulsion. *British Journal of Politics and International Relations* 11(1): 122–139.
- Andrew, Christopher (2004) Intelligence, international relations and ‘under-theorisation’. *Intelligence and National Security* 19(2): 170–184.
- Azar, Edward E (1980) The Conflict and Peace Data Bank (COPDAB) project. *Journal of Conflict Resolution* 24(1): 143–152.
- Bensahel, Nora (2006) A coalition of coalitions: International cooperation against terrorism. *Studies in Conflict and Terrorism* 29(1): 35–49.
- Bond, Doug; Joe Bond, Oh Churl, J Craig Jenkins & Charles Lewis Taylor (2003) Integrated Data for Events Analysis (IDEA): An event typology for automated events data development. *Journal of Peace Research* 40(6): 733–745.
- Clough, Chris (2004) Quid pro quo: The challenges of international strategic intelligence cooperation. *International Journal of Intelligence and Counterintelligence* 17(4): 601–613.
- Davies, John L & Chad K McDaniel (1994) A new generation of international event data. *International Interactions* 20(1–2): 55–78.
- Dumbrell, John (2009) The US–UK special relationship: Taking the 21st century temperature. *British Journal of Politics and International Relations* 11(1): 64–78.
- Earl, Jennifer; Andrew Martin, John D McCarthy & Sarah A Soule (2004) The use of newspaper data in the study of collective action. *Annual Review of Sociology* 30: 65–80.
- Engene, Jan Oskar (2007) Five decades of terrorism in Europe: The TWEED dataset. *Journal of Peace Research* 44(1): 109–121.
- Franzosi, Roberto (1987) The press as a source of sociohistorical data. *Historical Methods* 20(1): 5–16.
- Gerner, Deborah J; Philip A Schrodt, Ronald A Francisco & Judith L Weddle (1994) Machine coding of event data using regional and international sources. *International Studies Quarterly* 38(1): 91–119.
- Gerner, Deborah J; A Schrodt Philip, Yilmaz Ömür & Abu-Jabr Rajaa (2002) Conflict and Mediation Event Observations (CAMEO): A new event data framework for the analysis of foreign policy interactions. Annual Meeting of the International Studies Association, New Orleans, LA (<http://web.ku.edu/keds/papers.dir/gerner02.pdf>) (<http://eventdata.psu.edu/papers.dir/Gerner.APSA.02.pdf>).
- Hillard, Dustin; Stephen Purpura & John Wilkerson (2008) Computer-assisted topic classification for mixed-methods social science research. *Journal of Information Technology & Politics* 4(4): 31–46.
- Howell, Llewellyn D (1983) A comparative study of the WEIS and COPDAB data sets. *International Studies Quarterly* 27(2): 149–159.
- Howell, Llewellyn D & Gillian Barnes (1993) Event data for region-specific interactions: A research note on source coverage. In: Richard L Merritt, Robert G Muncaster & Dina A Zinnes (eds) *International Event-Data Developments: DDIR Phase II*. Ann Arbor, MI: University of Michigan Press, 45–54.
- Hussain, Imtiaz (2009) Post-9/11 Canada–US security integration: Of the butcher, the baker, and the intelligence-policy maker. *American Review of Canadian Studies* 39(1): 38–51.
- Huxtable, Phillip A & Jon C Pevehouse (1996) Potential validity problems in event data collection. *International Studies Notes* 21(2): 8–19.
- Jenkins, J Craig & Doug Bond (2001) Conflict-carrying capacity, political crisis, and reconstruction: A framework for the early warning of political system vulnerability. *Journal of Conflict Resolution* 45(1): 3–31.
- KEDS Project Software (2003) Text filter: nexispider.pl. web.ku.edu/keds (<http://web.ku.edu/keds/software.dir/filters.html>) (<http://eventdata.psu.edu/software.dir/filters.html>).
- KEDS Project Software (2005) Utility programs: One-A-Day_Filter. web.ku.edu/keds (<http://web.ku.edu/keds/software.dir/utilities.html>).

- KEDS Project Software (2008) Utility programs: Keds_Count (<http://web.ku.edu/keds/software.dir/utilities.html>) (<http://eventdata.psu.edu/software.dir/utilities.html>).
- King, Gary & Will Lowe (2003) An automated information extraction tool for international conflict data with performance as good as human coders: A rare events evaluation design. *International Organization* 57(3): 617–642.
- Koopmans, Ruud & Dieter Rucht (1999) Protest event analysis – where to now? *Mobilization* 4: 123–130.
- Lander, Stephen (2004) International intelligence cooperation: An inside perspective. *Cambridge Review of International Affairs* 17(3): 481–493.
- Lefebvre, Stéphan (2003) The difficulties and dilemmas of international intelligence cooperation. *International Journal of Intelligence and Counterintelligence* 16(3): 527–542.
- Leng, Russell J (1993) Automated and machine-assisted coding of event data: The BCOW approach. In: Richard L Merritt, Robert G Muncaster & Dina A Zinnes (eds) *International Event-Data Developments: DDIR Phase II*. Ann Arbor, MI: University of Michigan Press, 113–123.
- Matei, Florina Cristiana (2009) The challenges of intelligence sharing in Romania. *Intelligence and National Security* 24(4): 574–585.
- McCarthy, John D; Clark McPhail & Jackie Smith (1996) Images of protest: Dimensions of selection bias in media coverage of Washington demonstrations, 1982 and 1991. *American Sociological Review* 61(3): 478–499.
- McClelland, Charles A (1978) *World Event/Interaction Survey (WEIS) 1966–1978 (ICPSR 5211)*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research.
- Munton, Don (2009) Intelligence cooperation meets international studies theory: Explaining Canadian operations in Castro's Cuba. *Intelligence & National Security* 24(1): 119–138.
- Olzak, Susan (1989) Analysis of events in studies of collective actions. *Annual Review of Sociology* 15: 119–141.
- Ortiz, David G; Daniel J Myers, Eugene N Walls & Maria-Elena D Diaz (2005) Where do we stand with newspaper data? *Mobilization: An International Quarterly* 10(3): 397–419.
- Dieter, Rucht; Ruud Koopmans & Friedhelm Neidhardt (1999) *Acts of Dissent*. New York: Rowman & Littlefield.
- Rudner, Martin (2004) Hunters and gatherers: The intelligence coalition against Islamic terrorism. *International Journal of Intelligence and Counterintelligence* 17(2): 193–230.
- Schrodt, Philip A (2006) TABARI: Textual Analysis by Augmented Replacement Instructions Version 0.5. KEDS: Kansas Event Data System. (<http://web.ku.edu/keds/software.dir/tabari.html>) (<http://eventdata.psu.edu/software.dir/tabari.html>).
- Schrodt, Philip A & Deborah J Gerner (1994) Validity assessment of a machine-coded event data set for the Middle East, 1982–92. *American Journal of Political Science* 38(3): 825–854.
- Schrodt, Philip A; Shannon G Davis & Judy L Weddle (1994) Political science: KEDS – A program for the machine coding of event data. *Social Science Computer Review* 12(3): 561–588.
- Shellman, Stephen M (2008) Quantifying violence and nonviolence: Terrorism & Political Violence events data sets. *Electronic Newsletter of the ECPR-SG on Extremism & Democracy* 9(2) (http://www.tufts.edu/~dart01/extremismanddemocracy/newsletter/Article7_4.htm).
- Shapiro, Shlomo (2001) The communication of mutual security: Frameworks for European–Mediterranean intelligence sharing (<http://www.nato.int/acad/fellow/99-01/shapiro.pdf>).
- Sims, Jennifer E (2006) Foreign intelligence liaison: Devils, deals, and details. *International Journal of Intelligence and Counterintelligence* 19(2): 195–217.
- Strachan-Morris, David (2009) The future of civil–military intelligence cooperation based on lessons learned in Iraq. *Intelligence and National Security* 24(2): 257–274.
- Svendsen, Adam DM (2009) Connecting intelligence and theory: Intelligence liaison and international relations. *Intelligence and National Security* 24(5): 700–729.
- Tan, See Seng & Kumar Ramakrishna (2004) Interstate and intrastate dynamics in Southeast Asia's war on terror. *SAIS Review* 24(1): 91–105.
- Vincent, Jack E (1983) WEIS and COPDAB: Correspondence problems. *International Studies Quarterly* 27: 161–168.
- Wilkes, Rima & Danielle Ricard (2007) How does newspaper coverage of collective action vary? Protest by indigenous people in Canada. *Social Science Journal* 44(2): 231–251.
- Woolley, John T (2000) Using media-based data in studies of politics. *American Journal of Political Science* 44(1): 156–173.
- 9/11 Commission Report (2002) (<http://www.gpoaccess.gov/911/pdf/fullreport.pdf>).

ERSEL AYDINLI, b. 1967, PhD in Political Science (McGill University, 2001); Associate Professor, Bilkent University (2006–); current main interests: transnational security, non-state actors, and homegrown IR theorizing. Most recent book: *Emerging Transnational (In)Security Governance: A Statist-Transnationalist Approach* (Routledge, 2010).

MUSA TUZUNER, b. 1968, PhD in Political Science (Kent State University, 2009); Assistant Professor,

Turkish National Police Academy (2009–); Founding Director, Intelligence Studies Research Center (2009–) & Intelligence Studies Master Program (2010–), TNP Academy; current main interests: national and international security behavior, theorizing Turkish foreign policy, and international event data developments. Most recent book: *Intelligence Cooperation Practices in the 21st Century: Towards a Culture of Sharing* (IOS Press, 2010).