



Universiteit  
Leiden  
The Netherlands

## **Focusdata: foreign policy through language and sentiment**

Fisher, S.; Klein, G.R.; Codjo, J.

### **Citation**

Fisher, S., Klein, G. R., & Codjo, J. (2022). Focusdata: foreign policy through language and sentiment. *Foreign Policy Analysis*, 18(2).  
doi:10.1093/fpa/orac002

Version: Publisher's Version  
License: [Creative Commons CC BY 4.0 license](#)  
Downloaded from: <https://hdl.handle.net/1887/3486795>

**Note:** To cite this publication please use the final published version (if applicable).

# FOCUSdata: Foreign Policy through Language and Sentiment

SCOTT FISHER

*New Jersey City University, USA*

GRAIG R. KLEIN 

*Leiden University, the Netherlands*

AND

JUSTE CODJO

*New Jersey City University, USA*

Countries routinely translate official statements and state media articles from native languages to English. Over time, these articles provide a window into what each government is trying to portray to the world. The FOCUSdata Project provides years' worth of text and language sentiment ratings for hundreds of thousands of articles from state media and ministry of foreign affairs' websites from North Korea, China, Russia, and Iran. Information is an important foreign policy tool and national security strategists analyze how it influences the attitudes and behaviors of foreign audiences. This article introduces the FOCUSdata Project and shows how the sentiment data provide unique abilities to analyze Russia's and Iran's reactions to US policies and events and NGO human rights campaigns. Evaluating countries' official narratives improves understanding of government signals to outside actors, reactions to crises and foreign policy tools, and interests regarding (un)favorable developments. Governments' sentiment provides unique explanatory power.

Por lo general, los países traducen las declaraciones oficiales y los artículos de los medios de comunicación estatales de su lengua materna al inglés. Estos artículos ofrecen, con el tiempo, una ventana a lo que cada gobierno intenta mostrar al mundo. El Proyecto FOCUSdata ofrece años de clasificaciones de sentimiento de textos e idiomas de cientos de miles de artículos de medios de comunicación estatales y sitios web de ministerios de asuntos exteriores de Corea del Norte, China, Rusia e Irán. La información es una importante herramienta de política exterior, y los estrategas de seguridad nacional analizan cómo influye en las actitudes y comportamientos de los destinatarios extranjeros. Este artículo presenta el Proyecto FOCUSdata y muestra cómo los datos de sentimiento proporcionan capacidades únicas para analizar las reacciones de Rusia e Irán a las políticas y los eventos de Estados Unidos, así como a las campañas de derechos humanos de las ONG. La evaluación de las narrativas oficiales de los

---

Scott Fisher is a Security Studies Professor at New Jersey City University, United States. His research interests include information warfare, big data in national security, and open-source intelligence.

Graig R. Klein is an Assistant Professor at the Institute of Security and Global Affairs at Leiden University, the Netherlands. His research explores the instrumentality of political violence, primarily terrorism, protests, and government responses, to understand how dissident-government interactions inform tactical and strategic evolution in conflict processes, international security, and national security. He holds a Ph.D. in Political Science from Binghamton University, NY, United States.

Juste Codjo is an Assistant Professor of Professional Security Studies at New Jersey City University, United States. His research revolves around political violence, civil conflict, African security, political institutions in developing states, US national security, and US-Africa strategic relations, and has appeared in *African Studies Review* and *International Studies Review*.

Fisher, Scott et al. (2022) FOCUSdata: Foreign Policy through Language and Sentiment. *Foreign Policy Analysis*, <https://doi.org/10.1093/fpa/orac002>

© The Author(s) (2022). Published by Oxford University Press on behalf of the International Studies Association. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

países mejora la comprensión de las señales del gobierno a los actores externos, las reacciones ante las crisis y las herramientas de política exterior, así como los intereses en relación con los acontecimientos (des)favorables. El sentimiento de los gobiernos proporciona un poder explicativo excepcional.

Des pays traduisent régulièrement leurs déclarations officielles et articles des médias d'État de leur langue maternelle vers l'anglais. Au fil du temps, ces articles offrent une fenêtre sur ce que chaque gouvernement essaie de présenter au monde. Le projet FOCUSdata apporte la valeur d'années d'évaluations des sentiments et du langage des textes de centaines de milliers d'articles des médias d'État et des sites web des ministères des affaires étrangères de Corée du Nord, de Chine, de Russie et d'Iran. L'information est un outil de politique étrangère important et les stratèges chargés de la sécurité nationale analysent la manière dont elle influence les attitudes et les comportements des publics étrangers. Cet article présente le projet FOCUSdata et montre la façon dont les données sur les sentiments offrent des possibilités uniques d'analyser les réactions de la Russie et de l'Iran aux événements et politiques des États-Unis ainsi qu'aux campagnes des ONG en faveur des droits de l'homme. L'évaluation des discours officiels des pays améliore notre compréhension des signaux émis par les gouvernements à l'attention d'acteurs extérieurs, des réactions aux crises, des outils de politique étrangère et des intérêts concernant les évolutions (dé)favorables. Les sentiments des gouvernements offrent un pouvoir explicatif unique.

---

## Introduction

Every year, North Korean, Russian, Chinese, and Iranian state media take the time and trouble to write, translate and publish thousands of English-language articles. The foreign ministries of these four countries do the same for select speeches, communiques, and other communications. These data provide a window into what each government is trying to portray to the world: common versus ignored topics, narrative frameworks, and reactions to outside activities. Although these and other similar data can be helpful to analysts and researchers interested in public diplomacy, little has been done to collect them across time and space.<sup>1</sup>

The FOCUSdata Project (Fisher and Klein 2020), henceforth FOCUSdata, fills this gap by collecting years' worth of articles from state media and Ministry of Foreign Affairs (MOFA) websites, offering the collections as searchable databases, and providing article and topic sentiment.<sup>2</sup> Now, for example, anyone interested in North Korea can search 85,000 unique articles from the Korean Central News Agency (KCNA—North Korea's official state media outlet) from October 1, 2008–February 27, 2020, and explore what was reported, how topics were framed, and articles' sentiment.

In the remainder of the article, we first discuss how state media and communication generate and provide information as a foreign policy tool and its

---

<sup>1</sup>The field of public diplomacy, which is primarily concerned with how governments and non-state actors engage in purposive communication with foreign audiences in support of foreign policy objectives, has received growing attention in recent years (Sevin, Metzgar, and Hayden 2019). Yet, measurement of the elements that constitute the substance of international actors' public diplomacy efforts remains a challenge for scholars and practitioners (Sevin and Ingenhoff 2018). For additional information on the public diplomacy scholarship, see Jones and Mattiacci (2019), Huang (2016), and Sevin (2017).

<sup>2</sup>FOCUSdata Project (Fisher and Klein 2020) is funded by an Intelligence Community Center of Academic Excellence (IC CAE) grant from the Office of the Director of National Intelligence.

relevance to national security through the DIME framework (henceforth, DIME).<sup>3</sup> We then introduce the FOCUSdata by describing the data collection process, sentiment analysis, descriptive statistics, and highlight interesting preliminary findings. We demonstrate FOCUSdata's validity through two succinct case applications; one discussing Russian MOFA sentiment changes in reaction to US sanctions and NGO human rights campaigns and another describing the Islamic Republic News Agency's (IRNA) sentiment when reporting on the United States during the Obama and Trump administrations. Last, we summarize FOCUSdata's added value to policy development and national security.

### Information as a Foreign Policy Tool

We draw on the concept of information as an important foreign policy tool and introduce the FOCUSdata to help conduct related research. Broadly speaking, information is an element of "power over opinion" (Carr 1946) or "soft power" and "smart power" (Nye 1990, 2008, 2009). From a national security perspective, information is one of four critical instruments or elements of national power known as DIME: Diplomacy, Information, Military, and Economy.<sup>4</sup> Governments commonly rely on information, either in isolation or in combination with other DIME instruments, to achieve foreign policy or national security objectives. For instance, the United States uses its information power to "communicate America's story, its image, to the world" (Worley 2012, 278). Similarly, other countries rely on their information tools to influence the attitudes and behaviors of foreign audiences through public communication (Murphy 2012; Worley 2012). It is therefore important for scholars to pay attention to how countries use the information for national security and foreign policy purposes.

Governments rely on a variety of information tools including, but are not limited to, public diplomacy, international broadcasting, public affairs, and military information or psychological operations (Jones 2006; Murphy 2012; Worley 2012). For instance, international broadcasting agencies are commonly used by their governments to shape public opinion abroad and create favorable conditions for the implementation of their foreign policy (Mattiacci and Jones 2020; Snow and Cull 2020). Governments also rely on their public affairs and state media to persuade their domestic audiences about national security actions abroad (Jones 2006). Because governments control their information tools, it can be assumed that the communicated topics, language, and sentiment match an overarching national strategy.

The FOCUSdata features two information tools that are widely used, with varying degrees, by governments across the world to shape perception and behavior abroad. One is overt diplomatic information activities aimed at foreign audiences including communiqués by agencies in charge of foreign affairs, press conferences by diplomats, and other similar communication activities designed to affect hearts and minds in the international environment. The second is news articles in state-influenced, or controlled, media. This dimension of information power revolves around a government's ability to convey its message through news articles that are transmitted directly to a foreign audience rather than through an embassy (Jones 2006). In countries that lack freedom of speech and/or press,

<sup>3</sup> DIME is an acronym used by national security experts to describe what they consider to be the four elements of national power: diplomacy, information, military, and economy.

<sup>4</sup> The DIME concept emerged during the Cold War and has remained a central framework for defense and security strategy and policy (Joint Chiefs of Staff 2018). Counter-terrorism experts prefer the use of DIMEFIL arguing that the four traditional instruments emphasized in DIME fall short in both asymmetric warfare and hybrid warfare because they cannot sufficiently influence non-state actors seeking to defy traditional methods of state power projection (Oscarson 2017). They have pressed for the inclusion of three additional instruments: finance, intelligence, and law enforcement (Shellman, Levey, and Leonard 2011).

state-owned/controlled media outlets are an important branch of the government's information toolkit.

While these information tools and practices are cross-national, the impetus of the FOCUSdata is understanding how major US adversaries use information and construct narratives. This resulted in grant funding to collect primary documents (English language official MOFA communication and state-controlled news articles) from China, Russia, North Korea, and Iran. As none of the four countries use English as an official language, we consider the use of English a signal of intent to engage with the international community.

### Empirical Relevance of Informational Tools

Information as an element of national power has sparked an increasing interest among national security practitioners and international relations (IR) scholars leading to its widespread inclusion in strategic studies curricula at US military and academic institutions, especially DIME. It is also often the focus of analyses. For instance, DIME is used to examine a range of policy issues, including border disputes in Africa (Majani 2019) and the ability of the United States to coordinate its strategy in the executive and congressional branches of government (Armstrong 2019). Other practitioners have relied on the same construct to call for policy change regarding how information tools for national power are used in the United States, arguing that “we are falling behind our potential adversaries when using the information domain for national security” (Kozloski 2009; see also Morgan 2019).

The information instrument of power has also caught the attention of researchers. In a project funded by the Defense Advanced Research Projects Agency and the National Science Foundation, the Violent Intranational Political Conflict and Terrorism (VIPCAT) Research Laboratory at William & Mary relied on DIME to collect data and test the effects of the four instruments of US national power on political conflict abroad and how they could be used to achieve foreign policy objectives (Shellman, Levey, and Leonard 2011).

Others have focused solely on the information element. Many of them tend to investigate the influence of transnational media in world politics, such as anti-American sentiment abroad, an issue that IR scholars and US policymakers have long been interested in (Nisbet et al. 2004; Katzenstein and Keohane 2007; Jamal et al. 2015). This type of research often revolves around analyzing the relationship between foreign media exposure and opinion about the United States. For instance, Nisbet and Myers (2011) empirically evaluate the influence of transnational Arab TV channels such as Al-Jazeera on opinion formation about the United States in the Middle East and find a significant correlation between exposure to transnational Arab television and anti-American sentiment. Research on the influence of international news media has since expanded to other outcomes such as agenda-setting. Analyzing a dataset of 4,708 online news sources from 67 countries in 2015, Guo and Vargo (2017) demonstrate that wealthy countries are both the focus and narrator of global news and media. Such sources of opensource intelligence are a key component of national security intelligence collection and analysis, serve as a high-value “tool of first resort,” and could vault specialized branches of the intel-community, like the CIA's Open Source Enterprise, into critical enterprises (Clapper et al. 2020, 7).

With the rising policy and research interest in DIME, in general, and transnational media in particular, an increase in the demand for empirical data about the information environment that governments rely on to achieve their strategic goals is inevitable. To meet such demands, efforts must be made to collect and publish data on the nature, the content, and the activities of organizations dedicated to collecting, processing, disseminating, or acting on information across the world. In this context, the FOCUSdata is of great value. The availability of these data—the

raw text content and corresponding sentiment scores—is a valuable resource for researchers interested in measuring the information component of DIME and or analyzing anti-American sentiment.

### FOCUSdata

The FOCUSdata collected articles from ten sources representing four countries: two sources each from North Korea and Iran, and three each from Russia and China. Sources were selected based on government affiliation (i.e., foreign ministries) and status as official or state media that regularly publish in English. The documents were collected by “scraping” English-language articles from the associated websites—every article available on the site at the time of the scrape was collected into a spreadsheet with dates, headlines, and article text. We conducted some scraping ourselves using the software tools Parsehub (<https://www.parsehub.com/>) and ScrapeStorm (<https://www.scrapestorm.com/>); other sites (typically very large or difficult to access websites) were scraped by hiring a service: Marquee Data (<https://marqueedata.com/>) or Scrapinghub (<https://www.scrapinghub.com/>). We found these tools and services helpful, but there are dozens (if not hundreds) of similar capabilities on the market; we recommend selecting based on your needs, technical capabilities, and budget (Munzert et al. 2015; Monroe and Schrodt 2017; Wilkerson and Casas 2017). Once the source documents were collected, duplicate entries were removed and then sentiment analysis tools, primarily Google Natural Language (GNL) were applied to the media articles and MOFA documents. The sentiment analysis, ranging from very negative to very positive, provides an additional tool for those attempting to understand how governments perceive or frame a topic.

Social media, radio and TV broadcasts, and other data sources are valuable to researchers, but not the focus of this research. Official statements and media articles provide significantly more text and data to analyze than 140-character tweets and are not susceptible to bots and trolls in the same manner as social media posts and commentary. The documents included in FOCUSdata are ostensibly vetted and approved by the government without having the safety net of plausible deniability social media posts can provide.

In table 1, we summarize the collected/scraped data. The use of foreign media and official websites presented some limitations in the data collection. The Iranian MOFA site had recently been updated and English language content was only available dating back to January 1, 2018. The North Korean media websites, especially KCNA’s, were difficult to scrape as access was limited from a US IP-address; therefore, we used a virtual private network (VPN) that allowed the user or scraping software to appear to be located in Japan, enabling access to KCNA’s [kcnaco.jp](http://kcnaco.jp) website.

### *Sentiment Analysis*

Sentiment analysis is “the computational treatment of opinion, sentiment, and subjectivity in text” (Pang and Lee 2008, 6; Taboada 2016). In a sense, it is turning text into quantifiable data. By analyzing the sentiment of a report, or even better, tens of thousands of reports, researchers can quantify opinion on a topic (Ravi and Ravi 2015; Liu 2020). This offers researchers a measurement tool for comparing government reactions to foreign policy tools based on DIME categories: diplomacy, information, military, and economics.

Once our data were scraped, we applied two tools, MeaningCloud (<https://www.meaningcloud.com/>) and GNL (<https://cloud.google.com/natural-language>), to conduct sentiment analysis at the article level.

**Table 1.** Summary of FOCUSdata sources

Source	Website	Date range	<i>N</i>	Scrape date
Russia				
<i>Russia Today (RT)</i>	<a href="https://www.rt.com/">https://www.rt.com/</a>	June 21, 2006– January 9, 2020	211,923	January 2020
<i>Sputnik</i>	<a href="https://sputniknews.com/">https://sputniknews.com/</a>	June 2, 2019– January 16, 2020	28,727	January 2020
MOFA	<a href="https://www.mid.ru/">https://www.mid.ru/</a>	May 20, 2004– January 15, 2020	21,372	January 2020
Iran				
<i>Islamic Republic News Agency (IRNA)</i>	<a href="https://en.irna.ir/">https://en.irna.ir/</a>	December 22, 2011– February 15, 2020	103,983	February 2020
MOFA	<a href="https://en.mfa.ir/">https://en.mfa.ir/</a>	January 1, 2018– February 15, 2020	1,671	February 2020
China				
<i>People's Daily</i>	<a href="http://en.people.cn/">http://en.people.cn/</a>	June 29, 2007– February 22, 2020	483,268	February 2020
<i>Global Times</i>	<a href="https://www.globaltimes.cn/">https://www.globaltimes.cn/</a>	April 9, 2009– February 21, 2020	597,156	February 2020
MOFA	<a href="https://www.fmprc.gov.cn/mfa_eng/">https://www.fmprc.gov.cn/mfa_eng/</a>	November 15, 2000– December 25, 2019	762	December 2019
North Korea				
<i>Korean Central News Agency (KCNA)</i>	<a href="http://www.kcna.co.jp">www.kcna.co.jp</a>	October 1, 2008– February 27, 2020	85,313	February 2020
<i>Rodong Sinmun</i>	<a href="http://www.rodong.rep.kp/">http://www.rodong.rep.kp/</a>	January 2, 2018– December 31, 2019	7,102	January 2020

MeaningCloud works as an Excel add-in, making it easy for non-technical users—anyone who can use Excel can conduct sentiment analysis. GNL requires greater technical knowledge, for example, the ability to make Application Programming Interface (API) calls to Google's tools. The output of the two tools is also different. MeaningCloud returns values along a Likert scale: very positive, positive, neutral, negative, very negative, and none. GNL returns a number from  $-1.0$  (very negative) to  $1.0$  (very positive), with zero as neutral.

The two sentiment analysis systems also function in different ways. MeaningCloud attaches polarity to entities and concepts, identifies point(s) of view, with machine learning running “in the background” (Gonzalez 2021). GNL, with its massive size and data access, offers both pre-trained (our selection) and trainable models for sentiment analysis. Both systems can provide results at the article level and that is the level selected for our research; articles contain a variety of sentiments, but the value returned (and used for our analysis) is the article's overall sentiment. It is important to note that the goal of this research is not to compare or “prove” one method of sentiment analysis is superior to another; our goal is to apply existing, off-the-shelf capabilities to questions of political science and international/national security.

We found both tools valuable and would recommend either to other researchers. Those looking for additional options can find commercial sentiment analysis tools from Amazon, IBM, and Microsoft, try programming in R or Python, or search reviews of analytical text programs at resources like QuantText (<https://quanttext.com/>). Sentiment is only one type of text analysis. Because the

text of the original documents is included in FOCUSdata's raw datafiles, content analysis on a topic of information source can be conducted. It also means that word counts, word and phrase pattern analysis, language analysis, pronouns, emotions aside from a positive–negative continuum, and other language analysis tools and methods that have been used in similar research can be applied to FOCUSdata (e.g., National Research Council 2011; Bodine-Baron et al. 2016; Heller 2018; Barnum and Lo 2020; Gray and Baturu 2021).<sup>5</sup>

Select articles and their associated sentiment scores were analyzed by project researchers for both content and sentiment accuracy, but not on a systematic basis. One of the goals of making our research data available, including sentiment scores from both systems, is for other researchers to test our findings using their preferred methodology. Our sentiment scores are meant to assist our research by serving as “a score”, not “the score” for the articles. We earnestly await the analysis of our data by other researchers including those harnessing human-in-the-loop coding and analysis.

Finally, of our datasets, the following were analyzed using MeaningCloud: KCNA, Rodong Sinmun, Russia MOFA, and Iran MOFA. GNL tools were used for *Russia Today*, *Sputnik*, *Global Times*, *People's Daily*, China MOFA, and Iran's IRNA. The larger datasets were analyzed using Google's cloud-based tools; some of the smaller datasets were analyzed locally using MeaningCloud's Excel plug-in.

#### Public Data Access

To make our data widely accessible, we loaded the datasets into Tableau data visualization software (<https://www.tableau.com/>), then displayed the output on the project's website and on Tableau Public. Users can view baselines of article numbers and sentiment, conduct basic queries on keyword usage and sentiment over time, and create interactive visualizations suitable for both analysis and presentations. We also share the Iran and North Korea raw data (as spreadsheets) on Harvard Dataverse. Those interested in the raw data files for China and Russia should contact the authors.

#### Trends in FOCUSdata

MOFA sentiment is typically more positive, perhaps due to the diplomatic origin, than state media sentiment when discussing similar topics (Fisher 2019). MOFA communications rarely register negative sentiment, whereas state-controlled or influenced media often uses negative sentiment. For example, 1.1% of Chinese MOFA communiqués had negative sentiment, whereas articles in *Peoples' Daily* and *Global Times* were negative 22.5% and 25.3% of the time, respectively. The variation in sentiment suggests that significant nuances occur between official diplomacy and “government-speak” and what state-sanctioned media is approved to say. Further research in this area of discourse and sentiment differences is important.

For summary statistics, see table 2, and the following visualizations, the Likert scale produced by MeaningCloud is converted to a numerical scale ranging from –1.0 to 1.0 (very negative to very positive) that matches GNL's scale; the category of None (no sentiment detected) is dropped. We also include a recalculated Chinese MOFA converting the “continuous” GNL score into the categorical scoring produced by MeaningCloud.

From table 2, it is clear that Iranian MOFA has a substantively lower, but still positive, mean sentiment score. The Iranian MOFA data time period, as previously described, is shorter, which could result in the larger standard deviation, but this

<sup>5</sup>Raw text data are available on FOCUSdata's website (<https://focusdataproject.com/>) and Harvard Dataverse (<https://dataverse.harvard.edu/dataverse/focusdataproject>).

**Table 2.** FOCUSdata sentiment summary statistics

	N	Mean	Standard deviation	Minimum	Maximum
<b>Iran</b>					
<i>Islamic Republic News Agency (IRNA)</i>	102,505	0.069	0.122	-1	1
MOFA	1,647	0.139	0.449	-1	1
<b>Russia</b>					
<i>Russian Times (RT)</i>	211,864	0.031	0.115	-1	1
<i>Sputnik</i>	28,727	0.043	0.093	-0.8	1
MOFA	20,409	0.322	0.364	-1	1
<b>North Korea</b>					
<i>Korean Central News Agency (KCNA)</i>	83,308	0.274	0.467	-1	1
<i>Rodong</i>	6,779	0.448	0.351	-1	1
<b>China</b>					
<i>Peoples' Daily</i>	475,294	0.064	0.107	-1	1
<i>Global Times</i>	590,948	0.056	0.113	-1	1
MOFA	762	0.265	0.132	-0.3	0.7
MOFA (conversion scale)	762	0.465	0.162	-0.5	1

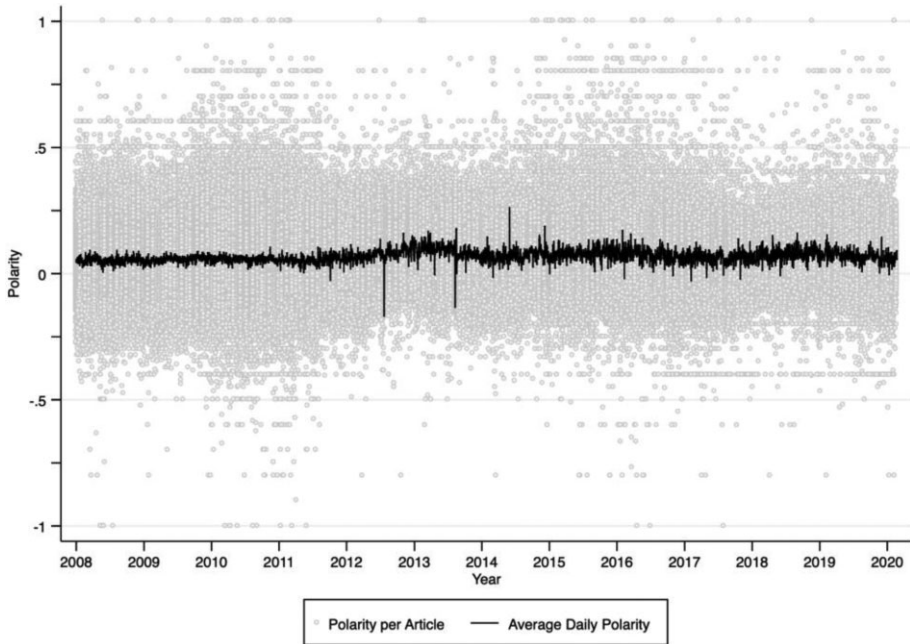
does not appear to be the cause of lower sentiment as we compared China and Russia's MOFAs sentiment in the same shortened time period (0.260 and 0.355, respectively).<sup>6</sup> The difference may be related to the content of the statements as they often discussed a rival—Israel—but Russian MOFA often discussed Kiev and North Atlantic Treaty Organization (NATO). Chinese MOFA more frequently discussed cooperation, trade, and global issues than the other two, but this did not make the average sentiment more positive than Russia's. Perhaps, as we discuss in the Iranian case application below, the less positive Iranian MOFA sentiment correlates, at least partially, with changing dynamics in its strategic rivalry with the United States as the temporal domain overlaps with the United States' policy of maximum pressure.

The summary statistics do not capture variation across time and there is, unsurprisingly, substantial variation in sentiment across state media outlets. To illustrate the variance around the average sentiment over time, we created several visualizations overlaying each media outlet's *daily* average sentiment on their per article sentiment.

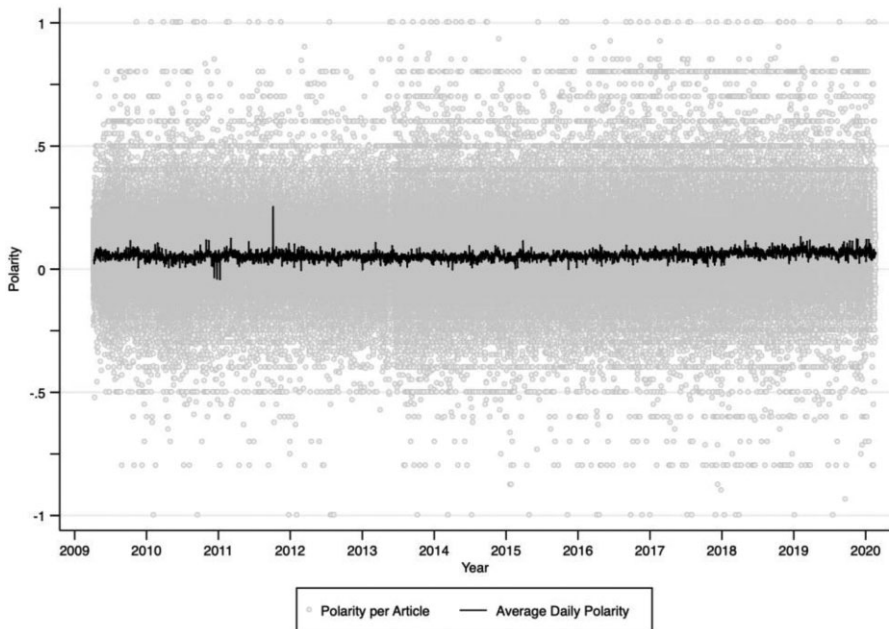
Because of space limitations, here we only present the visualizations for *Peoples' Daily* and *Global Times* in figures 1 and 2. The online appendix contains figures A1–A5 showing the other media outlet's daily average sentiment.<sup>7</sup> All the figures show there is distinct variance around the average sentiment across time and that while average daily sentiment is reasonably stable, there are some significant spikes and trends across time. In the figures, the small gray circles show sentiment per article and the black lines are the average daily sentiment for each source. In the *Peoples' Daily* sentiment visualization (figure 1), on the day of the two large negative spikes (July 22, 2012, and August 10, 2013), there were flash floods in Beijing and along the Russia–China border, respectively, whereas the three largest positive spikes (May 31, 2014, July 1, 2007, and December 7, 2014) coincide with negative political events

<sup>6</sup> Using the conversion scale, Chinese MOFA sentiment was 0.459.

<sup>7</sup> We conduct a similar exercise with the MOFA data. We produce figures for both daily and monthly average sentiment because multiple MOFA documents are rarely released on the same day. In the online appendix, figures A6–8 show MOFA's monthly average sentiment and figures A9–11 show daily average sentiment. In the average monthly sentiment figures, all three MOFAs are rarely negative. In the daily sentiment figures, MOFA sometimes uses negative sentiment, but there is usually a quick return to neutral or positive sentiment, except for Iran which we described when reporting table 1.



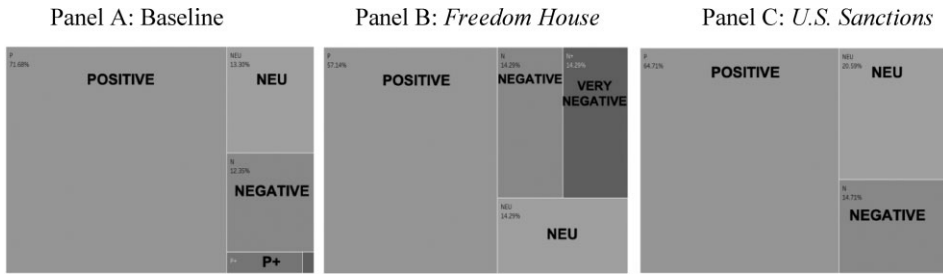
**Figure 1.** *Peoples' Daily* sentiment.



**Figure 2.** *Global Times* sentiment.

domestically and internationally—terrorism in Xinjiang, Hong Kong protests, and the release of the country's South China Sea position paper, respectively.

In the more nationalistic *Global Times*' sentiment (figure 2), the negative spikes (December 11, 2010, December 25, 2010, and January 9, 2011) and large positive spike (October 6, 2011) coincide with international events China appears to want



**Figure 3.** Russian Ministry of Foreign Affairs' sentiment, January 1, 2003–June 30, 2018.

to influence the narrative of: the 2010 Nobel Prize ceremony for Chinese dissident Liu Xiaobo on December 10 and Korean Peninsula defense talks with South Korea on December 26; on January 9, U.S. Defense Secretary Robert Gates arrived in China, and China and Russia vetoed a United Nations Security Council resolution condemning Syria's repression of protesters on October 4.

We briefly summarize some of the figures in the online appendix here. IRNA (figure A1) has some spikes in positive sentiment in mid-2013 and late-2016 with considerably more spikes in negative sentiment starting in 2017. We discuss this in more detail in the IRNA case application below. RT (figure A3) shows far more variance from 2007 to 2013 than after 2013, which happens to coincide with Putin's return to the presidency. As the standard deviations in [table 2](#) show, the sentiment in North Korean media outlets is the most variable. KCNA and Rodong sentiment (see figures A4 and A5) ebbs and flows, with some considerably larger spikes in Rodong.

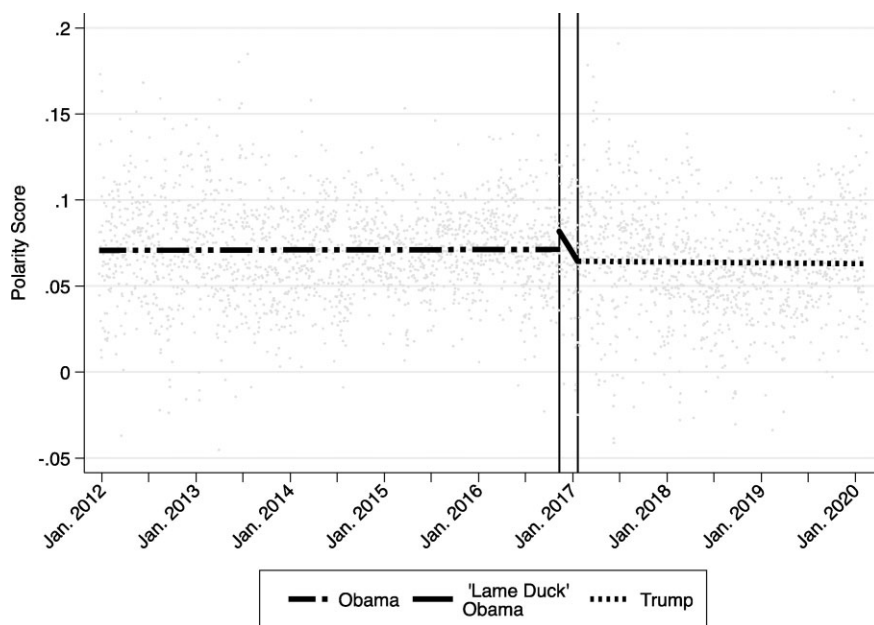
Even within countries, state-controlled (or influenced) media varies in what and how it is reporting, as seen in China, which may relate to how a government strategically applies the tool of information in terms of the intended audience, façade of a free press, and signaling to foreign governments. While we do not validate an empirical relationship, the pattern suggests that state-controlled media outlets may be used to generate positive narratives aimed at external audiences.

Next, we apply the data in two cases to showcase how it can be leveraged within DIME and how US adversaries' reactions to US foreign policy and politics can be measured through official and media sentiment.

### Case Application: Russian MOFA, Human Rights, and Sanctions

Previous research using DIME found Russia reacted more negatively to information tools—such as human rights naming and shaming ([Barry, Clay, and Flynn 2013](#); [Murdie and Peksen 2013, 2014, 2015](#); [Woo and Murdie 2017](#)) and attempted influence in Russia's domestic information environment—than diplomatic, military, or economic tools, including NATO military exercises and economic sanctions ([Fisher 2018](#)). By including additional terms, additional sources (e.g., TASS, Sputnik, and Russia Today), and comparisons with Russian activities (e.g., the timing and locations of Russian military exercises with NATO military exercises), [Fisher \(2018\)](#) revealed a multi-year effort by Russian state media to refocus criticism of Russian information controls/censorship onto criticisms of countries with which Russia was in conflict (e.g., Georgia, Ukraine).

Building from these counterintuitive findings—that Russia reacts more negatively to information tools than military tools—we leverage FOCUSdata to generate a baseline sentiment for Russian MOFA official communications from May 2004 to June 2018 (21,035 documents) in [figure 3](#), where P is positive, NEU is neutral, N is



**Figure 4.** *Islamic Republic News Agency's average Obama and Trump sentiment, January 1, 2012–February 15, 2020.*

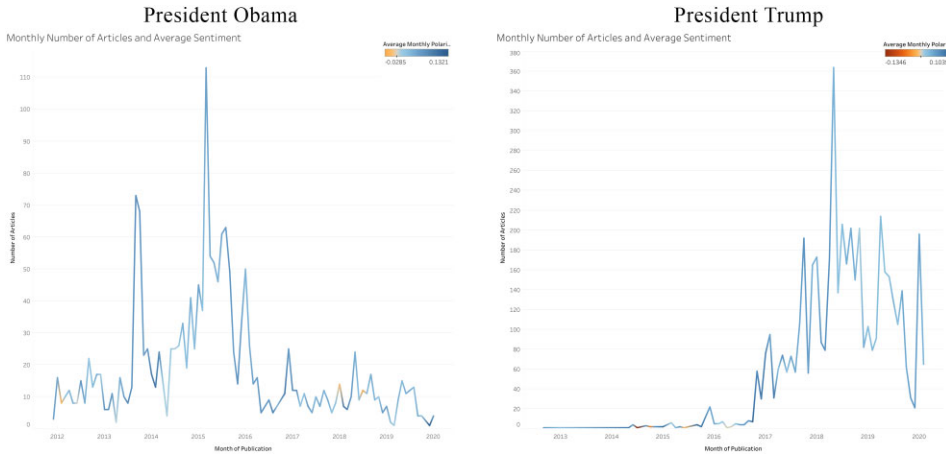
negative, P+ is very positive, and N+ is very negative. We then compare it with Russian MOFA's sentiment of documents mentioning *Freedom House* and *US Sanctions*.

Figure 3 shows that in the baseline (Panel A), MOFA is generally positive, and very positive is more common than very negative (very small, image lower-right). In Panel B, MOFA communications discussing or mentioning Freedom House are 15% less positive, contain no very positive sentiment, and show a clear increase in very negative sentiment. Contrast that with articles referring to US economic sanctions in Panel C where sentiment varied less from the baseline. The sentiment analysis suggests that Russia reacted, or at least “talked about”, more negatively to the international community's discussion of its human rights practices, primarily naming and shaming, than when reacting to sanctions imposed against it by the US government.

### Case Application: IRNA Coverage of the United States

US–Iran relations have often been an antagonistic rivalry, ranging from spates of militarized violence to cooperation culminating in the Joint Comprehensive Plan of Action (JCPOA) multinational agreement. These warmer relations ushered in by President Obama were debated in the US Senate (Kadkhodae and Tari 2019), relatively short lived, and a central focus of the 2016 US presidential elections with Republican candidates consistently criticizing the agreement. Candidate Trump was particularly vocal about JCPOA, vowing to tear up the “Iran Nuclear Deal” and re-establish aggressive US foreign policy toward Iran. President Trump then implemented a maximum pressure policy that included killing Iranian General Soleimani.

The variance in cooperation and conflict is reflected in how IRNA discussed the two US administrations. IRNA articles were, on average, more negative when discussing Trump than Obama. In figure 4, the daily sentiment scores are shown in light gray circles with the average sentiment for Obama, the “lame duck” period be-



**Figure 5.** *Islamic Republic News Agency's* average Obama and Trump sentiment by Month, January 1, 2012–February 15, 2020.

tween election and inauguration, and for Trump. During the “lame duck” period, a steep decline in sentiment is observed, which confirms expectations that Iranian state media would react to Trump’s maximum pressure policy with increased anti-US coverage.

Although IRNA maintained positive sentiment, on average, for both Presidents from January 2012 to February 15, 2020—0.071 for Obama and 0.064 for Trump—the official state media outlet was 9.9% more negative when discussing Trump compared to Obama. Submitting these average sentiment scores to an independent sample *t*-test shows that the different sentiment per administration is statistically significant ( $p \leq .001$ ).<sup>8</sup>

The data show that words matter. Our findings support the idea that changes in US (and presumably other states’) foreign policy can generate shifts in rivals’ propaganda. Our findings also support the significant attention in international relations to the effects of leadership change on state behavior, rivalries, and potential for cooperation (McGillivray and Smith 2008). Further evidence that leaders and leadership matter for international relations and security is shown in the monthly sentiment shown in figure 5. IRNA reporting regarding Obama was more consistent (ranged 0.161) and less negative than reporting regarding Trump (ranged 0.238).

Some of the most positive average monthly sentiment in IRNA regarding Obama was in early 2014 corresponding to the beginning of JCPOA implementation. The most positive the IRNA discussed Trump was at the beginning of his administration, suggesting that Iran may have given Trump a honeymoon period to adjust his policy. The sentiment is the most negative from mid-2018 through 2019 when the administration was negotiating a termination of JCPOA, designating the Islamic Revolutionary Guard Corps a foreign terrorist group, and the end of sanctions waivers for Iranian oil.

Further analysis is required to evaluate causal connections, but our findings provide clear evidence that the effects of leadership change on international and national security can be measured through analysis of state-controlled media and not only through kinetic behaviors.

<sup>8</sup> Results included in table A1 in the online appendix.

## Discussion

States use their foreign ministries and official media outlets to project an English-language narrative to outside audiences. Measuring and analyzing the articles that form this narrative—all of them, not just select articles—provide an additional tool for gaining insight into state interests. Evaluating these narratives, both the base-lines and those dealing with select terms/issues, allows researchers to better understand state signals to outside actors, how states react to crises and foreign policy tools, and state interests regarding (un)favorable developments. Especially when used alongside event databases (e.g., UN resolutions, military exercises), media data and related sentiment provide unique explanatory power.

The examples above demonstrate how the FOCUSdata aims to support researchers interested in four key states: China, Iran, North Korea, and Russia. While we recognize that public information provided by states is inherently noisy and is potentially propaganda, governments are strategically sending signals about their intentions and preferences among the noise. The case applications using Russian MOFA statements and IRNA articles provide evidence that governments use English language publications strategically. We conclude there is significant value in providing researchers with access to the wealth of raw information released by these states over the years housed in the FOCUSdata. This is especially true considering these countries recognize the strategic value of information and misinformation and make assessments using and operating within DIME. Because these governments are strategic players who purposefully engage with others, we can expect even their noisy communications to be directed toward specific goals. By examining the data, researchers can identify additional meaningful patterns or signals in the noise. We welcome efforts to harness our data, or challenge our conclusions and methodology, with the goal of developing improved tools and processes for better understanding these important international actors and the critical role of information in national and international security.

## Supplementary Information

Supplementary information is available at the *Foreign Policy Analysis* data archive.

## Acknowledgments

The authors would like to thank Aaron Ostrovsky, formerly of Rutgers University, for his work allowing us to harness Google's sentiment analysis capabilities. We would also like to thank Fabiola Soto, Rafael Mercedes, and Leslie Wilson—recent graduates of New Jersey City University—for their assistance cataloging and analyzing data and writing reports for the FOCUSdata webpage.

## Funder Information

United States Office of the Director of National Intelligence (ODNI), Project Number HHM402-19-FOA-399.

## References

- ARMSTRONG, ANDREW H. 2019. "Challenges to Coordinating the Instruments of National Power." MA Thesis, Johns Hopkins University.
- BARNUM, MIRIAM, AND JAMES LO. 2020. "Is the NPT Unraveling? Evidence from Text Analysis of Review Conference Statements." *Journal of Peace Research* 57 (6): 740–51.
- BARRY, COLIN M., K. CHAD CLAY, AND MICHAEL E. FLYNN. 2013. "Avoiding the Spotlight: Human Rights Shaming and Foreign Direct Investment." *International Studies Quarterly* 57 (3): 532–44.

- BODINE-BARON, ELIZABETH, TODD C. HELMUS, MADELINE MAGNUSON, AND ZEV WINKELMAN. 2016. *Examining ISIS Support and Opposition Networks on Twitter*. Santa Monica, CA: RAND Corporation.
- CARR, EDWARD H. 1946. *The Twenty Year' Crisis, 1919–1939: An Introduction to the Study of International Relations*. London: MacMillan & Co. Ltd.
- CLAPPER, JAMES, SUE GORDON, ROLF MOWATT-LARSEN, MIKE ROGERS, NORM ROULE, DAN HOFFMAN, AND BERN HUDSON et al. 2020. "An Intelligence Agenda for a New Administration." Policy Brief, Harvard Kennedy School, Belfer Center for Science and International Affairs, Boston, MA.
- FISHER, SCOTT. 2018. "Testing the Importance of Information Control to Pyongyang: How North Korea Reacts to the Diplomatic, Information, Military and Economic Tools of Statecraft." *International Journal of Korean Unification Studies* 27 (2): 66–110.
- FISHER, SCOTT. 2019. "Testing the Importance of Information Control: How Does Russia React When Pressured in the Information Environment?" *Journal of Information Warfare* 18 (1): 23–38.
- FISHER, SCOTT, AND GRAIG R. KLEIN. 2020. "FOCUSdata Project," 1st ed. Accessed December 3, 2021. <https://focusdataproject.com>.
- GONZALEZ, JOSE C. 2021. "Accuracy Measures in Sentiment Analysis: The Precision of MeaningCloud's Technology." Blog Posting, MeaningCloud.
- GRAY, JULIA, AND ALEX BATURO. 2021. "Delegating Diplomacy: Rhetoric across Agents in the United National General Assembly." *International Review of Administrative Sciences* 87 (4): 718–36.
- GUO, LEI, AND CHRIS J. VARGO. 2017. "Global Intermedia Agenda Setting: A Big Data Analysis of International News Flow." *Journal of Communication* 67 (4): 499–520.
- HELLER, REGINA. 2018. "More Rigor to Emotions! A Comparative, Qualitative Content Analysis of Anger in Russian Foreign Policy." In *Researching Emotions in International Relations*, edited by Maéva Clément and Eric Sangar, 75–99. Cham: Palgrave Macmillan.
- HUANG, REYKO. 2016. "Rebel Diplomacy in Civil War." *International Security* 40 (4): 89–126.
- JAMAL, AMANEY A., ROBERT O. KEOHANE, DAVID ROMNEY, AND DUSTIN TINGLEY. 2015. "Anti-Americanism and Anti-Interventionism in Arabic Twitter Discourses." *American Political Science Association* 13 (1): 55–73.
- JOINT CHIEFS OF STAFF. 2018. *Joint Doctrine Note 1-18*.
- JONES, BENJAMIN T., AND ELEONORA MATTIACCI. 2019. "A Manifesto, In 140 Characters or Fewer: Social Media as a Tool of Rebel Diplomacy." *British Journal of Political Science* 49 (2): 739–61.
- JONES, FRANK L. 2006. "Information: The Psychological Instrument." In *US Army War College Guide to National Security Policy and Strategy*, edited by J. Boone Bartholomees, Jr., Carlisle Barracks, PA: Army War College Strategic Studies Institute.
- KADKHODAEI, ELHAM, AND ZEINAB GHASEMI TARI. 2019. "Otherising Iran in American Political Discourse: Case Study of a Post-JCPOA Senate Hearing on Iran Sanctions." *Third World Quarterly* 40 (1): 109–28.
- KATZENSTEIN, PETER J., AND ROBERT O. KEOHANE, eds. 2007. *Anti-Americanism in World Politics*. Ithaca, NY: Cornell University Press.
- KOZLOSKI, ROBERT. 2009. "The Information Domain as an Element of National Power." *Strategic Insights* 8 (1): 1–9.
- LIU, BING. 2020. *Sentiment Analysis: Mining Sentiments, Opinions, and Emotions*, 2nd ed. New York: Cambridge University Press.
- MAJANI, AGNES B. 2019. "Applying Diplomatic, Information, Military and Economic Instruments of National Power to Finding a Solution to the Tanzania–Malawi Border Dispute." MMAS Thesis, US Army Command and General Staff College.
- MATTIACCI, ELEONORA, AND BENJAMIN T. JONES. 2020. "Restoring Legitimacy: Public Diplomacy Campaigns during Civil Wars." *International Studies Quarterly* 64 (4): 867–78.
- MCGILLIVRAY, FIONA, AND ALASTAIR SMITH. 2008. *Punishing the Prince: A Theory of Interstate Relations, Political Institutions, and Leader Change*. Princeton, NJ: Princeton University Press.
- MONROE, BURT L., AND PHILIP A. SCHRODT. 2017. "Introduction to the Special Issue: The Statistical Analysis of Political Text." *Political Analysis* 16 (4): 351–55.
- MORGAN, BRANDON. "Dropping DIMES: Leveraging All Elements of National Power on the Multi-Domain Battlefield." *Modern War Institute at West Point*. Accessed September 18, 2019. <https://mwi.usma.edu/dropping-dimes-leveraging-elements-national-power-multi-domain-battlefield/>
- MUNZERT, SIMON, CHRISTIAN RUBBA, PETER MEISSNER, AND DOMINIC NYHUIS. 2015. *Automated Data Collection with R: A Practical Guide to Web Scraping and Text Mining*. West Sussex: John Wiley & Sons Ltd.
- MURDIE, AMANDA, AND DURSUN PEKSEN. 2013. "The Impact of Human Rights INGO Shaming on Sanctions." *The Review of International Organizations* 8 (1): 33–53.
- . 2014. "The Impact of Human Rights INGO Shaming on Humanitarian Interventions." *Journal of Politics* 76 (1): 215–28.

- . 2015. “Women’s Rights INGO Shaming and the Government Respect for Women’s Rights.” *Review of International Organizations* 10 (1): 1–22.
- MURPHY, DENNIS M. 2012. “Strategic Communication: Wielding the Information Element of Power.” In *The US Army War College Guide to National Security Issues. Volume 1: Theory of War and Strategy*. edited by J. Boone Bartholomees, Jr. Carlisle Barracks, PA: Army War College Strategic Studies Institute.
- NATIONAL RESEARCH COUNCIL. 2011. *Threatening Communications and Behavior: Perspectives on the Pursuit of Public Figures*. Washington, DC: The National Academies Press.
- NISBET, ERIK C., MATTHEW C. NISBET, DIETRAM A. SCHEUFELE, AND JAMES E. SHANAHAN. 2004. “Public Diplomacy, Television News, and Muslim Opinion.” *The International Journal of Press/Politics* 9 (2): 11–37.
- NISBET, ERIK C., AND TERESA A. MYERS. 2011. “Anti-American Sentiment as a Media Effect? Arab Media, Political Identity, and Public Opinion in the Middle East.” *Communication Research* 38 (5): 684–709.
- NYE, JR., JOSEPH. S. 1990. “Soft Power.” *Foreign Policy* 80 (Fall): 155–71.
- . 2008. “Public Diplomacy and Soft Power.” *The Annals of the American Academy of Political Science and Social Science* 616 (1): 94–109.
- . 2009. “Smart Power.” *NPQ: New Perspectives Quarterly* 26 (2): 7–9.
- OSCARSON, KATERINA. 2017. “The Effectiveness of DIMEFIL Instruments of Power in the Gray Zone.” *NATO Open Publication* 1 (2): 1–16.
- PANG, BO, AND LILLIAN LEE. 2008. “Opinion Mining and Sentiment Analysis.” *Foundations and Trends in Information Retrieval* 2 (1-2): 1–135.
- RAVI, KUMAR, AND VADLAMANI RAVI. 2015. “A Survey on Opinion Mining and Sentiment Analysis: Tasks, Approaches and Applications.” *Knowledge-Based Systems* 89: 14–46.
- SEVIN, EFE. 2017. *Public Diplomacy and the Implementation of Foreign Policy in the US, Sweden and Turkey*. Cham: Palgrave Macmillan.
- SEVIN, EFE, AND DIANA INGENHOFF. 2018. “Public Diplomacy on Social Media: Analyzing Networks and Content.” *International Journal of Communication* 12: 3663–85.
- SEVIN, EFE, EMILY T. METZGAR, AND CRAIG HAYDEN. 2019. “The Scholarship of Public Diplomacy: Analysis of a Growing Field.” *International Journal of Communication* 13: 4814–37.
- SHELLMAN, STEPHEN M., BRIAN LEVEY, AND HANS H. LEONARD. 2011. “Countering the Adversary: Effective Policies of a DIME a Dozen.” In HSCB FOCUS 2011 Conference, The College of William and Mary, Williamsburg, VA.
- SNOW, NANCY, AND NICHOLAS J. CULL, eds. 2020. *Routledge Handbook of Public Diplomacy*. New York: Routledge.
- TABOADA, MAITE. 2016. “Sentiment Analysis: An Overview from Linguistics.” *Annual Review of Linguistics* 2 (1): 325–47.
- WILKERSON, JOHN, AND ANDREU CASAS. 2017. “Large-Scale Computerized Text Analysis in Political Science: Opportunities and Challenges.” *Annual Review of Political Science* 20 (1): 529–44.
- WOO, BYUNGWON, AND AMANDA MURDIE. 2017. “International Organizations and Naming and Shaming: Does the International Monetary Fund Care about the Human Rights Reputation of its Client?” *Political Studies* 65 (4): 767–85.
- WORLEY, D. ROBERT. 2012. *Orchestrating the Instruments of Power*. Raleigh, NC: Lulu Press Inc.