8-2024

Fantastic Trolls and Where to Find Them: Problem Framing in the Defense Against the Informational Dark Arts

Jayson M. Warren
Clemson University, jaysonw@g.clemson.edu

FANTASTIC TROLLS AND WHERE TO FIND THEM: PROBLEM FRAMING IN THE DEFENSE AGAINST THE INFORMATIONAL DARK ARTS

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Policy Studies

---

by
Jayson Michael Warren
August 2024

Accepted by:
Darren Linvill, Committee Co-Chair
Patrick Warren, Committee Co-Chair
Brandon Turner
Condoleezza Rice

ABSTRACT

In the wake of Russian interference in the 2016 U.S. presidential election, the notion of state-sponsored trolls engaging in information operations on social media has captured the attention of society. Yet eight years after this policymaking *focusing event*, sustainable solutions prove elusive and there is still much that is not known about the underlying phenomenon. This dissertation is part explanatory and part exploratory in its attempts to answer the overarching research question: *What is the Relationship Between How the Problem of State-Sponsored Trolls on Social Media is Framed & the Way U.S. Federal Policy Attempts to Solve It?* Using a convergent parallel design, this study first qualitatively (i.e., term frequency, content analysis) examines over a decade of news reporting about state-sponsored trolls to produce evidence that the dominant framings are monolithic and center around elections, Russia, and promoting an idea via generating user engagement. Second, this study quantitatively (i.e., OLS regressions) examines the entirety of the Twitter Information Operations Archive – only the second study to ever do so – along with portions of the Empirical Studies of Conflict's (ESOC) Trends in Online Influence Efforts to empirically demonstrate that state-sponsored trolls on social media are far more complex and heterogenous than commonly framed. Lastly, the study concludes by converging the qualitative and quantitative results in an examination of U.S. federal policies aimed at state-sponsored trolls and ultimately concludes that the oversimplified problem-framings do seem to have an impact on the development of policy. Consequently, these findings have implications for the intelligence community as professional problem framers and policy evaluators seeking more effective solutions.

DEDICATION

To my Mother – the woman who taught me to write and without whom this volume would have never been possible.

ACKNOWLEDGMENTS

To my wife, Grace, and daughters, Clara and Gwendelyn – thank you for enduring the ups and downs of this Ph.D. journey and encouraging me along the way. I could not have done this without you and I look forward to seeing what is next for us!

To my father, Jay – thank you for your continued support and for always being there. I just wish that Mom had lived long enough to come on this adventure with you.

To Colonel Jason King (USAF) and Steven Sheffield – thank you for setting things into motion that would ultimately trigger my ability to be at Clemson! Go Tigers!

To Darren Linvill, Patrick Warren, and the rest of the team at the Media Forensics Hub – thank you for taking a chance on me at a time where resources were scarce, for sponsoring my Ph.D. studies, and for enabling a "done dissertation" in record time. I am forever in your debt (until your next FOIA, that is, then I'll have to disavow).

To Brandon Turner – thank you for your mentorship throughout this endeavor and for the countless philosophical discussions. You are a gentleman and a scholar!

To Condoleezza Rice – thank you for being so generous with your time and for your willingness to take me on as a doctoral candidate. This was truly a life-changing experience for me and your contributions were immeasurable, especially in helping me organize my thoughts in a coherent research design. I look forward to the football game!

To H.R. McMaster, Nadia Schadlow, Jacquelyn Johnstone, Shana Farley, Yumi Higa, Chase Koontz, Denise Elson, and the rest of the Hoover Institution team – thank you for being some of my absolute favorite people and for allowing me the privilege of serving alongside you all as a colleague and friend. Hoover has become like home to me!

TABLE OF CONTENTS

Table of Contents (Continued) Page

LIST OF TABLES

List of Tables (Continued)

Table Page

LIST OF FIGURES

List of Figures (Continued)

List of Figures (Continued)

Figure Page

CHAPTER ONE

BACKGROUND & RESEARCH DESIGN

"Imperfect understanding is often more dangerous than ignorance" (Rowling 2001, xv). This sentiment, expressed by the magical beast taxonomist Newt Scamander from the Harry Potter literary universe, underpins the core premise of this dissertation and inspires the title *Fantastic Trolls and Where to Find Them*. We as the United States have a collective understanding that is imperfect – which in turn poses inherent dangers to the policymaking process and any attempt to address the problem with governmental interventions. As such, the following research seeks to remedy elements of said imperfect understanding.

To begin, it is important to bound this volume by establishing: *What is a troll?* Working from big to small, it is not a reference to mythological monsters with bridges for domiciles (e.g., the 19th Century Norwegian classic *Billy Goats Gruff*) but rather a shortform reference to internet troll. Yet even adding the "internet" contextual modifier still proves too imprecise because of the societal and etymological evolution of the term. According to the Oxford English Dictionary (2024), the first documented appearance of *troll*[1] in an online context was in 1992 on a Usenet post: "Maybe after I post it, we could go trolling some more and see what happens" (i.e., stir up trouble in internet communities). Through the 1990s and into the early 2000s, the term morphed into an

[1] "Intransitive. Computing slang. To post a deliberately erroneous or antagonistic message on a newsgroup or similar forum with the intention of eliciting a hostile or corrective response. Also transitive: to elicit such a response from (a person); to post messages of this type to (a newsgroup, etc.)."

unwieldy blanket term as the act of trolling "evolve[d] from causing annoyance as a result of your beliefs to simply believing in causing annoyance" (Feinberg 2014). Fast forward to today and the term is used so euphemistically that some people cannot differentiate between *trolling* and *cancel culture* (Cook, Tang, and Lin 2023) – setting conditions for the term to be a 21st Century rebranding of *Reductio ad Hitlerum* (i.e., everyone I disagree with is now a troll rather than Hitler).

The study contained herein adopts a more restrictive operationalization of *troll*, referring not to someone who is merely being inflammatory or belligerent on the internet but rather to a disingenuous account on social media that is concealing its true identity and taking its orders from a governmental entity (with particular emphasis on this latter criterion). Coordinated deception online is as old as the internet itself and attempts to detect it go back almost just as far (e.g., attempts in the 1990s to identify fake profiles on Usenet [Donath 1999]). However, the rise of social media platforms has lowered the cost of entry for nations to engage in information warfare to achieve their national objectives domestically and internationally – thereby creating a wicked problem for 21st Century policymakers to address in attempts to sustain free societies.

In keeping with the spirit of this dissertation's namesake, this research also seeks to set the stage for future scholars and practitioners to start formulating a taxonomy for understanding the problem of state-sponsored trolls on social media. There is a tendency to treat the problem of state-sponsored trolls and the information operations they wage as monolithic (e.g., all trolls seek to have believable personas, all information operations seek to influence/induce a behavioral response) – and thus the closest thing to taxonomy

is a rudimentary bifurcation of state-sponsored trolls as either *effective* or *ineffective* based on superficial observations regarding a profile's sophistication (Warren, Linvill, & Warren 2023). It is my contention that the efficacy of trolls should not be evaluated based on their appearance but rather on whether or not they accomplish their objectives, which may or may not be to assert active, direct influence upon unsuspecting users.[2]

But before I embark upon an exploration of the informational dark arts, it is important to contextualize where this interdisciplinary endeavor fits in the greater body of academic literature. The following will serve as more than the obligatory literature review by also incorporating some of my own personal motivation for how and why I chose to engage in this research in the manner I have designed. Towards that end, the subsequent chapters will attempt to weave theoretical, empirical, and personal strands into the mosaic tapestry that is the study of both policy and state-sponsored trolls.

Theoretical Thread – The Policy Literature

This dissertation builds upon Kingdon's (2011) *Multiple Streams Framework* which theorizes the existence of three independent "streams" within the policy process that must align for an issue to make its way onto the policymaking agenda: The *problem* stream, the *politics* stream, and the *policy* stream. While it is not always clear what causes the streams to come into sufficient alignment for policymaking to occur, Kingdon does conclude that an exogenous shock to the system (i.e., a "focusing event") creates a

[2] As an aside, this is also why I will utilize the term *information operation* rather than the term *influence operations*. In military contexts, an operation's nomenclature is typically derived from the means by which it is conducted (e.g., kinetic operations utilize physical munitions whereas non-kinetic operations are conducted through immaterial means such as the electromagnetic spectrum). By comparison all trolls do engage in operations using *information*, however it does not necessarily follow that all trolls wield direct *influence* in the Aristotelian or rhetorical sense. This nuance will become clearer in Chapter 3.

"window of opportunity"[3] for policymaking by elevating the prominence of an issue and establishing a sense of urgency. Kingdon further postulates that as the environment lurches towards the *problem*, solutions emerge from the "*policy* primeval soup" and these solutions are adjudicated within the *politics* stream.

Birkland builds upon Kingdon's theory in connecting his work to the literature on *policy learning*.[4] In *Lessons from Disaster*, Birkland (2006, 159) operationalizes a *focusing event* based "on a dramatic increase in mass and elite attention" and finds that said attention may be directed at "ideas that were generally unformulated before the event." These conclusions proved complimentary to his earlier findings in *After Disaster* (1997), namely that there is an interaction between the nature of a focusing event and the composition of policy actors who assemble to address the problem(s).

Which actors assemble further impacts how problems are framed and what ideas are combined with governmental action to create public policy (Akers 2014). Despite Frank Goodnow (1990) and Woodrow Wilson's (1887) aspirations for a professionalized public service that makes value-neutral decisions, bureaucracy researchers consistently find public administrators are at times more responsive to their own ideas than to public demand or even orders from the chief executive (e.g., James Q. Wilson 1989; Wood & Waterman 1994; Potter 2019). Foreign policy and national security communities are not immune from incorrect ideas or problem framings derailing the policymaking apparatus.

[3] Creating a window of opportunity by achieving agenda status should not be misconstrued as guaranteeing that new policies will be enacted.

[4] Birkland adopts Sabatier's (1987) argument that organizational/institutional learning is "metaphorical" on the basis it is actually individuals that possess the ability to learn. This also aligns with Nobel Laureate Elinor Ostrom's (1990) conceptualization of *institutional-choice* as aggregated human decision-making.

Military and foreign policy in Afghanistan that attempted to deal with the country as a singular nation-state struggled until James Gant's (2014) paper *One Tribe at a Time* was operationalized by General David Petraeus. Similarly, ongoing problem framings of great power competition as a deterrence struggle a la the Cold War are currently driving acquisition portfolios that at times make the U.S. less competitive against its adversaries because of its lack of emphasis on activities below the threshold of armed conflict (Haugh, Hall, & Fan 2020; Warren 2020b; Warren 2024).

Based on the literature and the nature of problem-solving, it seems reasonable that in order to understand policies one must first understand the ideas that fueled their origin and the ideas by which broader actions are deemed effective/ineffective – which means this dissertation also has potential implications for the policy implementation and policy evaluation bodies of literature. Pressman & Wildavsky (1979) noted policy implementation is a *de facto* hypothesis test in that a policy intrinsically implies theory (e.g., ideas about the nature of the problem, ideas about viable solutions) on the basis that it points to a causation chain.[5] As such, they argue that it becomes increasingly difficult to decouple the study of policy *implementation* versus policy *evaluation* because in the event of a policy failure researchers must first identify if it was a "control problem" (implementation) or a "policy problem" (evaluation) or both. J.Q. Wilson (1973) raised similar concerns in his two immutable laws of policy evaluation, paraphrased: All policy interventions succeed when evaluated by supporters and no policy interventions succeed

[5] Pressman & Wildavsky stressed this in the context of an exhortation against the "Great Leap Forward" psychology that believes anything written on paper can be substantively implemented in the real world without issue.

when evaluated by opponents. These immutable laws along with the broader implications of implementation and evaluation will be recurring themes throughout this dissertation – both as they pertain to understanding the development of policy solutions but also our "imperfect understanding" of the state-sponsored troll problem.

Empirical Thread – The State-Sponsored Troll Literature

The state-sponsored troll literature has little-to-no synthesis research examining the phenomenon holistically (e.g., I co-authored the first scoping study on this topic while developing this dissertation [Sarno et. al. 2024]). To the contrary, this body of literature is dominated by case studies of individual information operations pushing specific narratives in specific moments in space and time – which in turn subliminally reinforces the notion that information operations are a monolithic phenomenon, as alluded to previously. It was only recently that the Clemson University Media Forensics Hub (Linvill & Warren 2025 [pre-publication]) broadened the empirical aperture by creating an information operations tactics matrix and coining the term "objects of influence."

The tactics matrix emerged after realizing that there are essentially two pairs of mutually exclusive options with regards to what information operations are attempting to achieve and how trolls go about actualizing said goal. Since it is self-referentially incoherent to think that an information operation's goal is to maintain the status quo, it stands to reason an information operation is either seeking to promote a focal idea or to demote a focal idea. Likewise, since doing nothing will not move the information environment in the desired direction, trolls can promote/demote by either serving as a

direct mechanism or an indirect mechanism. These dual pairs of options converge to produce the tactics matrix in Table 1.1.

	Direct Mechanism	**Indirect Mechanism**
Goal: Promote Focal Idea	Promote Focal Idea by Strengthening Focal Idea	Promote Focal Idea by Weakening Alternative Idea(s)
Goal: Demote Focal Idea	Demote Focal Idea by Weakening Focal Idea	Demote Focal Idea by Strengthening Alternative Idea(s)

Table 1.1: Information Operations Tactics Matrix
(Linvill & Warren 2025)

In conjunction with producing the tactics matrix, the Media Forensics Hub also began drawing attention to the plurality of manipulatable items on social media by means of the term *object of influence*. In its simplest form, an object of influence is the aspect of the information environment towards which trolls are aligning their efforts. Trolls are by no means limited to only one object of influence at a time, however combining this new term with the tactics matrix begins revealing ways that certain objects of influence may be better suited for achieving certain promotion/demotion campaign goals (e.g., flooding a hashtag as the object of influence may be an effective approach to turning a conversation off whereas a robust persona as the object of influence may be an effective way of fostering dialogue in a specific manner).

Based on this state of affairs within the body of literature on state-sponsored trolls on social media, the following literature review will be constructed under the auspices of two framing questions: 1) Who is targeted by troll-waged information operations on social media? and 2) What are the objects of influence used by trolls in information operations on social media? The literature is systematically reviewed in this manner to

begin connecting what are otherwise disparate datapoints in a way that will effectively demonstrate how my dissertational research contributes to a more holistic understanding of the phenomenon of trolls.

Who is Targeted by Trolls?

With regards to *Who?* is targeted by troll-waged information operations, the macro answer is simple: Everyone! While this should not be misconstrued as implying that everyone is targeted all the time by every malicious actor simultaneously, it is paramount that the pervasiveness of information operations be understood in such a way that acknowledges a spectrum of targeting ranging from surgically-precise audiences to comprehensively global audiences. Naturally there are varying degrees of heterogeneity based on the directing actor – after all, it is generally well established that policy implementation often manifests the worldview idiosyncrasies of the implementing bureaucracy and/or greater society (Wilson 1989; Allison & Zelikow 1999; Warren 2017). That notwithstanding, the majority of said targeting variations arguably manifest as a result of the actor's goal(s) for the information operation. To demonstrate this point, Russia and China can be juxtaposed as two dipolar extremities along a spectrum of surgically-precise audiences versus comprehensively global audiences.

<u>Russia: The Exemplar for Surgically Precise Targeting</u>

In today's hyper-partisan social environment state-sponsored information operations are often superficially synonymized with Russia and the Internet Research Agency's (IRA) social media trolls during the 2016 elections; the so-called "Soviet ghost in the machine" (Valeriano, Jensen, & Maness 2020, 112-115) where Moscow continues

the KGB's legacy of active measures (*aktivnye meropriyatiya*) and disinformation (*dezinformatsiya*) to destabilize foreign governments from within by targeting populations (Rid 2020; Jonsson 2019; Haugh, Hall, & Fan 2020; Gerasimov 2013). Of note, although state-sponsored information operations are by no means unique to the cyber era, the internet has exponentially "amplified the dissemination" reminiscent to the invention of the Guttenberg printing press, radio, and television (Posetti & Matthews 2018, 1; Ellick, Westbrook, & Kessel 2018) – which thereby increased access/lowered costs for targeting audiences the world over.

Between the U.S.S.R.-turned-Russia's long history of political warfare (Rid 2020) and the *focusing event* (Kingdon 2011; Birkland 1997; Birkland 2006) of the 2016 U.S. election, the Russian use of social media trolls has become implicitly elevated as the *gold standard* by which all other information operations are measured (Warren, Linvill, & Warren 2023) – for instance, the *Politico* headline "China's Kremlin-style Disinformation Playbook" (Scott 2024) or the *Wired* article "Iran's New Facebook Trolls Are Using Russia's Playbook" (Lapowsky 2018). By extension, this also assumes that all state-sponsors are implicitly engaging in the same fundamental targeting process – impacting a specific target audience (e.g., Americans, British) with a specific targeted narrative/message (e.g., electoral politics, BREXIT) in such a way that triggers a targeted behavioral response (e.g., voting, changes in public sentiment).

As a result of these underlying tendencies within the literature, research often

analyzes account behavior and appearance.[6] Studies examining troll profile pictures of attractive women (Bastos, Marcea, & Goveia 2021); troll displays of humor and meme culture (DiResta et. al. 2017); troll manifestations of Black culture (Freelon et. al. 2020); or troll profile descriptions/bios (Krutka & Greenhalgh 2021) predominately rely on case studies and emphasize the targeted audience, the targeted narrative, and the targeted behavioral response. In that same vein these case studies also tend to examine how trolls function interpersonally with their targets via engagement metrics (e.g., likes, re-tweets).

To be clear, this scholarship has important implications for free societies, particularly at a time where global malign actors weaponize information to accomplish their national policy goals below the threshold of armed conflict (Haugh, Hall, & Fan 2020; Warren 2020b; Rid 2020). However, this research is also at times outright counterproductive when state-sponsored trolls are targeting global audiences by manipulating the information environment in the aggregate vis-à-vis meticulously targeted narratives curated to induce targeted behaviors in a targeted audience (i.e., understanding the phenomenon of state-sponsored trolls through the lens of a specific actor does not assist in understanding novel tactics developed by other actors, especially situations where non-western or domestic audiences are the targets)…Enter China as the dipolar opposite of the Russian approach to information operations.

China: The Exemplar for Global Audience Manipulation

If the aforementioned is conceptually understood as directly promoting a specific

[6] This will be a recurring theme throughout this dissertation, but for now the point of emphasis is this type of analysis can unintentionally reinforce circular presuppositions (e.g., *Trolls curate accounts to be believable because their goal is to be believed*).

idea, then much China's exploits can be inversely understood as directly demoting a specific idea. China actively endeavors to turn off conversations that are a threat to the Chinese Communist Party's (CCP) ontological security – such as suppressing social justice advocates from discussing the Uyghur genocide in Xinjiang province (Linvill et. al. 2021; Linvill & Warren 2021b; US Dept. of State 2022); disrupting calls for the boycott of the 2022 Beijing Olympics over the CCP's human rights abuses (Hundley et. al. 2022); digitally silencing politically inconvenient Chinese expatriates pointing to Beijing's culpability in the outbreak of COVID-19 (Fecher et. al. 2022), detracting from the U.S. Speaker of the House's visit to Taiwan in 2022 (Wells & Lin 2022), or subduing the pro-Hong Kong sentiments expressed by Houston Rockets' General Manager Daryl Morey (Cranmer et. al. 2024).

But just as there are bureaucratic, institutional, and cultural traits inherited from the Soviets within the Russian way of information warfare, the Chinese approach to information operations appears to manifest many of the same tactics and techniques that began within the domestic censorship apparatus used on Chinese citizens. Roberts (2018) finds that the so-called *Great Firewall of China* informational ecosystem is rampant with CCP-engineered diversions and distractions that ultimately control citizen behavior by controlling prevailing ideas. Offering a *Theory of Censorship*, Roberts (2018, 42-43) argues that Chinese censorship operates along three pillars: *Fear* (i.e., deterring people from engaging certain types of information), *Friction* (i.e., diverting attention and increasing the cost of accessing certain types of information), and *Flooding* (i.e., increasing the relative costs for prohibited information to compete by making

alternative information vastly cheaper). All three of these pillars are easily spotted within Chinese information operations, especially *Flooding* which "acts as a less observable form of censorship because it does not bring attention to the information the political entity is trying to hide" (Roberts 2018, 194).

Between the Dipoles: Information Operations Across the Spectrum

Ultimately, whether it is foreign efforts to sabotage faith in scientific researchers and institutions (Broad 2020); multinational manipulation of public sentiment in Latin America to boost sales of the Russian "Sputnik-V" COVID-19 vaccine (Linvill, Warren, & White 2022); Moscow's nefarious use of fake fact-checking during the ongoing invasion of Ukraine (Silverman & Kao, 2022); or autocratic regimes fortifying control over their own people (Linvill & Warren 2021a), information operations have near limitless applications in their efforts to subvert free societies around the world. Put more succinctly, information operations have the potential to impact every facet of day-to-day life. Yet it is precisely the near limitless applications of information operations that challenges the conventional wisdom of assuming every information operation will manifest in the same form/function.

What Objects of Influence do Trolls Use?

If one overemphasizes promoting ideas through engagement, then it would be easy to also overemphasize objects of influence that promote ideas with engagement. After all, one does not have to look far within the Psychology literature to encounter concepts such as Aristotle's persuasion triad of character (*ethos*), emotions (*pathos*), and argumentation/proof (*logos*) (Aristotle 350 B.C.; Rapp 2022; Geddes 2016); Chialdini's

(2006) influence principles of social validation, commitment/consistency, scarcity, authority, reciprocity, and friendship/liking; or other building blocks of in-group/out-group dynamics (Porpitakpan 2004; Halevy, Bornstein, & Sagiv 2008; Buttelmann & Bohm 2014; Duszak 2002; Huntington 1996). But in the same way that the limitless applications of information operations should raise questions about conceptualizing state-sponsored trolls as a monolithic enterprise, so too should the plurality of manipulatable objects on social media. Put more succinctly, there are simply too many avenues of attack for information operations to reasonably assume that there is only one way to be effective in a given campaign.

The primary objects of influence addressed in the state-sponsored troll literature are: Narratives; Hashtags; Personas; Website Domains; Users with Notoriety; Images; Videos; and Algorithms/Search Engines/Social Media Immune Systems. Each of these are addressed in turn, below.

Narratives

A narrative can be conceptually understood as the central idea that an information operation is pushing or the central idea which an information operation is seeking to turn off. When the narrative is the primary object of influence and it rises and falls on its own merit in the ecosystem of information, then influence can be achieved in similar ways to interpersonal communications – such as Aristotle's (350 B.C.) persuasion triad; Chialdini's (2006) influence principles; or other building blocks of in-group/out-group dynamics. A widely utilized tactic in this genre of activities is *astroturfing* which seeks to convince audiences that ideas originate from grassroots movements/have substantial

support amongst common people (e.g., the IRA coopting Instagram and Facebook's promotional algorithms to microtarget audiences along religious, racial, gendered, political preferences [Al-Rawi & Rahman 2020]; the South Korean secret service mimicking grassroot behaviors to influence the 2012 presidential election [Keller et. al. 2017]; or Turkish trolls masquerading as lay citizens to push pro-Erdogan/anti-opposition content in 2020 [Akca et. al. 2021]).

Hashtags

Since hashtags serve as a cross-referencing mechanism for social media platforms,[7] it is only logical that state-sponsored trolls would co-opt them for their own gains. Generally speaking, the only way to exact influence on a hashtag is to use it. That said, trolls can use hashtags to both promote and demote ideas (Linvill & Warren 2023). In terms of promotion, the IRA leveraged hashtag games about political (e.g., #IfIHadABodyDouble amidst questions of Hillary Clinton's health) and non-political topics (e.g., #ToDoListBeforeChristmas, #2016In4Words) in order to draw people into discussions online (Linvill & Warren 2020). In that same vein, the hashtag can contain the idea itself and making it trend is direct propagation of an idea (e.g., India's use of #BoycottMadeInChina [Grossman et. al. 2022,12]). As for demotion, the Chinese use hashtag flooding to censor speech and make a hashtag less useful to those critical of the CCP (Roberts 2018; Linvill & Warren 2021b; Linvill et. al. 2021).

[7] Although there is potential for overlap, it is the indexing properties that allow a hashtag to be differentiated from a *narrative*. While some narratives are spread via hashtags, not all hashtags are in and of themselves narratives.

Personas

Sometimes the object of influence is the account persona (i.e., convincing audiences of the credibility of the persona so anything it says/does then has credibility by the transitive property). Influence can be achieved in this regard by curating the account in such a way that resonates with targeted audiences (e.g., troll profile pictures of attractive women [Bastos, Marcea, & Goveia 2021]; troll displays of humor and meme culture [DiResta et. al. 2017]; troll manifestations of Black culture [Freelon et. al. 2020]; troll profile descriptions/bios [Grossman et. al. 2022, 5; Krutka & Greenhalgh 2021]) – an overarching tactic collectively known as *backstopping*. Using fake followers (Jamison, Broniatowski, & Quinn 2019), purchasing/hacking accounts with higher numbers of followers, or maintaining cross-platform personas can also bolster a persona's clout. Conversely, information operations can be used to undermine the credibility of real personas – such as the CCP creating hordes of fake accounts to detract from the online presence of both virologist Li Meng Yan and CCP critic John Churchill (Warren et. al. 2023, discussed further in Chapter 3).

Website Domains

Trolls can use website domains as an object of influence by distributing the hyperlinks to as many places on the internet as possible. Whereas someone may be suspicious of a troll persona, a website can offer a perceived sense of legitimacy that is also not subject to a social media platform's terms of service. It can substantially reduce the costs of producing content by hosting it on one central domain to share from and can enable trolls to co-opt real-world information for their own nefarious purposes. The IRA

created news aggregator Twitter accounts masquerading as regional U.S. news outlets that would tweet actual news with a pro-Russian positionality, replete with links to external news sources (Linvill & Warren 2020). India utilized links to global (e.g., youtube.com, facebook.com) and regional (e.g., defencenews.in, aninews.in) websites in their pro-Indian Army campaign (Grossman et. al. 2022, 4). Lastly, for a multi-layered approach to disinformation Russian and Iranian trolls have both been observed redirecting users to fake news websites run by the same actor (e.g., blackmattersus.com and donotshootus.us [Zannettou et. al. 2019, 359-60]).

Users with Notoriety

Exploiting people with notoriety as an object of influence is by no means new to the era of social media; observing such activities today pays homage to the KGB's notion of *useful idiots* and triggers memories of Dan Rather reading Russian disinformation about the AIDS virus leaking from a U.S. Army laboratory (Ellick, Westbrook, & Kessel 2018). To exact influence on said users, trolls must bring them into their activities. On social media, this can be attempted by something as simple as *tagging* (e.g., India tagged journalists and regional politicians [Grossman et. al. 2022, 5]). Similarly, it can also be done by surreptitiously entering the discourse in such a way that pulls in unwitting journalists (e.g., Russia tricking journalists into writing stories for them [Wanless & Walters 2020]) or by capitalizing on vulnerabilities inadvertently created by users with notoriety (e.g., IRA trolls entering and winning Eric Zorn's *Chicago Tribune* "Tweet of the Week" contest [Zorn 2019]).

Images

Text-based content is substantially easier to research in mass because it does not require inordinate amounts of digital storage and can be parsed expediently using automated tools. Images, on the other hand, have substantially larger file sizes and any embedded text requires processing by optical character recognition (OCR) before it can be entered into any type of analysis tool. As a result of these limitations (and data availability), the preponderance of literature on state-sponsored social media trolls has gravitated towards Twitter. Nevertheless, there are empirical examinations of images as an object of influence. Both the Russians (DiResta et. al. 2017) and the Chinese (Hundley et. al. 2022) make use of memes. Similarly, the Russian use of fake fact-checking during the ongoing invasion of Ukraine relied on doctored photographic evidence to convey the central theme of the information operation (Silverman & Kao, 2022). But this does not mean that images are restricted to promoting ideas – for example, Russian trolls on Reddit leveraged the photo hosting platform Imgur to use pictures as an opportunity to accumulate *karma* for their personas (Zannettou et. al. 2019, 359-60) and Chinese trolls use caricature-style cartoons to flood out politically inconvenient voices they deem a threat (Fecher et. al. 2022). More research into the use of images is needed, especially as text-based platforms give way to more multimedia-centric interfaces (e.g., Instagram).

Videos

All the limitations for analyzing imagery are also applicable to videos (e.g., requires massive amounts of storage space; requires voice-to-text extraction to do large

scale content analysis). While there have been documented instances of using video (e.g., China's video urging Taiwan's capitulation at approximately the same time the U.S. Speaker of the House's visited the island [Warren et. al. 2023]), more research is necessary as multimedia-based platforms rise in prevalence. More importantly, new tradecraft to research/counteract the threat of video-based information operations must transcend simply recycling old techniques (e.g., Bellingcat's tool for hashtags on TikTok [Wild 2022]), especially considering TikTok's first official covert information operations reports omit China and include Taiwan on the list of known perpetrators (Ryan 2023) – thereby highlighting the need for journalists/researchers/open-source practitioners to check the CCP's power over TikTok as one of the fastest growing platforms in the world.

Algorithms, Search Engines, & Social Media Platform Immune Systems

Finally, little-to-no research has been done on state-sponsored trolls being used to influence algorithms and search engines. This is arguably the case for two key reasons: 1) The platforms/data firms are not going to disclose their trade secrets and intellectual property to researchers; and 2) Attempting to empirically demonstrate a counterfactual with regards to algorithmic prioritization/search results would almost inevitably devolve to the logical fallacy of proving the negative. Yet some observations from the Clemson University Media Forensics Hub do appear to indicate that not only is there an effect but that the effect may actually be the ultimate goal of some information operations (Linvill & Warren 2021b; Warren et. al. 2023).

In the case of China and the #Xinjiang flooding campaign, not only did it impact the likelihood of the trending algorithm to present the social justice commentary about

the Uyghur genocide but it also made it nearly impossible to manually search for that same content. Similarly, when the flood of Li Meng Yang caricatures on Pinterest interacted with the algorithms controlling the platform's trust and safety moderation protocols, Pinterest removed all Li Meng Yang content (i.e., the platform's immune system did China's job for them). More research into algorithms, search engines, and social media platform immune systems as objects of influence is of extreme importance and will only become of increasing significance as technology advances.

Personal Thread – Intelligence Analysts as Professional Problem Framers

"What problem are you trying to solve, Jayson?" As a young Lieutenant at my first duty station, I heard this question daily from the unit's #2 – callsign: *Meat*. An old, washed-up fighter pilot, Meat had a no-nonsense personality that by default sought to pierce through the noise and chaff in order to devote time to things of actual significance. As such, whenever policy changes were being teed up for the Squadron Commander to make a decision, Meat would always ask me as the Executive Officer what problem the policy change was seeking to solve. When there was a clearly articulated causal relationship connecting the problem and proposed policy, Meat would open the administrative gate and allow the decision folder into the Boss's office. However, when the connective tissue between problem and policy was ostensibly absent, Meat would refer the action officer (along with his/her efforts) back to the proverbial drawing board.

Meat would also tell me: "You are an officer *first*, an intelligence analyst *second*!" While I found it readily apparent how an emphasis on problem-framing applied to my primary function, I did not fully appreciate at the time how Meat's mantra had far broader

applications than just officership or running the day-to-day within the largest squadron of its kind in the U.S. Air Force (and I definitely did not foresee that I would be using it as introductory material for my doctoral dissertation a decade later). As policymakers seek to solve problems in the national security domain, it is our responsibility as the Intelligence Community to frame the problem in such a way that fulfills our mandate to *Speak Truth to Power*. When the problem is framed correctly it meaningfully energizes the policymaking process and *provides for the common defence*; when the problem is framed incorrectly (or worse, actively misrepresented) it drives the policymaking process off the tracks and inevitably leads to wasted resources, missed opportunities, and the survival of the actual problem.

With these experiences forming my earliest "core memories" as an officer, to borrow a turn of phrase from the Disney-Pixar movie *Inside Out* (one of my daughters' favorites in the early years), I continued to climb the ranks. My next assignment was as aircrew on one of the U.S. Air Force's intelligence-collection aircraft and it was here I was entrusted with my first command opportunity. While in the seat, I strove to channel my inner-Meat daily and routinely had my fellow Flight Commanders stopping by my office for assistance with challenges (effectively actualizing another one of Meat's common refrains: "Your goal should be that you cannot walk from one end of the building to the other without someone stopping you to ask a question"). However the Wing as a whole had a significant number of problems, none the least of which being its chronic dysfunction and underperformance. This problem was the direct result of what James Q. Wilson (1989) refers to as *organizational culture* – and culture problems are

solved through accountability and standards. But my superiors went out of their way to frame it as a resource problem which meant that, rather than hold people accountable and/or enforce standards, they spent their time attempting to increase manpower and bonuses through whatever means necessary. More manpower and money did come, but the dysfunction continued and even worsened because it normalized underperformance from the top-down by legitimizing Pavlovian retorts of "we are undermanned" as a universal *get out of jail free* card for anyone wanting to shirk a task/perform mediocrely.

Upon completion of my airborne tour, I was humbled to receive a "by name request" assignment from a General Officer to be a part of the 16th Air Force (Air Forces Cyber) stand-up. As the first-of-its-kind organization dedicated to Information Warfare, we as the initial cadre sought to build institutional practices and norms that would enable the U.S. government to holistically (i.e., not just militarily) engage our geopolitical competitors. In so doing, I encountered yet another problem framing issue. While we as an organization were discussing engaging adversaries below the threshold of armed conflict *today*, much of the DOD was preparing for possible combat *tomorrow*. This triggered two very different proposed resourcing portfolios. We advocated for increased cyber, information, and Interagency cooperation capabilities that could compete against China and Russia now while the preponderance of the Department solicited new weapons systems in hopes they would deter adversaries from armed conflict despite the reality our adversaries were undeterred and actively making strategic gains without warfare (e.g., China seized *de facto* control of the South China Sea without firing a shot; Russia annexed portions of Ukraine in 2014 and 2016 without being considered an *invasion* or

illegal *act of aggression*). Although incremental gains were made in the former, anecdotally it always felt that the latter by and large prevailed – which in turn created more potential vulnerabilities for America and her allies.

Interestingly enough, after a long series of providentially-orchestrated events, I continued operating in the great power competition context at my next assignment when I was given the once-in-a-lifetime opportunity to study as a full-time Ph.D. student at a civilian institution for three years. In partnership with the Clemson University Media Forensics Hub, I began conducting interdisciplinary research into the phenomenon of state-sponsored trolls on social media. True to form, I once again encountered matters of problem-framing that raised more questions than answers. The more I read and the more I spoke with the community of interest, the more it felt as if trolls were overwhelmingly framed as a completely new/novel phenomenon unlike anything experienced previously; as a uniquely Russian phenomenon; and as a phenomenon operating primarily within the setting of an election. The more I found this to be seemingly commonplace, the more I questioned it. Are these really the intrinsic qualities of the problem?

Research Purpose

In weaving these three aforementioned strands (i.e., theoretical, empirical, and personal) together, I have three principal goals for this dissertation as they pertain to both my academic pursuits and my professional mandate as a warrior-scholar within the U.S. Intelligence Community. First and foremost, this endeavor seeks to produce a more holistic understanding of the problem of state-sponsored trolls – validating what we do know, challenging what we think we know, and producing insights for that which we do

not yet know. Second, this study seeks to provide intelligence analysts (and other professional problem framers, particularly within national security policymaking forums) with a tangible exemplar of what happens when policymaking is energized by correct/incorrect mental models of problems. Third, this project seeks to demonstrate the importance of decisionmakers understanding and challenging assumptions prior to acting.

Overarching Research Design

The following study is part explanatory and part exploratory. Ultimately, it seeks to offer some explanatory insights into the overarching research question: *What is the Relationship Between How the Problem of State-Sponsored Trolls on Social Media is Framed & the Way U.S. Federal Policy Attempts to Solve It?* In pursuing an answer, this study contains three research questions under the umbrella of the primary:

- RQ_1: *How is the Problem of State-Sponsored Trolls on Social Media Framed?*
- RQ_2: *How are State-Sponsored Trolls on Social Media Operationalized to Conduct Information Operations?*
- RQ_3: *What Are Policies the U.S. Federal Government Uses to Solve the Problem of State-Sponsored Trolls on Social Media?*

By engaging with these three related RQs independently, it I will use their answers as the requisite pieces to empirically pursue the overarching research question.

Chapter 2 builds upon the literature review and synthesis above by qualitatively and quantitatively exploring non-academic commentary via the research question:

- RQ_1 = *How is the Problem of State-Sponsored Trolls on Social Media Framed?*

By approaching this research question through a dataset consisting of media coverage, Chapter 2 seeks to identify problem-framing trends and idiosyncrasies within the public

discourse surrounding state-sponsored trolls on social media.

Once the framing tendencies have been established, Chapter 3 will attempt to establish a more holistic understanding of the challenges presented by *social media trolls* through the following research question:

- RQ_2: *How are State-Sponsored Trolls on Social Media Operationalized to Conduct Information Operations?*

To produce empirical evidence to answer this question, the following two sub-questions are posed in order to attain more granular datapoints:

- $RQ_{2.1}$ = *What are Ways State-Sponsored Trolls Invest in the Development of their Online Personas?*

- $RQ_{2.2}$ = *What is the Relationship Between the State-Sponsored Troll Persona Investment and the Goals of the Information Operations they Conduct?*

Because these questions deliberately seek generalizable findings, Chapter 3 approaches them using a novel persona-investment scale and two *ordinary least squares* (OLS) models to quantitatively analyze the Twitter (n.d.) Information Operations Archive in its entirety (N = 87,437 trolls). Although such an approach is limited to a single social media platform, it nevertheless provides a unique opportunity for generalizable findings regarding the phenomenon of *social media trolls* by using the entirety of the Twitter dataset as opposed to a selective subset of the trolls operating within a single case study.

Chapter 4 then attempts to explore the governmental response to state-sponsored information operations through the question:

- RQ_3 = *What are Policies the U.S. Federal Government Uses to Solve the Problem of State-Sponsored Trolls on Social Media?*

This portion of the study is inherently exploratory and does not produce exhaustive findings – which in and of itself highlights the need for policy research and policy evaluation in this particular field. Because there is no lead agency for the problem of state-sponsored information operations, the disjointed federal response lacks synchronization across all the departments involved (e.g., Defense, State, Homeland Security, Justice). Consequently, Chapter 4 seeks to compile assorted policy responses in order to extrapolate how the federal government's attempted solutions seem to conceptually frame the problem.

Lastly, Chapter 4 will also serve as the conclusion by using the answers to the three related RQs in order to provide insights into the overarching question: *What is the Relationship Between How the Problem of State-Sponsored Trolls on Social Media is Framed & the Way U.S. Policy Attempts to Solve It?* More specifically, this section compares/contrasts the way in which the problem is framed in the news (Chapter 2) and the generalizable findings of the phenomenon (Chapter 3) to see if the problem is conceptually understood the same way. Given that the surveyed literature above suggests the possibility that the dominant framing trends may be an incomplete representation of reality, Chapter 4 will examine the governmental response in order to determine if the federal understanding of the problem is more akin to the findings of Chapter 2 or Chapter 3 – which in so doing will ultimately provide insights into the relationship between problem framing and the governmental response.

CHAPTER TWO

DON'T FIGHT THE PROBLEM, DECIDE IT

"Don't fight the problem, decide it." These words, attributed to former U.S. Secretary of State George Marshall (Diaz-Plaja & Polchar 2023), highlight the potential pitfalls that can emerge from attempting to conform a problem to one's own will vis-à-vis making a deliberate decision regarding the objective essence of the problem and crafting policy mechanisms/instruments accordingly. In many ways, these words are evocative of the forcing-function question I routinely encountered from Meat in the previous chapter: "What problem are you trying to solve?"

In light of this, the following chapter begins pursuing the overarching research question (i.e., *What is the Relationship Between How the Problem of State-Sponsored Trolls on Social Media is Framed & the Way U.S. Federal Policy Attempts to Solve It?*) by qualitatively and quantitatively exploring the discourse surrounding *state-sponsored trolls* via the research question:

- RQ_1 = *How is the Problem of State-Sponsored Trolls on Social Media Framed?*

This research question is approached by applying content analysis techniques to a novel dataset created from major newspaper coverage about state-sponsored trolls engaging in information operations on social media.

Data

In a completely transparent and accessible world, this research would exploit official governmental records (e.g., threat bulletins, intelligence reports, briefings, interdepartmental memoranda) in order to produce insights regarding how the problem of

state-sponsored information operations on social media is framed to policymakers. Yet given the nature of this particular issue and its corresponding touchpoints with the national security apparatus, such data is shrouded in the secrecy of classification and/or obscured by redactions – for legitimate reasons (e.g., protecting collection sources and methods; preserving decision advantage for civilian/military leadership). In such a transparent and accessible world, one would also expect social media platforms themselves to make timely ground-truth information available regarding malicious activities – but there is no incentive to do so given the potential adverse effects that highlighting vulnerabilities can have on revenue, share prices, and consumer trust.

In lieu of such data, media reporting can provide substantive datapoints regarding problem framing and is not without justification in the policy literature. While Kingdon (2011, 57-61) finds that the media plays a generally insignificant role in getting issues on the policy agenda, once a problem is on the agenda (e.g., after a focusing event) then the media is often reporting on what is transpiring within policy communities. Similarly, Downs's (1972) issue-attention cycle and Schattschneider's (1960) emphasis on issue visibility underpinned Baumgartner and Jones's (2009) work on *Punctuated Equilibrium Theory* – which methodologically relied on the *Readers' Guide to Periodical Literature* and *New York Times Index* data to test for instances of Downsian (public support) or Schattschneider (public opposition) mobilizations/policy framings. Lastly, Birkland (2009) makes considerable use of media reporting from *The New York Times* as evidence of *policy learning* within the policymaking process.

Data Aggregation & Selection Criteria

Using the ProQuest database *US Major Dailies*, a novel dataset of news reporting about state-sponsored trolls on social media was created from five widely-circulated American newspapers: *The New York Times*, *The Washington Post*, *The Wall Street Journal*, *Los Angeles Times*, and *Chicago Tribune*. For articles to be included they needed to be published between 2010-2023; have the full-text available; be in the print or online versions of the newspapers (i.e., not a subsidiary blog, podcast, etc.); and be indexed in the database as either *news*, *commentary*, or *editorial.*

With regards to content, a Boolean logic formula was developed to bound the article aggregation based on Operational Domain, Origin/Activity, and Digital Deception. Flags were developed iteratively in collaboration with other experts on state-sponsored trolls on social media and assessed on their tendencies to trigger false positives (e.g., "online" as an Operational Domain yielded considerable false positives for articles indexed on the "online" version of newspapers; "foreign" as an Origin/Activity was too simplistic and produced false positives for "foreign policy" and reporting from the "foreign desk"). The final Boolean logic is reflected in Table 2.1 and articles must use at least one term from each of the three columns to be included in the dataset.

Operational Domain	Origin/Activity	Digital Deception
Social Media	Nation State	Troll(s)
Social Network(s)	Nation-State	Bot(s)
Social Networking	State Sponsor(ed)	Fake Account(s)
Internet	State-Sponsor(ed)	Fake User(s)
	Foreign Influence	Fake Profile(s)
	Foreign Actor	
	Interference	
	Information Operation(s)	
	Influence Operation(s)	
	Influence Campaign(s)	
	Malign Actor	
	Malign Activities	
	Malign Influence	

Table 2.1: Boolean Selection Criteria for News Dataset

Robustness Checks

After determining the database filters and selection criteria, the aggregated articles were subjected to three robustness checks to ensure the Boolean logic performed correctly prior to analysis. The first was a chronology evaluation to see if the distribution of articles by year is intuitive based on real-world circumstances. This will be discussed in greater depth in the Method & Findings section (see Figure 2.1), but for the purposes of a robustness check the 23 articles between 2010-2016, a spike in 2017, a peak in 2018, and 2019-2023 never returning to pre-2017 numbers are reconcilable with the dual reality that the problem existed before the 2016 U.S. presidential elections but the focusing event of Russian interference drew considerable coverage for an enduring problem.

The second check was an estimation of the false-positive rate. 50 articles were manually reviewed, the first 25 published and the last 25 published. In the first 25, there

were five false positives (i.e., two referred to the cyber deviant behavior of "trolling" to illicit a response; two contained reference to Hillary Clinton's emails combined with either a botnet cyberattack or calling Dinesh D'Souza a "troll;" and one mentioned the ruling of Advocate General Yves *Bot* [emphasis added] in a data security case before the Court of Justice of the European Union) – but it is important to note that these false positives were at least in the correct genres of discourse (e.g., data privacy, social media, information security, cyber). Likewise there were five on the margins of inclusion (i.e., one referred to fake profiles on Gab attempting to taint the image of the platform; two discussed Tunisian dissidents using nascent virtual private network [VPN] technologies to mask personas from the government; and two mentioned Clinton and Trump having fake followers on Twitter). The remaining 15 of the first 25 published and all of the last 25 published were correctly on topic which ultimately yielded a false positive rate of 10% (i.e., 45 out of 50).

It is important to note that this 10% is likely an overestimate given: 1) the deliberate oversampling of the period where false-positives are the most probable (i.e., 2010-2016), and 2) the last 25's perfect selection seems to indicate the Boolean performs better after the focusing event publicizes the issue. Consequently, the third and final check was a cursory review of the dataset during export. Because ProQuest caps downloads to increments of 100, each batch was skimmed within the *US Major Dailies* interface during cuing and the vast majority of the articles appeared to be on topic – thus further supporting the assumption that the false-positive rate is well below 10%.

Method

Upon finalizing the dataset, it was exploited using an exploratory approach to content analysis. More specifically, the data was first analyzed by identifying observable patterns within the distribution of publications annually and within the *US Major Dailies* database indexing tags. The articles were then compiled into a *.CSV* file for term frequency analysis in the opensource statistics software *R*, using both the patterns in the database indexing as well as the existing literature (Chapter 1) to inform the development of content markers for quantification.

Of note, embedded within my methodological design is a desire to produce more nuanced findings than could otherwise be attained through a rudimentary approach to term frequency (i.e., simple counting). Towards that end, I blend term frequency with the econometric concepts of *extensive margin* and *intensive margin*. From a supply-side economics perspective, the *extensive margin* can be understood as how many firms produce a certain commodity or good (i.e., all the firms in an economy are evaluated on a dichotomous 0/1 based on whether they produce the commodity or good in question and the resulting sum is the *extensive margin*). On the other hand, the *intensive margin* expresses how much of said commodity or good the firms collectively produce. Applying these concepts to term frequency, the metaphorical *economy* in this study is all the articles in the dataset which in turn means the *extensive margin* examines how many articles contain certain terms while the *intensive margin* reflects how many times those terms are collectively used.

Findings

Systematically approaching the data in this manner yielded four principal findings and they are presented in the order identified so as to annotate how initial findings combined with existing theory to produce additional findings. The four findings are: Annual publication rates have considerable variations; indexing tags emphasize elections, Russia; term frequency emphasizes elections, Russia; and term frequency suggests a possible emphasis on tactics employed in promoting an idea.

Finding #1: Annual Publication Rates Have Considerable Variations

The first finding from the news reporting dataset is the existence of prominent variations in article publication patterns over time. Figure 2.1 portrays the distribution of articles (N = 1,896) according to the year they were originally published.

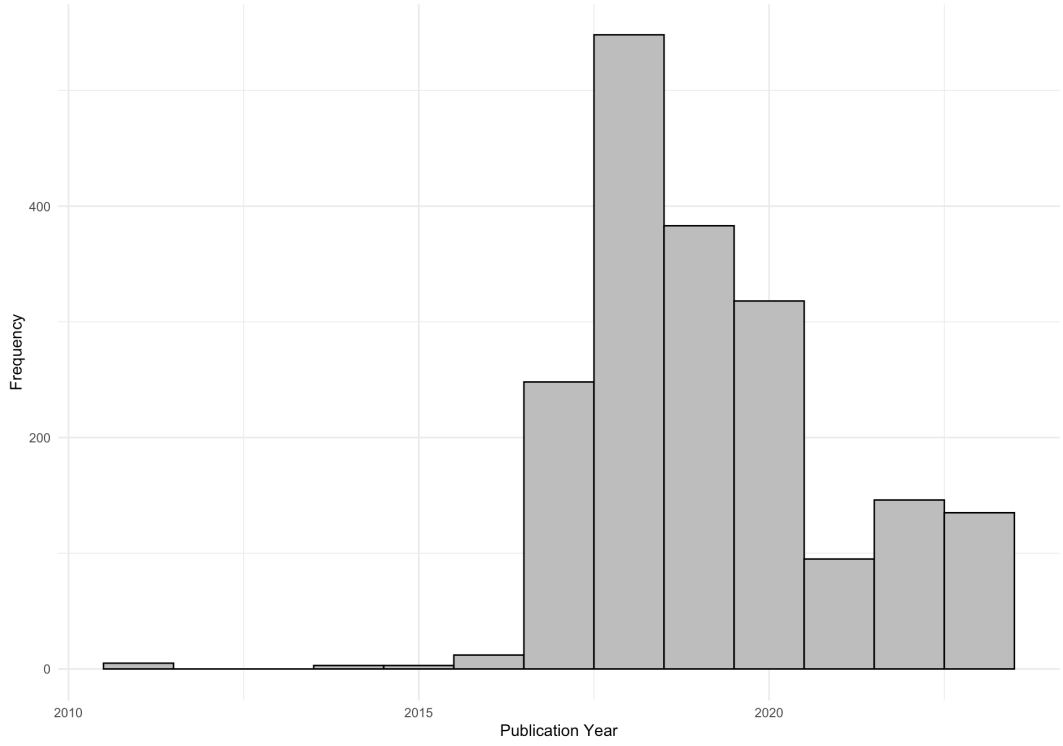

Figure 2.1: Annual Publication Distribution

Based on this distribution, the dataset arguably self-divides itself into three distinct time periods: a low rate of publication period from 2010-2016 (23 articles, 1% of dataset), a high rate of publication period from 2017-2020 (1,497 articles, 79% of dataset), and a medium rate of publication period from 2021-2023 (376 articles, 20% of dataset).

Finding #2: Indexing Tags Emphasize Elections, Russia

Initial indications of how the problem of state-sponsored trolls on social media might be framed emerged in two different ways by means of the *US Major Dailies* database indexing (i.e., the tags attached to articles that facilitate in cross-referencing and narrowing search engine results by topic). First, the indexing tags pertaining to elections are very prominent within the "Subject" field. Out of the top 25 "Subject" keywords in the dataset, 11 of them are directly connected to electoral dynamics (i.e., *presidential elections*, *political campaigns*, *politics*, *political advertising*, *elections*, *presidents*, *political parties*, *democracy*, *election results*, *political activism*, *candidates*). By extension, almost all of the top 25 "Person" keywords are individuals involved in American domestic politics (see Table 2.2). Interestingly, having commensurate indexing as Joe Biden, Barack Obama, and Bernie Sanders combined is Vladimir Putin – which, along with Yevgeny Prigozhin at #24, offers evidence of another potential framing trend.

The *US Major Dailies* indexing tags seem to also imply that the problem of state-sponsored trolls on social media primarily originates from Russia as the principal perpetrator. In addition to the prominent place Vladimir Putin holds in the "Person" tags, Russia dominates the "Location" keyword indexing with 1,184 articles tagged. In fact,

Russia has over three times more articles tagged than China (China [291] + Beijing China [86] = 377); yet China tags are not substantially more than Ukraine tags (271), which are likely in and of themselves an extension of reporting about Russia. Another indicator of Russia's prominence within the dataset, coming in at #4 of the "Company/Organization" tags is the late Yevgeny Prigozhin's *Internet Research Agency* as the only non-U.S. based entity in the top 10 (and one of only 5 in top 25).

However, it is paramount to realize that these aggregated database indexing tags are for the dataset as a whole – 79% of which is clustered in the 2017-2020 time period at the highpoint of news coverage of the 2016 election interference. As a result, it is possible that news reporting about the focusing event is disproportionately skewing the distribution of election and Russia tags, thereby limiting their potential efficacy for explanatory power regarding the framing of state-sponsored trolls wholistically.

Subject	Tags	Person	Tags
Social Networks	1,440	Trump, Donald J	716
Presidential Elections	511	Mueller, Robert S III	245
Political Campaigns	489	Clinton, Hillary Rodham	243
False Information	412	Putin, Vladimir	231
Politics	332	Zuckerberg, Mark	164
Political Advertising	266	Biden, Joseph R Jr	93
Internet	261	Obama, Barack	74
Elections	249	Sandberg, Sheryl	67
Propaganda	236	Sanders, Bernard	64
National Security	216	Manafort, Paul	54
Congressional Committees	209	Barr, William P	46
Presidents	207	Dorsey, Jack	41
Intelligence Gathering	184	Stamos, Alex	41
Criminal Investigations	178	Coats, Dann	38
Researchers	170	Comey, James B	37
Political Parties	162	Timberg, Craig	37
Intelligence Services	160	Musk, Elon	35
Democracy	153	Sessions, Jeff	34
Conspiracy	149	Burr, Richard	31
Election Results	129	Jones, Alex	31
Political Activism	125	Cohen, Michael D	29
Computer Security	114	Schiff, Adam B	29
Candidates	109	Podesta, John	28
Information Warfare	109	Prigozhin, Yevgeny	27
Cybercrime	107	Warner, Mark R	25
Location	**Tags**	**Company/Organization**	**Tags**
United States	1,477	Facebook Inc	688
Russia	1,184	Twitter Inc	524
New York	309	Congress	386
China	291	Internet Research Agency	325
Ukraine	271	FBI	243
Iran	181	Google Inc	239
Europe	140	Senate	215
United Kingdom	117	Department of Justice	128
Silicon Valley-California	113	New York Times Co	127
California	99	YouTube Inc	126
Beijing China	86	Committee on Intelligence, Select Senate	124
Germany	84	Democratic Party	109
France	76	Republican Party	96
Texas	64	Department of Homeland Security	95
Florida	60	Democratic National Committee	93
Middle East	60	European Union	82
Saudi Arabia	54	Wikileaks.org	78
Virginia	53	National Security Agency	72
Israel	50	North Atlantic Treaty Organization	63
Africa	49	Cyber Command-US	58
Hong Kong	43	Cambridge Analytica	57
Georgia	42	Wall Street Journal	56
Australia	41	Central Intelligence Agency	51
India	41	CNN	49
North Korea	36	Microsoft Corp	43

Table 2.2: US Major Dailies Indexing Keywords

Finding #3: Term Frequency Emphasizes Elections, Russia

Prompted to dig deeper into the dataset based on the prevalence of certain database indexing terms, the full text of all 1,896 articles was analyzed in the opensource software *R*. Using the *wordcloud* and *tm* (i.e. Text Mining) packages, Figure 2.2 depicts the top 80 words in the corpus (after removing stop words).

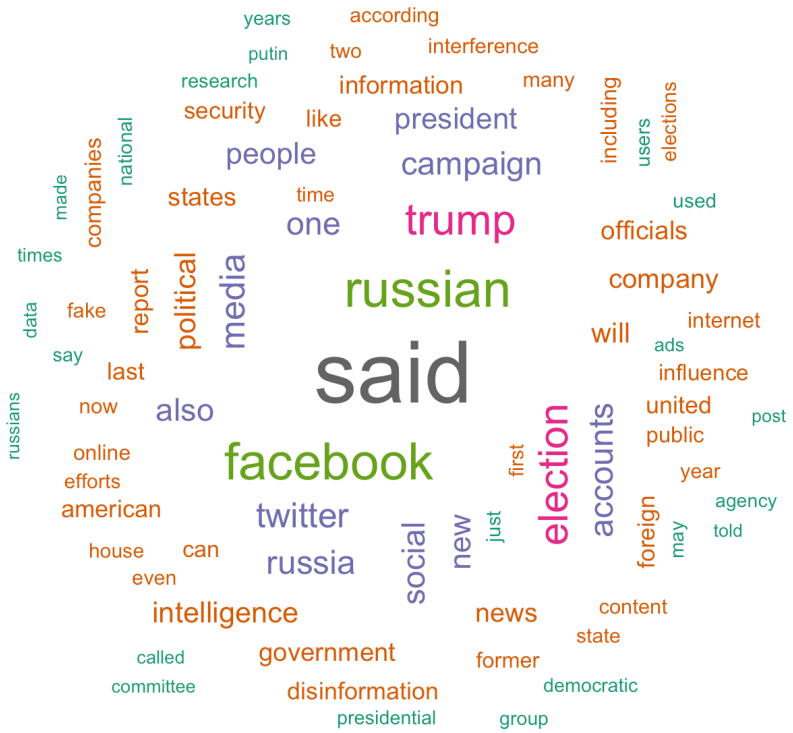

Figure 2.2: News Article Text Corpus Word Cloud

Figure 2.2 yields results consistent with the data from the database indexing tags, namely the prevalence of election-related terms (e.g., "election," "elections," "political," "president," "presidential," "democratic") and references to Moscow (e.g., "Russia," "Russian," "Russians," "Putin"). Nevertheless, the presence of "Trump" also highlights that simple term frequency could be skewed by coverage of the 2016 elections — hence

the use of the economics concepts extensive margin and intensive margin to produce more tractable and substantive insights.

Election Extensive & Intensive Margins

To begin the econometric approach, I first probe mentions of *elections* based on their prominence in the database indexing tags. To cross the threshold for being included as a *firm* dealing in *elections*, an article must contain at least one reference to: "election(s)," "voter(s)," or "vote(s)." The resulting extensive margin is 1,706 articles referencing elections. Given that this margin is 90% of the dataset, plotting this distribution over time would be essentially a recreation of the annual distribution of articles in Figure 2.1. Instead, Figure 2.3 offers a distribution of the 190 articles failing to trigger the election marker.

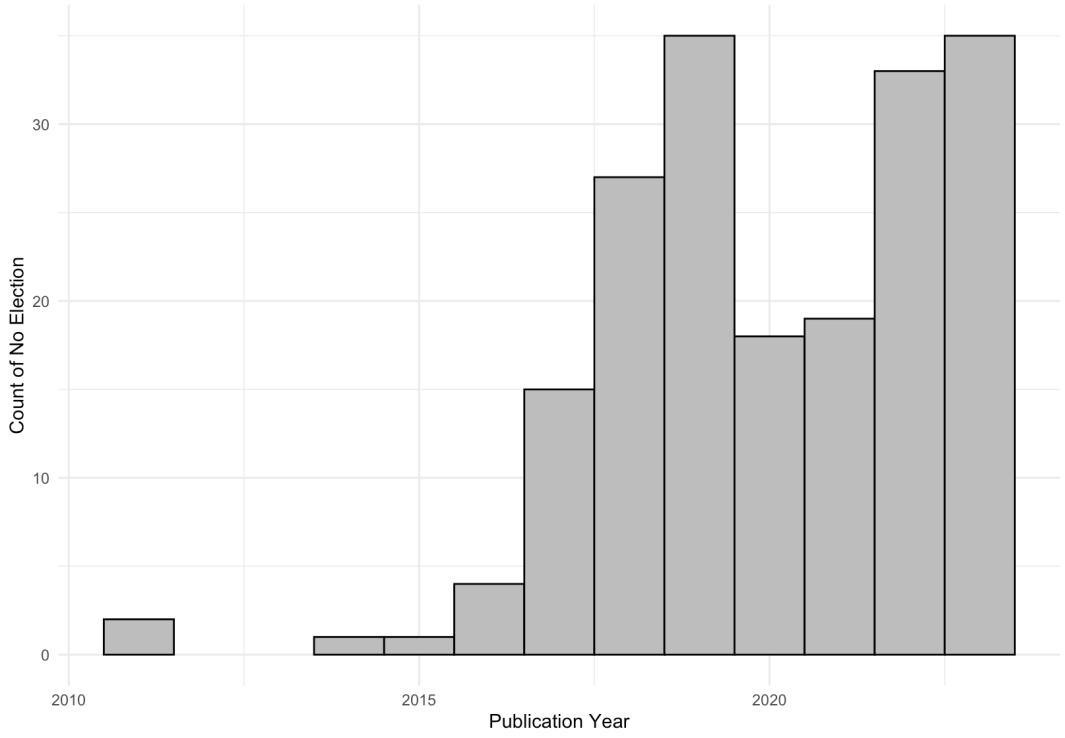

Figure 2.3: Annual Articles Not Mentioning Elections

Based on the absence of a discernable pattern in Figure 2.3, it would seem that the ultra-minority of articles making no reference to *elections* are not a new phenomenon but rather have always been present (albeit infrequently) within the discourse about state-sponsored trolls on social media.

As for the *intensive margin*, I use the *dplyr*, *stringr*, and *readr* packages to tokenize the dataset and look for total mentions of the same markers used in the *extensive margin*. In total, the 1,706 articles referring to *elections* produce a combined 13,646 mentions of elections.

Russia Extensive & Intensive Margins

In addition to examining *elections*, the database indexing also suggests a closer look at Russia is warranted. However, unlike *elections* which does not have intuitive counterparts by which to facilitate comparison, Russia can be easily contrasted with other known geopolitical belligerents active in the information environment. For the purposes of comparing extensive and intensive margins, this section will leverage China, Iran, North Korea, and Venezuela.[8]

Once again, in terms of extensive margins, the threshold for being included as a *firm* dealing in Russia, China, Iran, North Korea, or Venezuela is simply that an article must contain at least one reference to the country in question. The criteria for the markers and the results are in Table 2.3 and the distribution over time in Figure 2.4.

[8] China, Iran, and North Korea are selected due to their adversarial positions as the principal geopolitical competitors with the United States. Venezuela is included as a known state-sponsor of trolls that also provides a representative from Latin America and the Global South.

Country	Criteria	Articles
Russia	Russia, Russian(s), Moscow, Kremlin, Putin	1,674
China	China, Chinese, Beijing, CCP, Xi Jinping	578
Iran	Iran, Iranian(s), Tehran, Ayatollah	374
North Korea	North Korea, North Korean(s), DPRK, Pyongyang, Kim Jong Un	100
Venezuela	Venezuela, Venezuelan(s), Caracas, Maduro	67
Note: Because articles can reference more than one country, these markers are not mutually exclusive categories.		

Table 2.3: Extensive Margins by Country

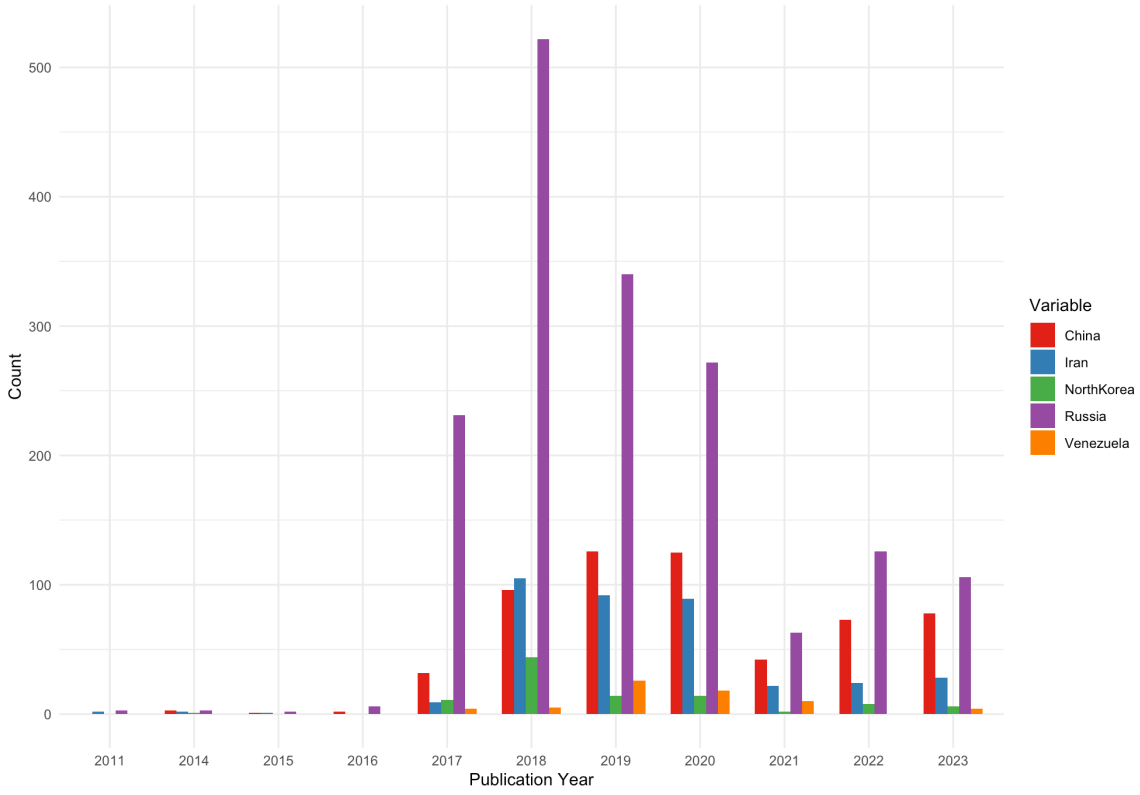

Figure 2.4: Annual Extensive Margins by Country

Russia overwhelmingly dominates the extensive margins for the majority of the dataset. Iran exceeds China in 2018 but then falls further and further behind China with time, particularly as China attempts to close the gap with Russia from 2019-2023. Likewise,

North Korea and Venezuela jockey for position relative to one another between 2018-2023; yet relative to the other countries they are both consistently less.

While Russia holds a nearly three times larger extensive margin than the next closest country (i.e., China), this differential becomes even greater when examining the intensive margins. As seen in Table 2.4, Russia's intensive margin is over six times greater than China's; nearly 20 times greater than Iran's; and more than 100 times greater than both North Korea's and Venezuela's.

Country	Criteria	Mentions
Russia	Russia, Russian(s), Moscow, Kremlin, Putin	26,826
China	China, Chinese, Beijing, CCP, Xi Jinping	4,102
Iran	Iran, Iranian(s), Tehran, Ayatollah	1,359
North Korea	North Korea, North Korean(s), DPRK, Pyongyang, Kim Jong Un	254
Venezuela	Venezuela, Venezuelan(s), Caracas, Maduro	162

Table 2.4: Intensive Margins by Country

Finding #4: Possible Emphasis on Tactics Employed in Promoting an Idea

Based upon the disproportionate emphasis on Russia (Finding #3) and the existing literature on state-sponsored trolls on social media (Chapter 1), one final set of content analysis markers was developed to try and examine whether or not the newspaper articles seem to emphasize the tactics commonly employed by Russian trolls over the tactics of others, particularly China with the next closest intensive and extensive margins.

These markers proved challenging to develop and admittedly pushed the limits of term frequency, due in part to the academic jargon of disinformation scholars not lending itself to news reporting targeted towards the average reader. For instance, *astroturfing* (i.e., mimicking grassroots movements to bolster the perceived salience of ideas) only

appeared in seven articles whereas *backstopping* (i.e., providing fake profiles with backstory details) only appeared in six articles. Further complicating marker development, other terms are co-opted from commonplace words that have different meanings in different contexts (e.g., *hijack* appeared in 55 articles but upon closer examination only two of them were specifically referring to "hashtag hijacking").

Given these circumstances, markers were developed in conjunction with the existing literature (Chapter 1) and researcher experience. More specifically, markers were developed to serve as pseudo-proxy measures for the words not used by journalists (e.g., *Hijack*, *Astroturfing*, *Backstop*). Exploratory content markers were utilized to probe the data for potential trends in three broad areas: 1) Promoting ideas (i.e., *Deceive*, *Believe*, *Propaganda*); 2) Emphasis on Russian-style profile development (i.e., *Persona*, *Engagement*, *Active Measures*, *Soviet*); and 3) Chinese tactics deviating from Russian norms (e.g., *Flood*, *Spam*). The criteria for the markers and the results are in Table 2.5.

Marker	Criteria	Articles
Deceive	Deceive(d), Deceiving, Deception(s), Trick(ed), Tricking, Fool(ed), Convince(d), Convincing, Dupe(d), Duping, Sway	602
Believe	Believe(d), Believable, Belief(s)	780
Persona	Persona(s), Impersonate, Impersonating, Fake Profile(s), Fake Account(s)	699
Propaganda	Propaganda, Propagandize	676
Soviets	Soviet(s), KGB, U.S.S.R., USSR	220
A. Measures	Active Measures	41
Flood	Flood(ed), Flooding	173
Spam	Spam, Spammy, Spamming, Spammed	174
Engagement	Likes, Reshare(s), Re-share(s), Retweet(s), Re-tweet(s), Repost(s), Re-post(s), Upvote(s), Downvote(s), Followers	595

Table 2.5: Extensive Margins by Exploratory Content Marker

The *Deceive* and *Believe* markers rank among some of the highest, but are admittedly constructed with less-precise terminological criteria which in turn limits their generalizable efficacy. Legacy references to the *Soviets* exceed both *Flood* and *Spam* while niche references to *Active Measures* are six-times those of *Backstopping* and *Astroturfing* – thereby providing an additional layer of evidence in support of Finding #3. *Flood* and *spam* rank amongst the lowest.

Based on these initial probes, the following were selected for deeper examination: *Persona, Propaganda, Engagement*, *Flood*, and *Spam*. The extensive margins were reported in Table 2.5, but Figure 2.5 provides the distribution of articles over time.

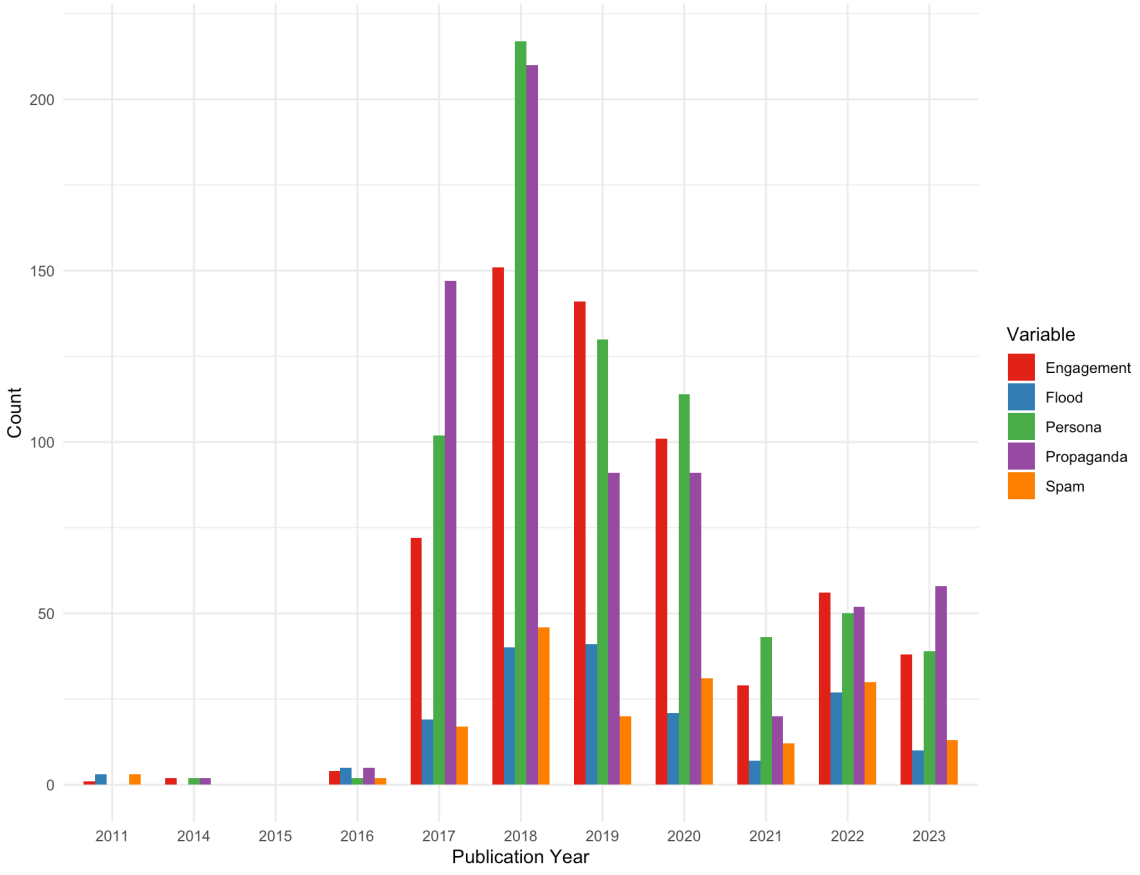

Figure 2.5: Annual Extensive Margins by Content Marker

While the patterns and differentials within this figure are not as distinct as the country extensive margins in Figure 2.4, there are things that can be gleaned from this figure. Of note, *Engagement* and *Persona* often vary together over time. However the peaks and valleys of the *Flood* and *Spam* markers do not follow the same pattern as the China extensive margins in Figure 2.4, indicating the possibility that the news reporting may make ubiquitous references to flooding and spamming in non-Chinese troll contexts.

Generally speaking, the extensive margin distribution provides possible evidence that the framing of state-sponsored trolls places more emphasis on the illegitimate account and promoting ideas than it does on known ways that Chinese trolls turn off conversations (i.e., demote ideas). The intensive margins for these markers carry similar implications (see Table 2.6).

Marker	Criteria	Mentions
Propaganda	Propaganda, Propagandize	1,314
Persona	Persona(s), Impersonate, Impersonating, Fake Profile(s), Fake Account(s)	1,264
Engagement	Likes, Reshare(s), Re-share(s), Retweet(s), Re-tweet(s), Repost(s), Re-post(s), Upvote(s), Downvote(s), Followers	1,071
Spam	Spam, Spammy, Spamming, Spammed	304
Flood	Flood(ed), Flooding	190

Table 2.6: Intensive Margins by Content Marker

The *Flood* marker's extensive margin (170) is almost the same as its intensive margin, similarly the *Spam* marker's intensive margin is less than double its extensive margin. Meanwhile the *Persona*, *Propaganda*, and *Engagement* mentions range from three to six times greater than *Flood* and *Spam* – thereby providing further evidence for a potential

framing that emphasizes fake accounts promoting ideas via generating engagement with unsuspecting social media users.

Discussion

The three principal elements to discuss from the findings are the variations in publication rate over time (Finding #1); the emphasis on elections and Russia (Findings #2-#3); and the potential emphasis on promoting ideas (Finding #4).

Variations in Publication Rate Over Time

To attempt to understand why the publication rates vary over time, it is important to contextualize them with real world events occupying the same moments in time.

2010-2016: The Rise of Social Media

Reddit was invented in 2005. A year later, Twitter came online and Facebook opened its digital doors to the general public (i.e., it stopped being only for college students with .edu email addresses). Thus, because social media platforms were on the rise and becoming increasingly mainstream around the world it reasonably follows that this period in the dataset should have the smallest number of articles – which it does.

With the advent of these nascent social networking platforms also came optimism regarding how they might enhance democracy – such as a commentator from Brookings Institution hailing the potential for "civic conversations," "citizen feedback," increased "political agency," and mechanisms for "confidence-building" (West 2011). This optimism notwithstanding, to say that the risk to free societies created by social networking was not known would be a distortion; to the contrary, the proverbial writing was already on the wall well before the U.S. elections in 2016.

Manually reviewing the dataset articles from this time period reveals numerous indications and warnings that the rise of social media brought with it vulnerabilities for U.S. national security and the interests of other like-minded nations:

- October 2011: The U.S. military "recognized Twitter as a new battlefield for information warfare" and explored the possibility of a $42M effort "to detect 'persuasion campaigns' and 'influence operations'" (Hotz 2011).

- December 2011: A journalist reports on Russia using fake accounts to praise Putin/Russia and to spoof a leader of the domestic Russian opposition so it could post fake electoral concessions (Cullison 2011).

- November 2014: Discussions of "How to Fight the Internet's State-Sponsored Trolls" emerge, making reference to Russian trolls being joined by those from China and Iran. It also makes note that these efforts "aren't supplying classic propaganda" but rather "new tactics of disinformation" (Applebaum 2014).

- September 2015: Reports surface in the context of Putin's skirmishes with dissident Alexy Navalny that "small armies of hacktivists and trolls on the Kremlin payroll got busy harassing liberals online" (Beckerman 2015).

- January 2016: A commentator offered the exhortation to "Beware of Hackers, Not Assassins" and posited that troll efforts to "dismay, divide and distract the West may be more destructive than any one assassination" (Galeotti 2016).

Ultimately, the key takeaway from this period in the dataset is that the modern problem of state-sponsored trolls on social media was not a failure of imagination but rather a failure to heed the available warnings and prepare.

2017-2020: The Focusing Event

As Birkland (2006, 33) observed in *After Disaster*: "Applying Kingdon's streams metaphor, the September 11 attacks simply focused attention on a previously existing problem stream" (i.e., transnational terrorism). The same can be said of state-sponsored trolls on social media; the Russian interference in the 2016 U.S. presidential elections directed attention towards a problem that already existed (as demonstrated in the previous

time period). In the wake of this focusing event, there is a rapid spike in news coverage in 2017 that flows directly into the dataset's apex in 2018 and ultimately results in 79% of the dataset being within this period (likely due to the Mueller Probe, the 2018 midterm elections, and the 2020 presidential election).

2021-2023: Age of Disinformation?

While the rapid rise in news coverage surrounding the focusing event is reasonable and to be expected, what sets this period apart as an outlier compared to the other two is the fact that news coverage regarding the problem of state-sponsored trolls plummets by more than 50% starting in 2021. Given that 2021 is not an election year, a decline is not completely unreasonable in and of itself – but if the drop were solely due to election patterns, this does not explain why 2022 (146 articles) is not that much more than 2021 (95 articles) or why 2022 and 2023 (135 articles) are basically the same.

What cannot go without mention is the fact that this tremendous drop in articles occurs despite experts such as Thomas Rid (2020) predicting an "age of disinformation" where information operations are a "forever war" (Nunberg 2019) due to an ever-increasing number of geopolitical adversaries utilizing low-cost digital tools such as trolls to further their national objectives at home and abroad. Similarly, the U.S. Intelligence Community assessed at the time that Iran and China desired a Biden presidency (CNN 2020) and that both used trolls to drive towards their desired outcomes (US Department of Justice 2021; Tabatabi 2018; Stone 2020; Nimmo et. al. 2020).

This begs the question: *Why are we corporately discussing the problem of state-sponsored trolls so much less at the same time experts assess that the threat is*

increasing? If this were completely a result of the COVID-19 pandemic, the decline should theoretically be in 2020 rather than 2021. While additional research is needed to provide a more definitive answer, one possibility is politicizing the problem of state-sponsored trolls in the wake of the polarizing presidency of Donald Trump – particularly during the years where conspiracy theories abounded that he was a Russian "puppet" (Wang 2019; Carpenter 2019; Lutz 2020) in the Oval Office (i.e., the problem became quantified on a scale of political winners and losers).

The highpoints of the dataset distribution roughly coincide with Robert Mueller's appointment as special counsel (May 17, 2017) and his subsequent delivery of his final report to the Department of Justice (March 22, 2019). Despite the Mueller report concluding there was no evidence of a conspiratorial agreement between the Kremlin and the Trump Campaign, news coverage continued to surge with publications in 2019 and 2020 exceeding those of 2017. Yet when Trump departs 1600 Pennsylvania Avenue, news coverage plummets by over half.

Emphasis on Elections & Russia

A potential politicization dynamic may also be a contributing factor for why *elections* are a dominant problem framing theme within this dataset – or at the very least, why they are a recurring point of contextualization by which other troll-waged campaigns are understood. While the implications of framing the problem of trolls relative to elections will be discussed in greater detail in Chapter 4, for now the data seems to support that such a framing exists given its presence in nearly the entire dataset.

As for the Russia framing, the focusing event of the 2016 elections has seemingly set conditions for Moscow to be viewed as the originator of state-sponsored trolls on social media – even though there were advanced warnings that social media created national security vulnerabilities, as discussed in the *Rise of Social Media* period above. Put another way, as mentioned previously Russia has not only been elevated as the standard by which all trolls on social media are evaluated (Warren, Linvill, and Warren 2023) but because the Russians are a type of *first-mover* then all subsequent state-sponsors of trolls must be copying them. Exemplars of this tendency from the newspaper dataset include:

- Similarly, reporters covering the 2018 midterm elections stated that "now mischief makers in other countries appear to be following the Russian playbook" before discussing Iranian exploits (Isaac & Frenkel 2018).

- One headline reads "Iranians, Others Take Russian Cue in Election Disinformation" and the article states that "many more countries had developed similar capabilities based in part on the Russian playbook" – identifying not only Venezuela, but also Saudi Arabia, Israel, and the United Arab Emirates (Timberg & Romm 2019).

- One article notes that China's activities "represent a troubling effort reminiscent of Russia's attempts to sow discord during the 2016 presidential election." This comparison becomes even more concrete when it states: "They're copying the Kremlin's playbook" (Volz 2021).

- In discussing a pro-Huawei campaign, the author overlooks Huawei's connections to the Chinese government in saying: "Tactics once used mainly for government objectives – like Russia's interference in the 2016 American presidential election – are being adapted to achieve corporate goals" (Satariano 2021).

- Amidst the 2023 Hawaii wildfires, a journalist notes that "China appeared to have adopted Russia's playbook for influence operations" (Sanger & Myers 2023).

Acknowledging that this tendency exists, Figure 2.6 shows the extensive margins for the same five countries discussed previously, only under slightly different conditions.

Russia is quantified on a 0/1 dichotomy based on making one or more mention of Russia; China, Iran, North Korea, and Venezuela are quantified on a 0/1 based on making one or more mention of themselves while also making **no** reference to Russia whatsoever Manipulating the data under these conditions identifies that 64 articles mention China without referencing Russia; 24 articles mention Iran without referencing Russia; and 2 articles respectively mention Venezuela and North Korea without referencing Russia.

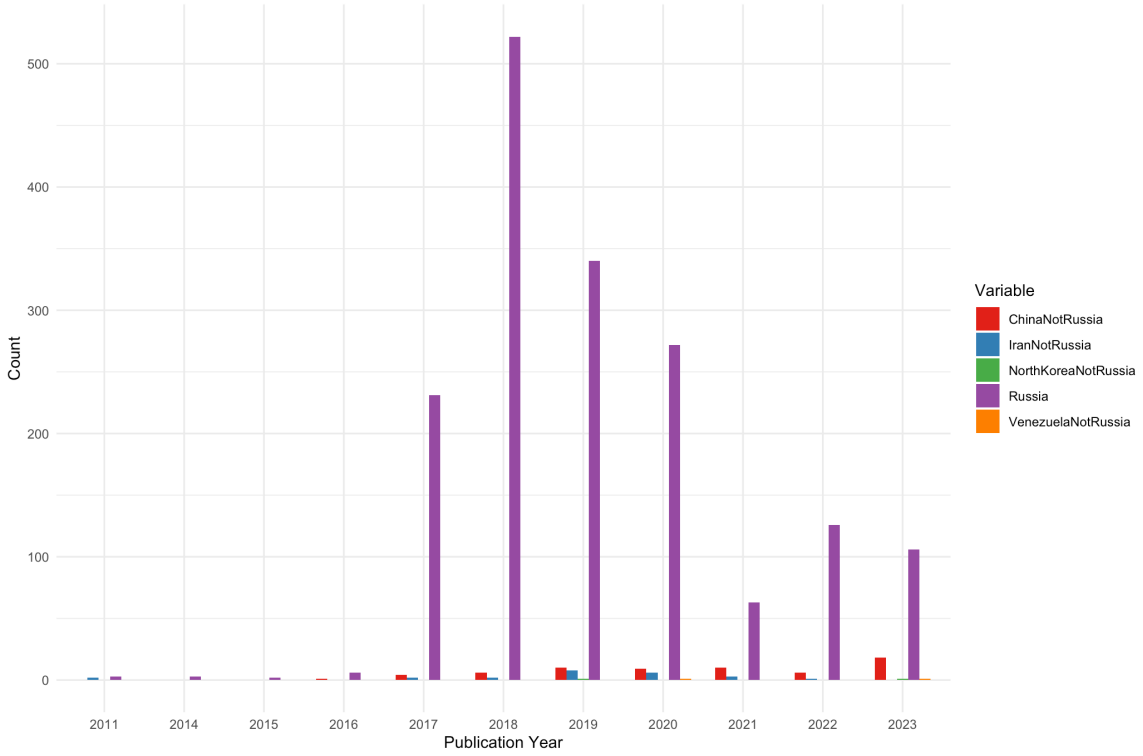

Figure 2.6: Extensive Margins Without Also Mentioning Russia

While this is likely an overcompensation from Figure 2.4 in the Findings section above (i.e., the truth probably lies somewhere in the middle), this does offer additional evidence that much of the problem framing for state-sponsored trolls is rooted in the collective experience with Russa. Put more succinctly: In many ways, there is a very real

possibility that even when people are talking about Chinese trolls…they are still talking about Russian trolls.

To be clear, the issue is not the fact that comparisons are being made. To the contrary, "explanation requires comparisons, because in order to answer the 'why' question, you need to find a rule that applies to all the members of one category, but not to the members of the other" (Markman 2019). Comparing a country to Russia is epistemologically valid when the given country is in fact doing the same thing as Russia. Yet if and when they are doing something different than Russia, making unwarranted equivocations is at times dangerous and inadvertently creates new vulnerabilities for free societies, hence why it is critical to weigh the implications of different countries using different tactics to accomplish different goals.

In an attempt to informally sanity check the validity of the aforementioned findings, the dataset was uploaded to a premium subscription of the ChatGPT 4.0 large language model (LLM). After directing ChatGPT to "read" all 1,896 articles in the dataset, the model was asked to identify five themes and to sort the articles into the themes it developed (i.e., ChatGPT was essentially directed to engage in *emergent coding*). Table 2.7 contains the theme names and descriptions that ChatGPT developed independently of any researcher inputs as well as the number of articles sorted into each.

Theme Name	Theme Description	# of Articles
Social Media & Politics	Involves discussions on platforms like Facebook and Twitter, particularly in relation to political influences and Russian accounts	634
Russian Influence & Cybersecurity	Focuses on Russian activities, cybersecurity issues, and political figures like Putin	483
U.S. Political Campaigns	Centers around key U.S. political figures like Trump and Clinton, and includes discussions on campaigns and Russian influence	407
Tech Industry & Public Policy	Concerns speeches from tech industry leaders and interactions with public policy, including hearings and regulations in the tech sector	216
China's Geopolitical Influence	Covers topics related to China, including statements by Chinese officials, and Beijing's influence in geopolitical matters	156

Table 2.7: ChatGPT 4.0 Emergent Coding Results

Consistent with Findings #2 and #3, the LLM was so inundated with references to *elections* and *Russia* that it produced three overlapping categories and distributed 80% of the dataset across them, leaving policy solutions to address the problem of trolls (11% of articles) and the threat of China (9% of articles) as the secondary and tertiary themes.

Possible Emphasis on Idea Promotion & Engagement

As discussed in Chapter 1, there is an underlying tendency within the body of academic literature on state-sponsored trolls to emphasize account appearances and how trolls interact with users to promote ideas. Finding #4 offers some preliminary evidence that this trend may also exist in newspaper reporting as it appears to place emphasis on fake accounts promoting ideas and generating user engagement.

This evidence becomes potentially more apparent when appropriately contextualized within the limitations of term frequency analysis. First, the engagement marker is likely an underestimate given that otherwise valid engagement metric terms

were conservatively omitted to avoid false positives (e.g., *views* would trigger on "political views;" *engagement* would trigger on "Global Engagement Center;" the singular *like* and *share* are commonplace; the singular *comment* would hit for a reporter's "request for comment"). When the marker included variations of such words in the 0/1 criteria the resulting number of articles was over 1,000 – some likely valid and others invalid, hence the more restrictive inclusion criteria that still identified 595 articles.

Second, the quantitative delta between *Persona*/*Propaganda*/*Engagement* and *Flood*/*Spam* is underrepresented in Finding #4 above on the basis that the *Flood* and *Spam* markers overestimate discussions of Chinese tactics specifically. This can be observed by, similar to the previous section, adding some additional criteria to the annual distributions. Figure 2.7 depicts the total number of articles mentioning China as well as the same *Flood* and *Spam* markers presented previously; the *ChinaFlood* and *ChinaSpam* variables, however, are those articles that satisfy the criteria for both *China* and *Flood* (60 articles) or *China* and *Spam* (61 articles) respectively.

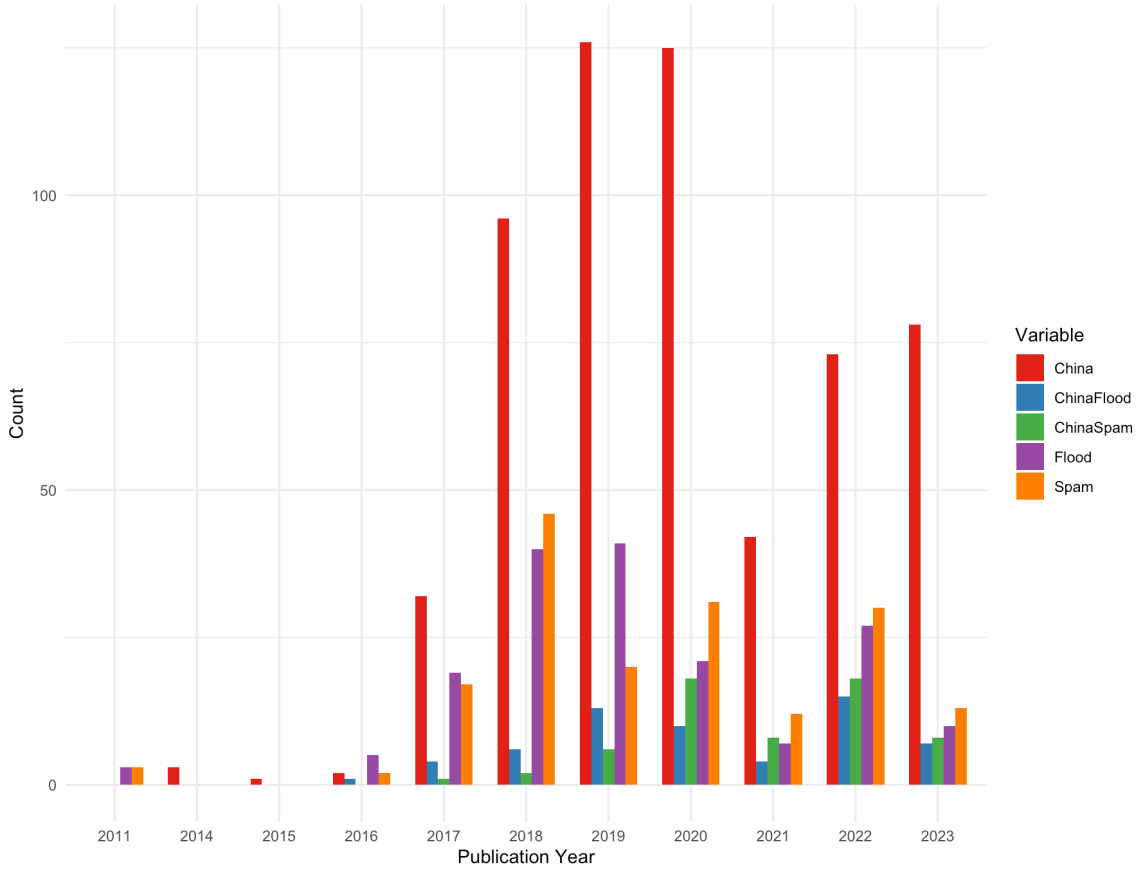

Figure 2.7: Extensive Margins for China References

From 2017-2019 the vast majority of articles containing *Spam* or *Flood* make no reference to China, suggesting generally imprecise and ubiquitous uses of these terms vis-à-vis specific references to tactics (e.g., "Russian operatives that flooded Facebook with disinformation around the 2016 election" [Dwoskin & Romm 2018a]; "spamming Facebook groups with identical pieces of content and using fake profiles" [Dwoskin & Romm 2018b]). It is not until 2021-2023 that these markers begin to covary with half or more mentions of *Spam* or *Flood* also mentioning *China* – albeit in a small number of articles by the time this alignment emerges. Moreover, while this co-occurrence does

begin to suggest that the discourse shifts over time to associate these tactics with the Chinese, relative to the overarching number of articles mentioning *China* discussions of specific tactics seem to be only in an extreme minority of news articles. This in turn does seem to provide some corroborating evidence that the framing of state-sponsored trolls places more consistent and monolithic emphasis on fake accounts promoting ideas and generating user engagement.

Once again, as an informal sanity check of my assessment, ChatGPT was asked to explain *What is an Information Operation?* based on the news articles within the dataset. ChatGPT's answer was twofold. First it responded with the following paragraph:

> Based on the context of the articles, especially those that discuss topics like political interference and online manipulation, an "Information Operation" refers to coordinated activities aimed at influencing, confusing, or controlling public opinion through the dissemination of propaganda, disinformation, or misinformation. These operations are typically conducted by state actors or organized non-state groups and are intended to achieve strategic and political objectives.

Second, ChatGPT went on to provide what it considered to be five key features of an Information Operation (consolidated in Table 2.8).

Key Feature	Feature Description
Strategic Intent	Unlike casual dissemination of false information, information operations are deliberate and strategically timed to influence specific events, such as elections, political decisions, or public sentiments on contentious issues.
Multi-Platform Coordination	These operations often utilize a variety of media channels—social media, traditional news outlets, fake news websites, and sometimes even hacking and leaks—to spread their narratives and maximize their reach.
Psychological Warfare	The goal is often to sow confusion, distrust, and fear among the population, making it more difficult for individuals to discern truth from falsehood, thereby weakening the societal cohesion and response to the operations.
Manipulation Tactics	Information operations frequently employ sophisticated psychological tactics to appeal to emotional responses, biases, or prejudices. This can include framing narratives in ways that exploit existing tensions within communities or nations.
Targeted Campaigns	These operations can be highly targeted, aiming at specific demographic or psychographic segments of a population to exploit known vulnerabilities or influence key opinion leaders.

Table 2.8: ChatGPT 4.0 Key Features of Information Operations

Consistent with the indications in Finding #4, ChatGPT's synthesis of the news article dataset arrives at similar conclusions. The five key features especially very easily accommodate a "strategic," Russian-style approach to "sophisticated" fake accounts promoting ideas and appealing to "emotional responses, biases, or prejudices" to generate user engagement within "highly targeted" audiences – but these features offer no substantive insights into the concept of trolls being used to demote ideas or engage in a form of algorithmic censorship. Although more research is necessary to demonstrate to what degree Finding #4 is a dominant framing of trolls, for the purposes of this dissertation it is my contention that this chapter demonstrates that sufficient evidence

exists to at least incorporate the implications of such a framing in the research design and discussion of subsequent chapters.

Bridging the Gap Between Media & Policymaking

While the Data section above provides literature-derived justification for using media reporting as a proxy measurement of an issue's salience within policy communities and the broader societal agenda, it would be epistemologically valuable to have additional data from inside the government so as not to inadvertently imply the media framings are the sole catalysts driving policy formulation. After all, if the problem framings conveyed by the media also appear in the machinations of governmental proceedings then there can be an even higher degree of empirical confidence in Chapter 4's efforts to answer the overarching research question (i.e., *What is the Relationship Between How the Problem of State-Sponsored Trolls on Social Media is Framed & the Way U.S. Federal Policy Attempts to Solve It?*). With regards to state-sponsored trolls on social media, evidence from Congressional hearings would suggest that the aforementioned trends transcend the news and are the prevailing problem framings.

Using the same selection criteria in Table 2.1 on the federal repository *govinfo.gov* yields congressional hearings from 2010-2023 (N = 322) that are reasonably assumed to contain a discussion of state-sponsored trolls. In examining the extensive margins of hearing transcripts mentioning certain countries, the overall shape of the distribution (Figure 2.8) resembles that of the news articles (i.e., a low period, a high period, and a medium period). The comparative intensities of countries also resemble the news reporting dataset with the top three being Russia, China, and Iran respectively.

However there is one notable exception to the resemblance between Figure 2.4 and 2.8 – in 2021 and 2023, China's extensive margin exceeds that of Russia.

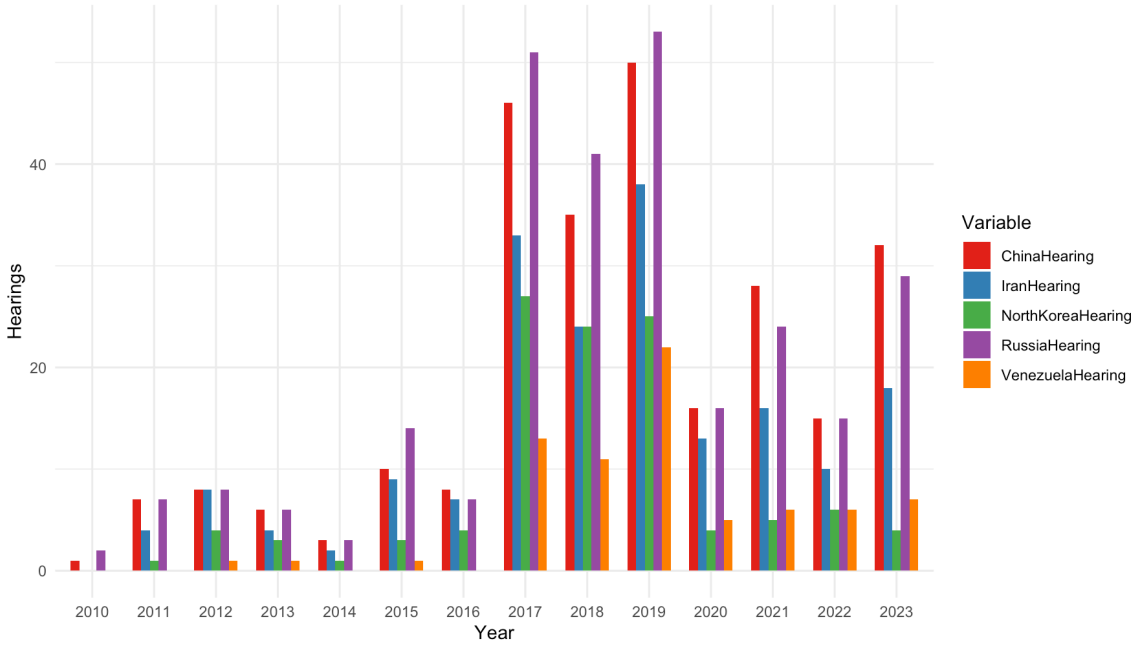

Figure 2.8: Hearing Extensive Margins by Country

In light of this difference, the intensive margins become increasingly important to understanding the problem-framing within the discourse of these hearings. Although Iran is ranked third in the extensive margins, Table 2.9 differs from Table 2.4 in that North Korea and Venezuela both exceed Iran's intensive margin. More importantly, despite the parity within the extensive margins, the intensive margins confirm that Russia is discussed roughly two-times more than China (Table 2.9). Given that technological hearings can cover a range of topics with one visit from an expert (e.g., data privacy, child exploitation, telecommunication infrastructure), it is possible the delta between Russia and China is actually much larger as it pertains to state-sponsored trolls waging

information operations on social media. ChatGPT 4o suggests such an overestimate – after consuming all 322 transcripts, the LLM produced Table 2.10 in response to being prompted to summarize the difference between how Russia and China are discussed in proceedings on Capitol Hill.

Country	Criteria	Mentions
Russia	Russia, Russian(s), Moscow, Kremlin, Putin	32,292
China	China, Chinese, Beijing, CCP, Xi Jinping	17,187
North Korea	North Korea, North Korean(s), DPRK, Pyongyang, Kim Jong Un	3,287
Venezuela	Venezuela, Venezuelan(s), Caracas, Maduro	1,869
Iran	Iran, Iranian(s), Tehran, Ayatollah	1,359

Table 2.9: Hearing Intensive Margins by Country

Country	Context
Russia	Disinformation and Influence: Russia is frequently mentioned in the context of disinformation campaigns and influence operations targeting the United States and other countries.
	Cybersecurity Threats: Discussions often involve Russian cyber activities, including hacking and cyber espionage.
	Political Interference: There are references to Russian attempts to interfere in political processes and elections.
	Geopolitical Tensions: Russia is discussed in terms of its geopolitical actions, such as military activities and foreign policy maneuvers.
China	Economic Competition: China is often mentioned in the context of economic competition and trade issues, including intellectual property theft and unfair trade practices.
	Geopolitical Influence: Discussions include China's efforts to expand its influence globally through initiatives like the Belt and Road Initiative (BRI).
	Human Rights: China is frequently mentioned in relation to human rights concerns, including the treatment of ethnic minorities and political dissidents.
	Technology and Espionage: There are references to China's technological advancements and concerns about espionage and data security.

Table 2.10: ChatGPT Hearing Content Analysis

The first three ChatGPT context descriptions for Russia all point to the issue of state-sponsored trolls on social media (e.g., disinformation, hack-and-leak, political interference) while none of the China contexts would seem to indicate a similar emphasis. Additionally, the third context for Russia points to the possibility of framings around elections. Applying the same extensive and intensive margins criteria as before, 300 of the 322 hearings contain at least one reference to elections and in total there are 19,589 total references (i.e., there are more *election* mentions than there are *China* mentions). Thus, it does appear the same emphasis on *Russia* and *elections* that exists in the news dataset also exists within Congressional hearings.

Lastly, with regards to a possible emphasis on believable accounts promoting ideas, there is some evidence to suggest that such a framing exists (Table 2.11). Even though *Flood* is higher in Table 2.11 than in Table 2.6 (likely an overestimate based on generic usage of the term), the sustained use of *Propaganda* and *Engagement* (likely an underestimate for the same reasons stated previously) would suggest a framing of idea promotion – an assertion further justifiable considering the mentions of *disinformation* (3,484) and *misinformation* (1,602). While a more in-depth study would be necessary to know for certain, for the purpose of this dissertation's overarching research question the dominant problem framings across both the newspaper dataset and the Congressional hearings appear to be elections, Russia, and promoting ideas via generating engagement.

Marker	Criteria	Mentions
Propaganda	Propaganda, Propagandize	1,874
Flood	Flood(ed), Flooding	560
Engagement	Likes, Reshare(s), Re-share(s), Retweet(s), Re-tweet(s), Repost(s), Re-post(s), Upvote(s), Downvote(s), Followers	518
Persona	Persona(s), Impersonate, Impersonating, Fake Profile(s), Fake Account(s)	450
Spam	Spam, Spammy, Spamming, Spammed	282

Table 2.11: Hearing Intensive Margins by Content Marker

Limitations & Way Forward

Blending a term frequency approach to content analysis with the economics principles of extensive and intensive margins enabled this study to efficiently parse a large text corpus dataset (N = 1,896 articles) for contextual markers that begin to answer RQ_1: *How is the Problem of State-Sponsored Trolls on Social Media Framed?* Yet in identifying the emphases on elections, Russia, and fake accounts promoting ideas through user engagement, the principal limitation of this approach is that it offers no insights into the tone or sentiment within the discourse surrounding these framings. For instance, while this approach tells us that the gap between the number of articles mentioning Russia and the number of articles mentioning China has been substantially reduced in recent years, it is unable to reveal whether or not China and Russia are talked about in the same manner (e.g., Since content markers suggest the possibility that comparisons to Russia are often made, are there value-laden judgments that assert some trolls are better than others? If so, why is it that one country's trolls are superior or inferior?).

For now, this chapter's principal contribution is identifying the apparent emphasis on elections, Russia, and fake accounts promoting ideas while generating user

engagement within the discourse of state-sponsored trolls on social media. The inability of term frequency to identify tone and sentiment in this chapter will be mitigated in Chapter Three by deliberate citations from beyond the newspaper dataset to further explore problem framing trends as they relate to how state-sponsored troll accounts are operationalized to conduct information operations in the real world as opposed to solely their portrayal in public discourse.

CHAPTER THREE

WE HOLD THESE LIES TO BE SELF-EVIDENT, NOT ALL TROLLS ARE CREATED EQUAL

"Chinese operators do not appear to have done the psychological or ethnographic research required to create convincing accounts on Western platforms. They additionally appear relatively unconcerned about getting no engagement" (DiResta et. al. 2022, 44). Embedded within this quote from the Stanford Internet Observatory are two underlying assumptions that merit open discussion:

1) The Chinese Communist Party (CCP) operators are systematically ignorant with regards to how to conduct information operations effectively (i.e., generate engagement); and

2) The CCP operators are entrenched in maintaining extremely high rates of information output to such a degree that they are impervious to the feedback that what they are doing is not generating engagement (e.g., likes, re-tweets, shares) within a pre-specified target audience and is therefore (supposedly) ineffective.

While the accuracy and prudence of such an assessment is questionable, it is nevertheless an increasingly mainstream position.

Because of the pervasive presupposition all social media troll's prime directive is to promote ideas and generate real-user engagement, scholars and commentators alike have a predisposition to label the Chinese troll apparatus as "more simplistic" (Kurlantzick 2020), "more primitive" (Huang 2019), or "still improving" (Harold, Beauchamp-Mustafaga, Hornung 2021, 40). In some instances one does not even have to look beyond the headline and subtitle, such as *Wired* magazine's "Why China Is So Bad at Disinformation: China's State-Sponsored Disinformation Campaign Has Been Running at a Massive A Scale for Seven Years – but No One is Looking at it" (Gilbert

2024). Others describe this genre of activities as a "post and pray" strategy that drops content throughout the social media ecosystem in the hopes that an otherwise unspecified audience will stumble upon it (Nimmo & Hutchins 2023) – an assessment incorrect on two counts: 1) The CCP members are staunch atheists that do not pray; and 2) it reductively assumes that the objective is to promote ideas and generate engagement.

Yet if the actual goal is more analogous to the *Theory of Censorship* (Roberts 2018) discussed in Chapter 1 and these expendable CCP accounts are turning conversations off, then flooding can be likened to barrage noise jamming in electronic warfare (i.e., "blinding a system by filling the display with noise" [Linvill & Warren 2021b; see also Warren et. al. 2023]). More importantly, if the goal is to turn off conversations than it necessarily follows that engagement is an irrelevant metric for gauging effectiveness – in fact, the absence of engagement may even demonstrate such an information operation is succeeding rather than failing.

Thus, the present chapter seeks a more holistic understanding of the challenges presented by state-sponsored trolls through the following research question:

- RQ_2: *How are State-Sponsored Trolls on Social Media Operationalized to Conduct Information Operations?*

In order to address this primary question, the following two sub-questions are posed:

- $RQ_{2.1}$ = *What are Ways State-Sponsored Trolls Invest in the Development of their Online Personas?*

- $RQ_{2.2}$ = *What is the Relationship Between the State-Sponsored Troll Persona Investment and the Goals of the Information Operations they Conduct?*

These questions are deliberately structured to pursue generalizable findings that could empirically test the notion state-sponsored trolls and the information operations they

wage are not a monolithic phenomenon but rather display considerable heterogeneity to the point of needing a taxonomical framework.

Data

Data inaccessibility is arguably the single greatest impediment to meaningful research into the problem of state-sponsored trolls on social media – particularly given the inherent aspect of deception within the phenomenon. As mentioned in Chapter 2, classified threat reporting is unavailable for obvious reasons and the social media platforms have no incentive to make timely ground-truth information available given the potential adverse effects that highlighting vulnerabilities can have on corporate bottom lines. Although there has been at least one Russian troll defector to provide first-hand testimony regarding the tradecraft behind some surreptitious accounts (Troianovski 2018), such occurrences are extremely rare and they have limited potential energy for broader generalizability.

This data shortage notwithstanding, it is important to keep in perspective that these types of hurdles are not unique to troll researchers. To the contrary, sociological research examining deviant behaviors (e.g., gangs, organized crime, prostitution, drug usage, cyber bullying) must go to great lengths to effectively collect data while simultaneously insulating it from corrosive second-order effects (e.g., the Hawthorne Effect where participants change their behavior as a result of their awareness of being observed). With this in mind, the present chapter uses two secondary datasets: the Twitter Information Operations Archive in its entirety and a cross-section of the Empirical Studies of Conflict's (ESOC) Trends in Online Influence Efforts. Both

datasets are the best of their kind – but since "their kind" is an objectively small peer group, it is critical to understand the strengths and weaknesses of these datasets as it pertains to this research endeavor.

Twitter Information Operations Archive

In October 2018, Twitter began operating the first large-scale data repository of state-sponsored information operations through its Twitter Moderation Research Consortium. Within this archive are 46 individual datasets representing more than 35 known operations and/or clusters of nefarious networks that violated the platform's content/behavior standards. In total, this archive represents the coordinated activities of 87,437 individual troll accounts sponsored by 20 different countries and tweeting more than 200 million times (over nine terabytes of multimedia).

The magnitude and robustness of this dataset notwithstanding, there are objective data quality issues in need of acknowledgement. First, Twitter (pre-Elon Musk and post-Elon Musk) has not disclosed its methodological approach or philosophical rationale behind which trolls do and do not find themselves in the public disclosure repository. Similarly, Twitter also does not disclose the tactics, standards, or practices that are employed in making formal attribution of the state sponsoring the trolls. Lastly, the lack of transparency on the part of Twitter has also fostered circumstances where there are a few accounts belonging to real individuals that are erroneously included (Linvill & Warren 2020, 449) and other attribution mistakes have occurred (Elgin 2019). That said, in spite of these aforementioned issues regarding the construction and underlying nature of the information operation archive, it remains the greatest troll dataset of its kind.

The preeminence of this dataset naturally implies that it is the most common source for data on information operations (which it is [Cima et. al. 2024]), but as discussed in Chapter 1 the majority of these ventures are case studies whereby researchers examine specific trolls (more often than not, Russian trolls) operating in specific moments in space-time (often before and/or after elections). There is nothing inherently wrong with this approach, but the limitation of disparate case studies is that they do not move the proverbial needle towards generalizable findings about the problem of state-sponsored trolls on social media holistically. To avoid further contributing to this particular limiting factor within the existing body of literature, the research design of this chapter intentionally shifts the focus from the behavior of trolls to the actual accounts themselves in order to compare and contrast like kinds across the various state sponsors.

More specifically, this chapter takes all 46 individual datasets and combines them into one comprehensive dataset where each observation is an account (i.e., the unit of analysis is individual trolls, N = 87,437). To my knowledge, this study is only the second research study leveraging the entirety of the Twitter Information Operation Archive as a singularly unified dataset (Linvill et. al. 2024). In adopting this approach with my data, it is my intent to produce generalizable findings that can extend beyond the monolithic and homogenous framing of state-sponsored trolls discussed in Chapter 1 and Chapter 2.

ESOC Trends in Online Influence Efforts

The Empirical Studies of Conflict maintains a database of information operations entitled “Trends in Online Influence Efforts” (Martin, Shapiro, & Ilhardt 2020/2023). This database is comprised of secondary reporting about information operations globally

and includes descriptions of state-sponsors, targeted populations, which social media platforms the campaign took place on, topics emphasized, digital tactics, operational goals, and overarching strategies. Similar to the Twitter Information Operations Archive, this too is the best and only dataset of its kind…but there are issues here as well.

The principal limitation of this dataset is that it is entirely derived from secondary reporting. Consequently, the various information operation attributes coded by ESOC are indirectly inferred rather than directly known. Put another way, had this data come from interviews or focus groups with troll farm operators we might have a certain degree of confidence in qualitatively coding goals because the trolls themselves would be offering first-hand accounts of why they did what they did. Instead, the ESOC dataset is entirely dependent upon what third-party observers (mostly journalists) think trolls are doing based solely on observing account behaviors and messaging. This in turn means that the dataset is potentially tainted with circular presuppositions (e.g., *Trolls curate accounts to be believable because their goal is to be believed*), the homogenous framing tendencies identified via the newspaper dataset in Chapter 2, and the superficial assumptions in the introduction above that contend a troll's efficacy lives or dies by how much real user engagement it produces. The limitations notwithstanding, the ESOC dataset remains the best and only one of its kind.

For the purpose of the present chapter, $RQ_{2.2}$ (*What is the Relationship Between the State-Sponsored Troll Persona Investment and the Goals of the Information Operations they Conduct?*) is probed by leveraging the data my colleagues and I produced in the first study exploiting the entirety of the Twitter Information Operation

Archive (Linvill et. al. 2024). If ESOC mentioned Twitter as being involved in a given information operation, we considered it a qualitative match for a Twitter Information Operation Archive release if the country of origin (i.e., state-sponsor) matched and the Twitter accounts used qualifying keywords in their tweets that corresponded with the ESOC campaign description. Fusing the Twitter and ESOC datasets in this manner resulted in our pairing 27,701 individual troll accounts with amplifying descriptive information from ESOC.

Method

Methodologically, these large datasets are examined in a quantitative manner using Ordinary Least Squares (OLS) regression models. The dependent variable is a novel measurement scale for account investment operationalized by means of the Twitter Information Operation Archives dataset whereas the independent variables are dummy variables representing each country of origin and each campaign goal as derived from either the Twitter or the ESOC datasets.

Account Investment Dependent Variable

The dependent variable is a representation of the aggregated decisions made by the troll operator in conjunction with constructing the Twitter profile (i.e., when the troll was created, the operator chose to give the account certain attributes). Because account *investment* possesses an inherent level of subjectivity and abstraction, this dependent variable is operationalized using a first-of-its-kind scoring system that examines each individual troll profile along six separate markers:

- ❖ The troll persona has a profile description/biographical data (Y/N).
- ❖ The troll persona has claimed a geographic location (Y/N).
- ❖ The troll persona has provided a personal URL (Y/N).
- ❖ The troll persona has at least 10 other Twitter accounts it follows (Y/N).
- ❖ The troll persona has at least 50 other Twitter accounts it follows (Y/N).
- ❖ The troll persona has at least 100 other Twitter accounts it follows (Y/N).

Each one of these novel *investment* components are *objectively* discernable at the individual account level (i.e., the account either has the characteristic or it does not) rather than being *subjectively* derived from the account's behavior. This ensures consistency in both operationalizing the variable and measuring it across all observations.

Once these markers are each scored based on present/not-present criteria where a YES = 1 and a NO = 0, the markers are aggregated using the following formula to produce an *investment* score ranging from 0 to 6:

Description + Location + URL +Follow 10+ Follow 50 + Follow 100 = Investment

With each individual troll scored accordingly, the *investment* score is then treated as a continuous variable[9] in order to analyze the variation in scores relative to the independent variable. The univariate *investment* score summary statistics can be seen in Table 3.1 and the corresponding distribution can be seen in Figure 3.1.

[9] Treating the integer-based scale as a continuous variable does introduce some heteroscedasticity to the dependent variable, in doing so it allows the OLS regressions to yield an analysis of variance (ANOVA) for investment score relative to the independent variables.

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
All Trolls	0	0	2	2.107	4	6

Table 3.1: Investment Score Descriptive Statistics

Figure 3.1: Distribution of Investment Scores

State Sponsorship Independent Variable

The state sponsorship independent variable is defined as the nation state sponsoring the troll (i.e., the troll's country of origin). State-sponsorship is operationalized on a nominal scale as a categorical variable with 20 distinct options: Armenia, Bangladesh, China, Cuba, Ecuador, Egypt, Honduras, Indonesia, Iran, Mexico,

Russia, Saudi Arabia, Serbia, Spain,[10] Tanzania, Thailand, Turkey, Uganda, UAE, and Venezuela. Because each individual observation has a directing country attributed to it by Twitter, the independent variable objectively self-sorts to yield the number of trolls per country within the entirety of the Twitter dataset. The univariate *state sponsorship* distribution can be seen in Figure 3.2.

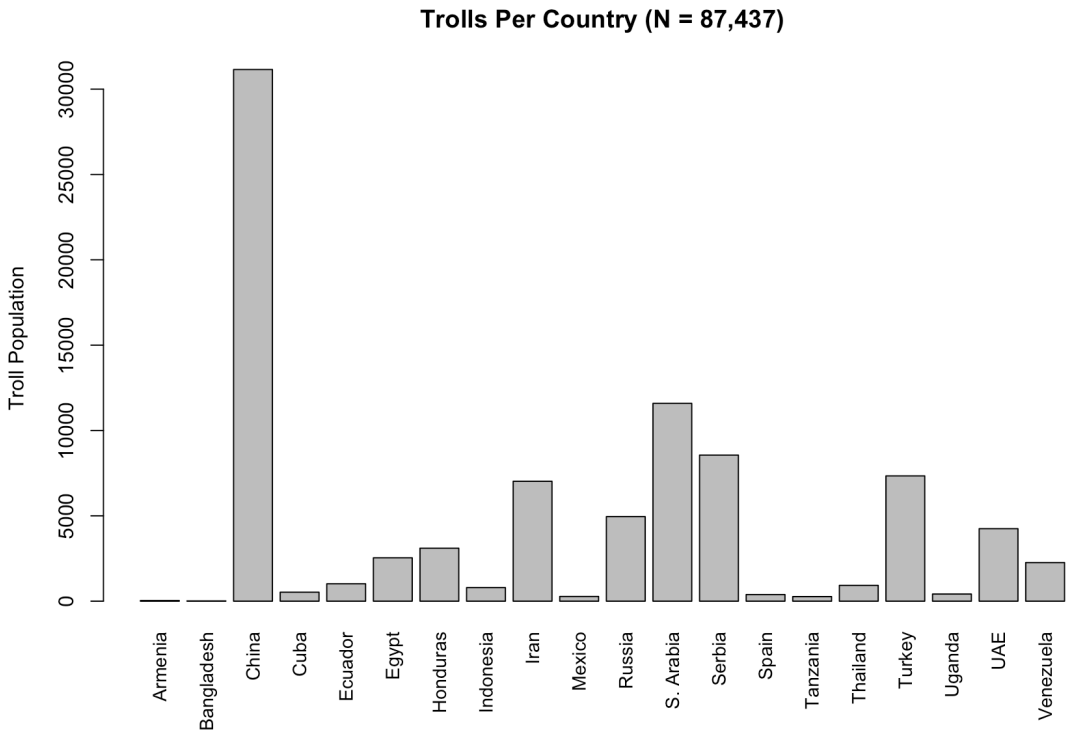

Figure 3.2: Distribution of Trolls Per Country

[10] The trolls attributed to Spain are actually operated by the Catalonian independence movement. They will be referred to as "Spain" throughout because that is nomenclature Twitter attaches to these accounts. Even though these trolls are not sponsored by Madrid, since the Catalonia freedom movement functions as a pseudo government-in-exile with political objectives it is not empirically disingenuous to include them as a state-sponsor for the purpose of seeking generalizable findings about state-sponsored trolls.

Information Operation Goal Independent Variable

Operationalized from the information operations tactics matrix from Chapter 1 (Linvill & Warren 2023), the *goal* independent variable is conceptualized as either promotional or demotional. This variable is derived from the qualitative coding in the ESOC database. If an information operation is coded as having the goal Influence, Support, or Spread, then the corresponding trolls are coded as promotional (i.e., increasing the attention of a set of ideas). Conversely, if an information operation is coded as having the goal Discredit or Hinder, then the corresponding trolls are coded as demotional (i.e., decreasing the attention of a set of ideas). Since it is possible that troll accounts can be used to contribute to multiple information operations, the operationalization of promotional versus demotional is not mutually exclusive.

Hypotheses & Statistical Models

In the spirit of where this dissertation began (i.e., "Imperfect understanding is often more dangerous than ignorance"), I argue the aforementioned claim that all trolls need to have a certain appearance or meet an otherwise unspecified standard of sophistication in order to be effective is illogical on the basis that it is grounded in the assumption all trolls seek to influence audiences through direct engagement. Just as the potential applications of trolls are limitless and only constrained by the imagination of their operators, it is premature to discount the potential efficacy of unsophisticated trolls without definitive evidence to the contrary. Thus, this chapter further theorizes that trolls are not a homogenous phenomenon but rather a heterogenous one in desperate need of taxonomical classification.

To begin assessing said theory, I pair this chapter's research questions with null hypothesis testing statistical models. For $RQ_{2.1}$ (*What are Ways State-Sponsored Trolls Invest in the Development of their Online Personas?*), I hypothesize that countries will invest differently from one another as opposed to demonstrating homogenous patterns that mirror one another across the various components of the investment score variable:

- H_1 = *There is a significant difference between the troll account investments from the various state-sponsors of trolls.*

Testing this hypothesis is done via the following OLS regression model with 19 categorical dummy variables and China as the reference category:

$$\text{Investment}_i = \beta_0 + \beta_1\text{Armenia}_i + \beta_2\text{Bangladesh}_i + \beta_3\text{Cuba}_i + \beta_4\text{Ecuador}_i + \beta_5\text{Egypt}_i + \beta_6\text{Honduras}_i + \beta_7\text{Indonesia}_i + \beta_8\text{Iran}_i + \beta_9\text{Mexico}_i + \beta_{10}\text{Russia}_i + \beta_{11}\text{SaudiArabia}_i + \beta_{12}\text{Serbia}_i + \beta_{13}\text{Spain}_i + \beta_{14}\text{Tanzania}_i + \beta_{15}\text{Thailand}_i + \beta_{16}\text{Turkey}_i + \beta_{17}\text{Uganda}_i + \beta_{18}\text{UAE}_i + \beta_{19}\text{Venezuela}_i + u_i \quad \text{(Equation 1)}$$

In the event the null hypothesis for $RQ_{2.1}/H_1$ is rejected, this chapter further theorizes that understanding the ways state-sponsors invest in their trolls might begin to offer clues regarding the goals of information operations. Accordingly, for $RQ_{2.2}$ (*What is the Relationship Between the State-Sponsored Troll Persona Investment and the Goals of the Information Operations they Conduct?*), I hypothesize that different goals will correlate with different troll investment scores:

- H_2 = *There is a significant difference between the troll account investments for promotion information operations versus demotion information operations.*

Testing this hypothesis is done via the following OLS regression model with two categorical variables and the reference category is those accounts we were unable to make a definitive attribution of goal:

$$\text{Investment}_i = \beta_0 + \beta_1\text{Promotion}_i + \beta_2\text{Demotion}_i + u_i$$

(Equation 2)

Findings

The outputs of the statistical models revealed definitive evidence that countries invest differently in the trolls that they sponsor and that there is a relationship between investment and an information operation's goal(s) – which stands in stark contrast to the conventional wisdom that frames the problem of state-sponsored trolls as a monolithic/homogonous phenomenon (Chapter 1 and Chapter 2).

Finding #1: Countries Invest Differently

The F-statistic of the regression has a p-value of <.01 which triggers a rejection of the null hypothesis that all of the coefficients are equal to zero. Analyzing the bivariate distribution of *investment* score relative to state-sponsor and the results of the OLS model for $RQ_{2.1}/H_1$ reveal considerable heterogeneity in the ways troll personas are cultivated. Table 3.2 provides the bivariate summary statistics for each country's *investment* score and Figures 3.3-3.22 visually depict the *investment* score distributions by country.

	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Armenia	2	3	5	4.314	5	6
Bangladesh	1	2	3	3.000	4	6
China	0	0	0	0.4241	0	6
Cuba	0	4	4	3.922	5	6
Ecuador	0	0	1	1.802	3	6
Egypt	0	1	3	2.886	4	6
Honduras	0	0	1	1.541	2	6
Indonesia	0	1	3	2.804	4	6
Iran	0	2	4	3.403	5	6
Mexico	0	3	3	3.261	4	6
Russia	0	3	4	3.909	5	6
Saudi Arabia	0	3	4	3.493	5	6
Serbia	0	1	2	2.309	3	6
Spain	0	1	3	2.714	4	6
Tanzania	0	0	1	1.250	2	5
Thailand	0	0	0	0.703	1	4
Turkey	0	2	4	3.244	5	6
Uganda	0	3	4	3.823	5	6
UAE	0	2	3	3.064	5	6
Venezuela	0	2	3	3.408	5	6

Table 3.2: Investment Summary Statistics by Country

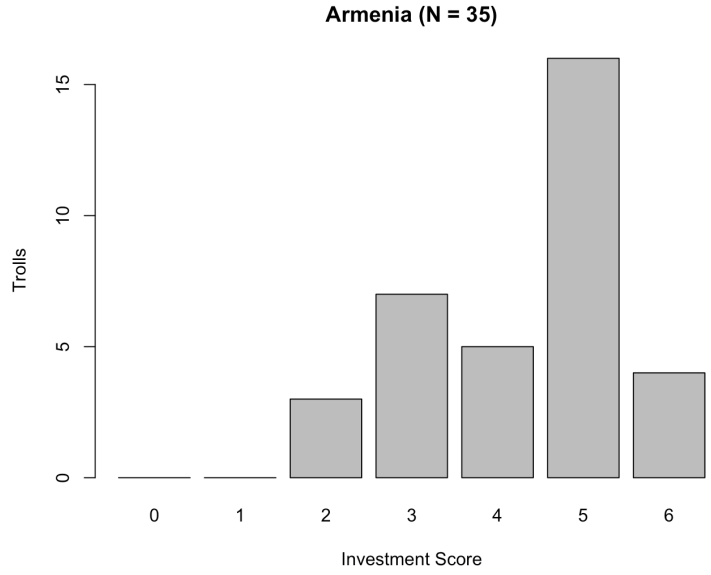

Figure 3.3: Armenia Investment Distribution

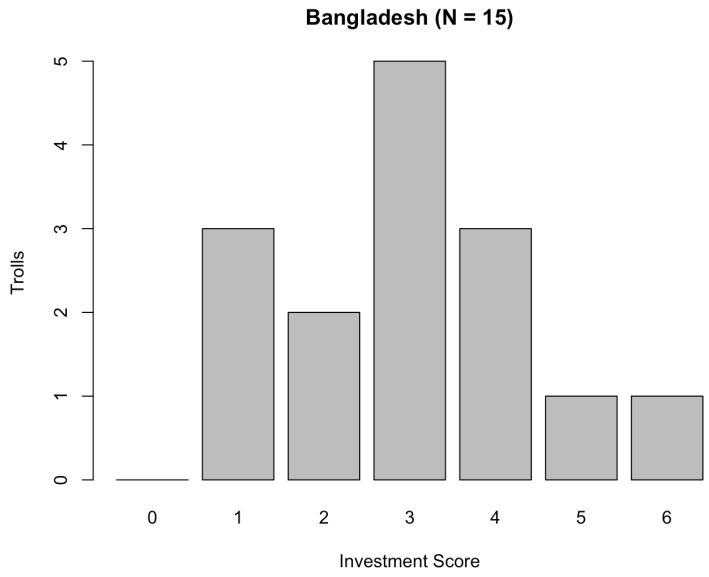

Figure 3.4: Bangladesh Investment Distribution

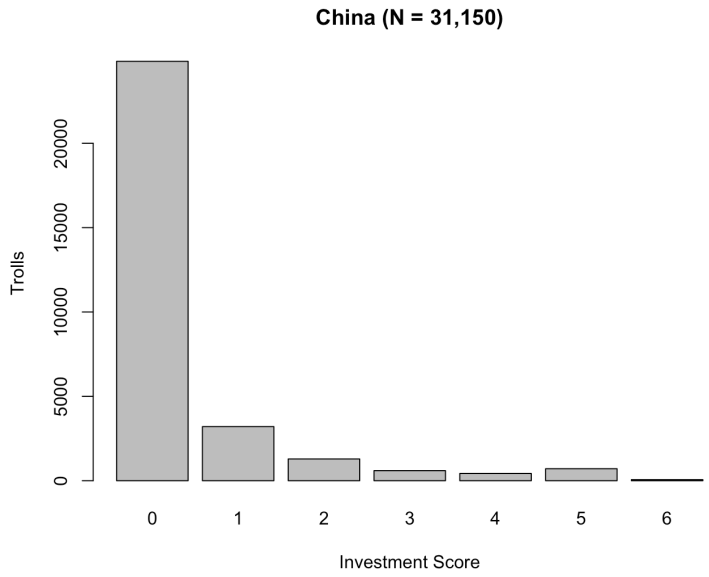

Figure 3.5: China Investment Distribution

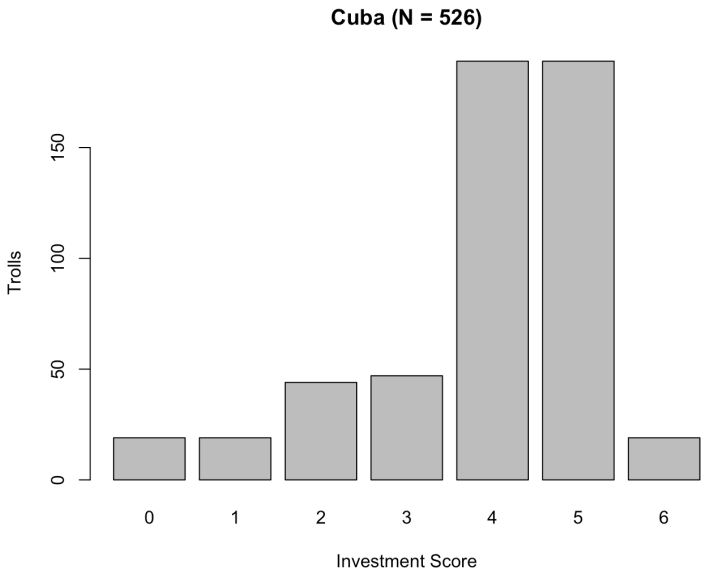

Figure 3.6: Cuba Investment Distribution

Figure 3.7: Ecuador Investment Distribution

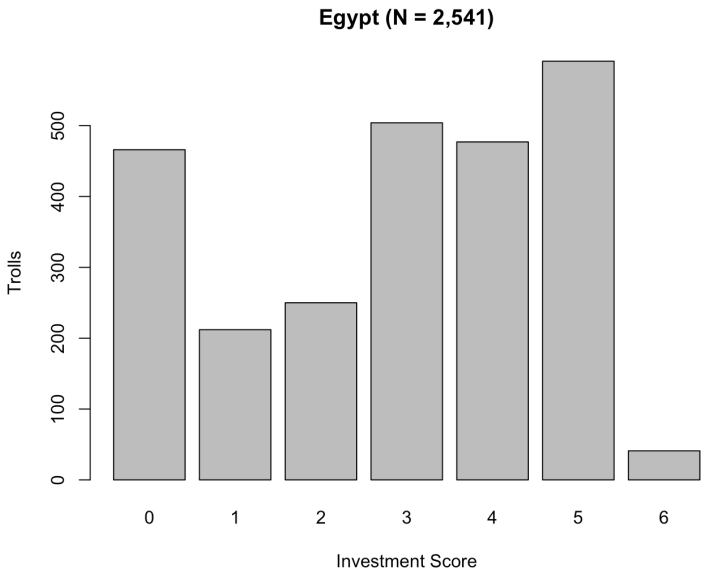

Figure 3.8: Egypt Investment Distribution

Figure 3.9: Honduras Investment Distribution

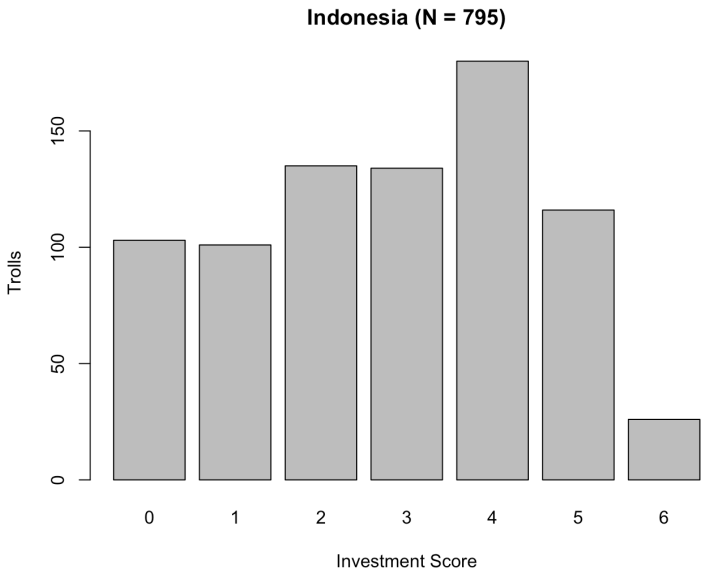

Figure 3.10: Indonesia Investment Distribution

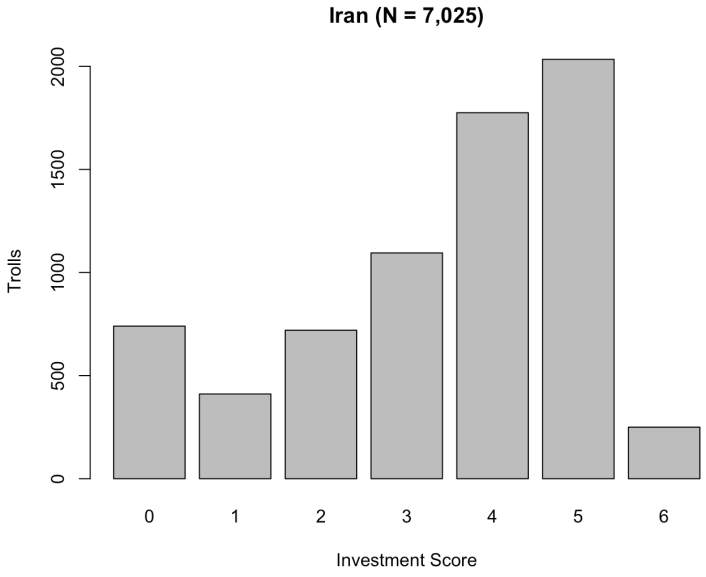

Figure 3.11: Iran Investment Distribution

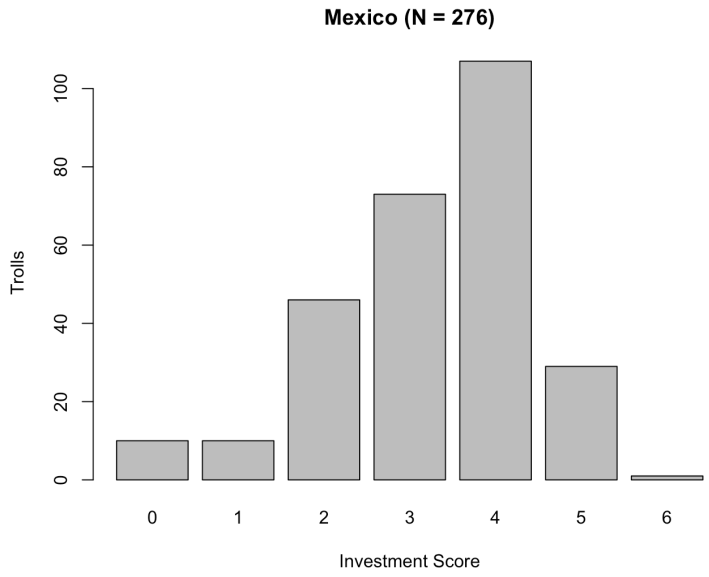

Figure 3.12: Mexico Investment Distribution

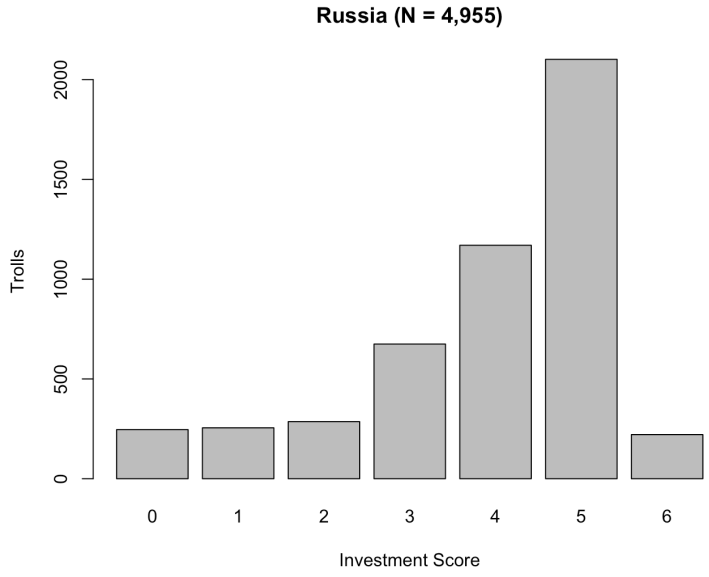

Figure 3.13: Russia Investment Distribution

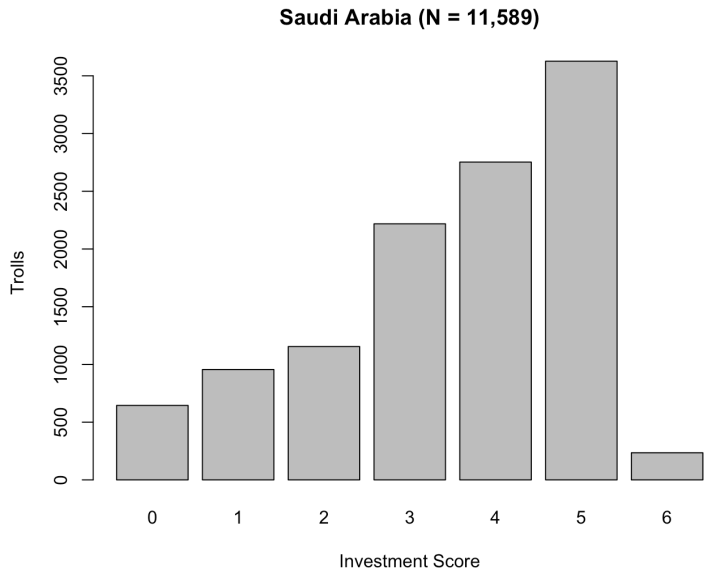

Figure 3.14: Saudi Arabia Investment Distribution

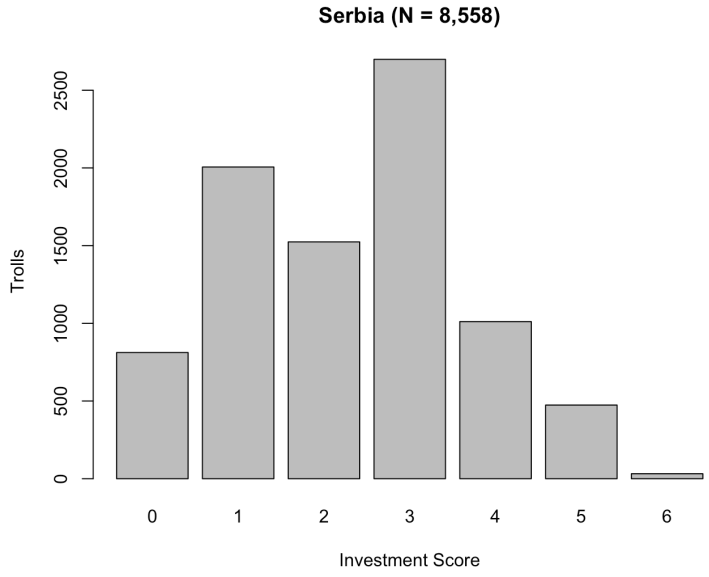

Figure 3.15: Serbia Investment Distribution

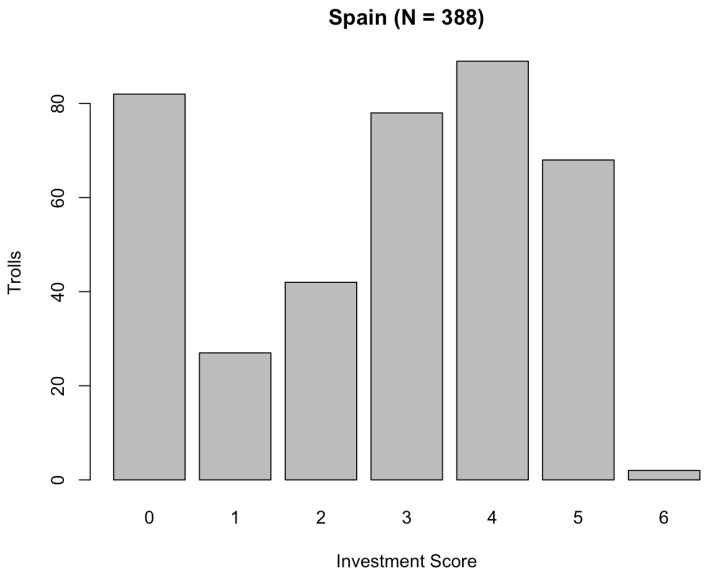

Figure 3.16: Spain Investment Distribution

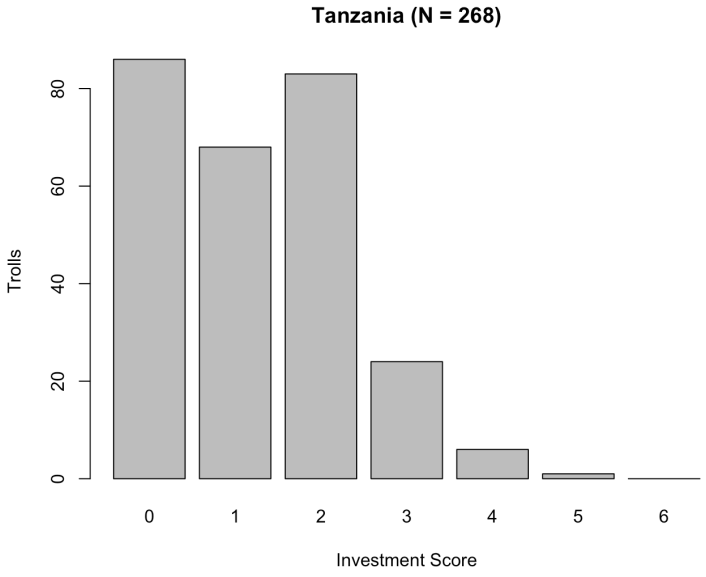

Figure 3.17: Tanzania Investment Distribution

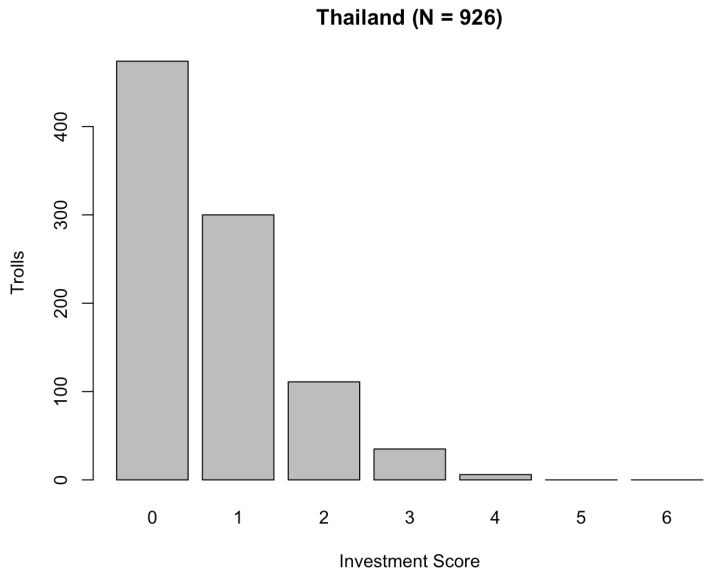

Figure 3.18: Thailand Investment Distribution

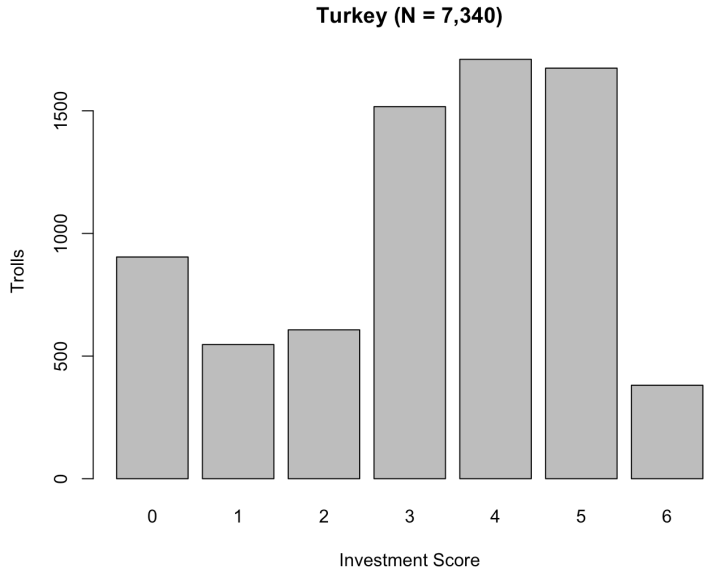

Figure 3.19: Turkey Investment Distribution

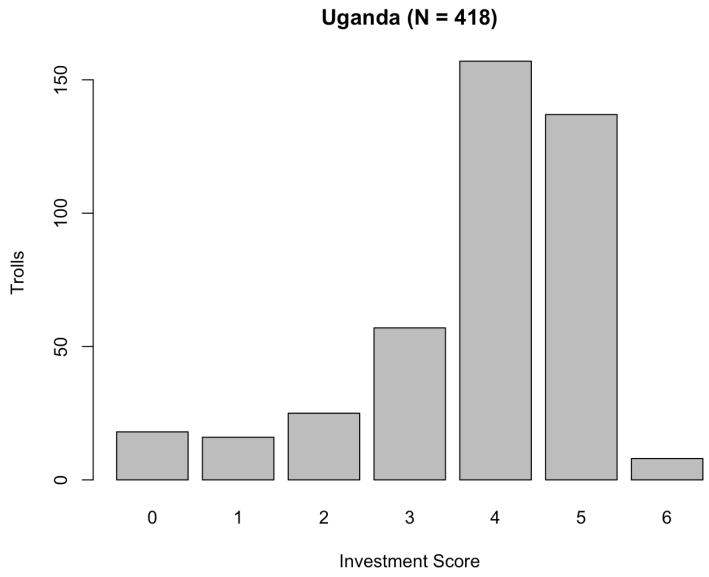

Figure 3.20: Uganda Investment Distribution

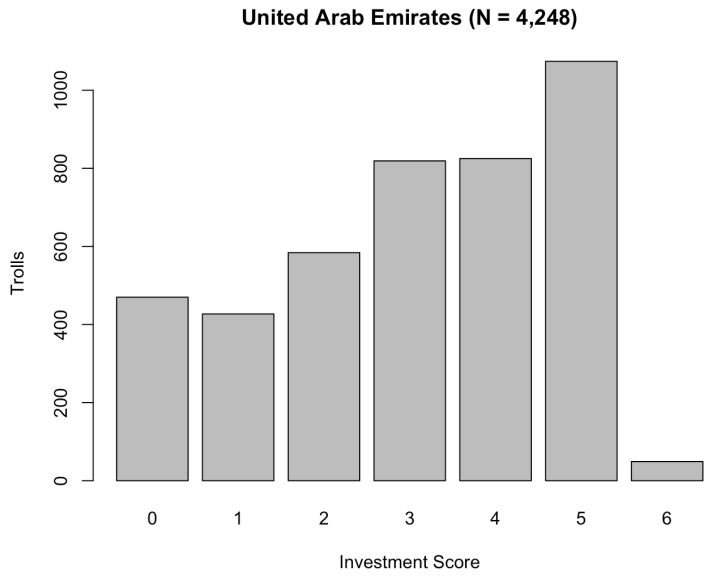

Figure 3.21: UAE Investment Distribution

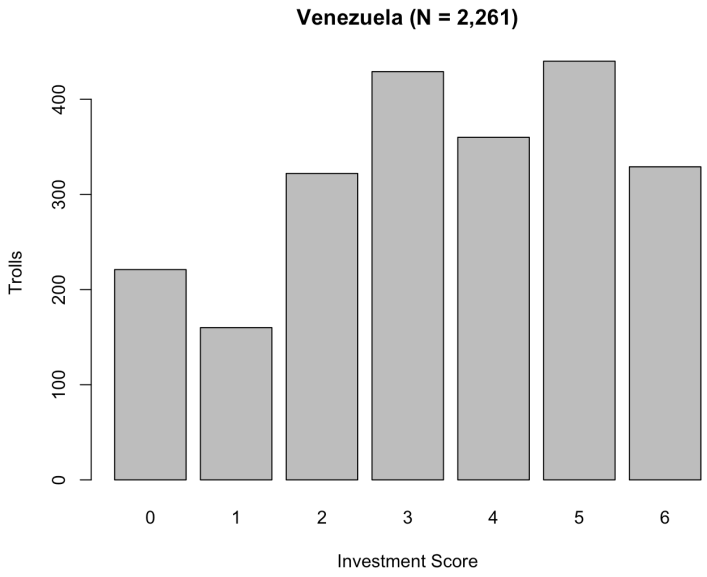

Figure 3.22: Venezuela Investment Distribution

Similarly, Figure 3.23 compares national investment patterns via a box-and-whiskers plot where the black lines depict the median investment score; the boxes depict the interquartile range (IQR; i.e., the first quartile to the third quartile or where 50% of the data resides); the whiskers depict a range extending from 1Q - (1.5*IQR) to 3Q + (1.5*IQR); and the circles depict the existence of at least one troll with a score beyond the +/- 1.5*IQR whisker range.

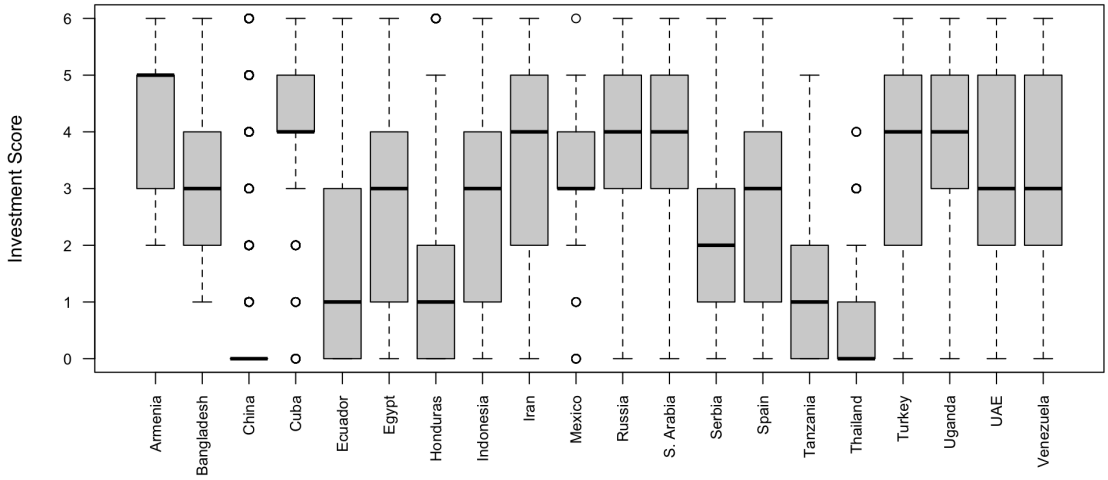

Figure 3.23: Box-and-Whiskers of Troll Investment by Country

As for the statistical results of the OLS model, the coefficient values for the 19 dummy variables can be found in Table 3.3 with China as the reference category (i.e., *constant*) and the statistical significance threshold being 95% (i.e., p-values < 0.05). While it is important to note that the null hypothesis is rejected due to all 19 country variables possessing a statistically significant difference than the reference category (i.e., China), it is arguably more important to draw attention to the reality that the preponderance of countries are statistically different from one another – as seen in the OLS coefficient plot (Figure 3.24).

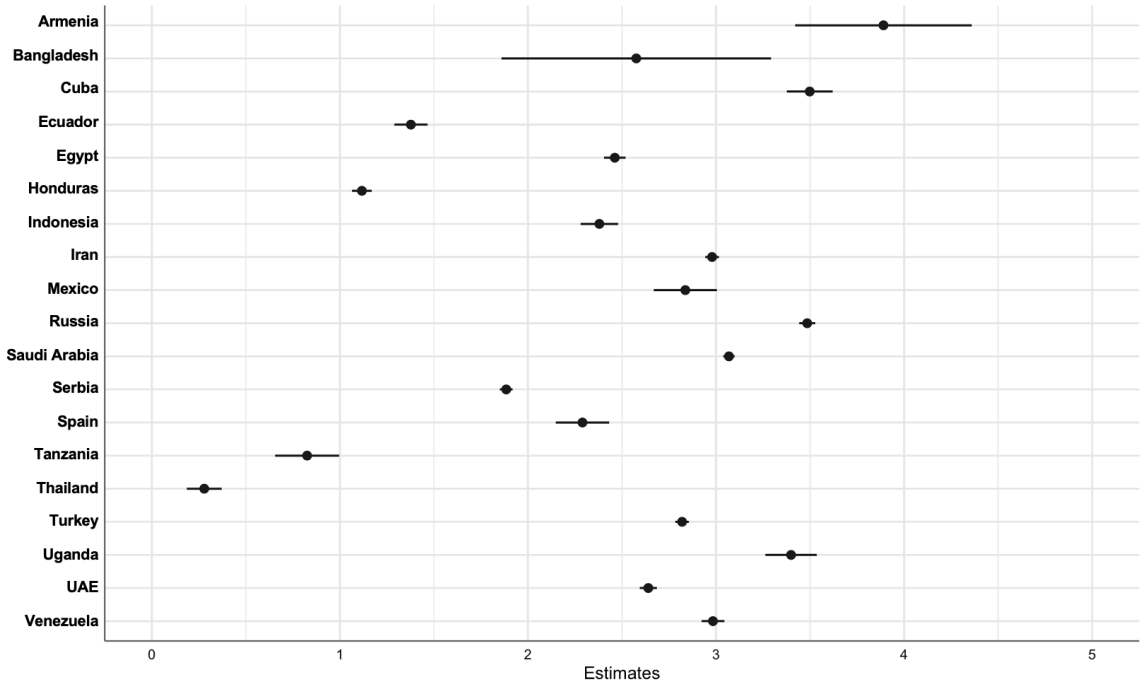

Figure 3.24: OLS Model Coefficient Plot for Investment~Country

	Investment Score		
Constant	0.424***	**Russia**	3.485***
	(0.008)		(0.022)
Armenia	3.890***	**S. Arabia**	3.069***
	(0.239)		(0.015)
Bangladesh	2.576***	**Serbia**	1.885***
	(0.366)		(0.017)
Cuba	3.498***	**Spain**	2.290***
	(0.062)		(0.072)
Ecuador	1.378***	**Tanzania**	0.826***
	(0.045)		(0.087)
Egypt	2.462***	**Thailand**	0.279***
	(0.029)		(0.047)
Honduras	1.117***	**Turkey**	2.820***
	(0.027)		(0.018)
Indonesia	2.380***	**Uganda**	3.399***
	(0.051)		(0.070)
Iran	2.979***	**UAE**	2.640***
	(0.019)		(0.023)
Mexico	2.837***	**Venezuela**	2.984***
	(0.086)		(0.031)
Observations	87,437		
R^2	0.484		
F Statistic	4,320.301***		
	(df = 19; 87,417)		

Note: * p < 0.05, ** p < 0.01, and *** p < 0.01. The constant is China as the reference category and estimates the mean *Investment Score* (i.e., the dependent variable on a 0-6 scale) for Chinese trolls The remaining coefficients estimate how much more or less than China the corresponding state-sponsor is invested in their trolls, all else being equal. Parenthetical numbers are the coefficient standard errors. The F-statistic of the regression tests the null hypothesis that all of the coefficients are equal to zero.

Table 3.3: Mean State-Sponsor Troll Investment Scores

Finding #2: Nuanced Relationship Between Investment & Goal

After easily identifying that significant differences exist between the average investment scores of the various state-sponsors, attempting to correlate investment scores with the goals of a given information operation proved more challenging. Taking the fused Twitter-ESOC data as-is and parsing it through the second OLS regression yields the outputs listed in Table 3.4, Panel 1. With the intercept being those accounts that could not be coded as either *promotional* or *demotional* due to data limitations, this regression would trigger a rejection of the null hypothesis by suggesting that on average and all else equal *demotional* trolls have higher mean investment than their *promotional* counterparts.

However, there is a potential dispute with placing the data as-is into the regression, namely the 10,899 Chinese trolls that my colleagues and I argue ESOC coded incorrectly as *promotional* over *demotional* based on the monolithic framing tendencies surrounding the problem of state-sponsored trolls (Linvill et. al 2024; see also Chapter 1). Recoding China as *demotional* prior to rerunning the same OLS regression yields the outputs in Table 3.4, Panel 2. Once again the constant is those accounts that could not be coded as either *promotional* or *demotional* due to data limitations, but this regression triggers a rejection of the null hypothesis by suggesting that on average and all else equal *promotional* trolls have higher mean investment than their *demotional* counterparts.

Although this second iteration of the model is arguably more defendable in terms of data accuracy, the volume of Chinese trolls admittedly dominates all other countries and thus p-hacking critics would not be wrong in pointing out that whichever operational

goal category China is in will by default have the lower mean investment score. To account for this possibility, the model is run a third time without any Chinese trolls (N=16,800) and the results are in Table 3.4, Panel 3. Statistically this third iteration of the model would also result in a rejection of the null hypothesis.

	Investment Score		
	(1)	(2)	(3)
Constant	1.788***	1.426***	3.769***
	(0.040)	(0.038)	(0.030)
Promotion	-0.080**	2.473***	0.077***
	(0.038)	(0.033)	(0.026)
Demotion	2.348***	0.139***	0.309***
	(0.033)	(0.036)	(0.025)
Observations	27,699	27,699	16,800
R^2	0.280	0.298	0.011
F Statistic	5,393.008***	5,877.929***	94.267***
	(df = 2; 27,696)	(df = 2; 27,696)	(df = 2; 16,797)

Note: * p < 0.05, ** p < 0.01, and *** p < 0.01. Panel 1 codes China as "Promote" per original ESOC dataset; Panel 2 re-codes China as "Demote;" Panel 3 omits China. The "constant" is the estimated *Investment Score* (i.e., the dependent variable on a 0-6 scale) for trolls unable to be coded as promote or demote. The remaining coefficients estimate how much more or less than the constant those types of trolls are, all else being equal. Since trolls can be both, promotion/demotion are not mutually exclusive. The F-statistic of the regression tests the null hypothesis that all of the coefficients are equal to zero.

Table 3.4: Mean Information Operation Goal Investment Scores

Given that all three iterations of the OLS model cause the null hypothesis to be rejected, albeit each for different reasons, we can still have confidence in saying there is in fact *a significant difference between the troll account investments for promotion information operations versus demotion information operations* – while simultaneously acknowledging that more work needs to be done in order to understand why that is true.

Discussion

Based upon these quantitative findings, the answer to RQ_2 (i.e., *How are State-Sponsored Trolls on Social Media Operationalized to Conduct Information Operations?*) proves to be the quintessential (and anti-climactic) academic refrain: It Depends!

Fact: Countries Invest Differently

This study provides considerable evidence that countries invest differently in their trolls, which directly challenges the existing tendencies of framing the problem of state-sponsored trolls on social media as monolithic and homogonous. However, merely demonstrating that the variations in overall investment offers little insight into why this seems to occur…but this is also why the RQ_2 was designed as an *exploratory* question.

In examining the patterns in the first OLS model, there are ostensibly high-investors (e.g., Armenia, Russia, Cuba), however there is general consensus about why countries would want artisanal accounts that present as real people on many cross-sectional facets – namely that they invest in order to wield surreptitious influence (e.g., Aristotelian *ethos*, *logos*, and *pathos*; the Cialdini Influence Principles). On the other hand, with the exception of the sustained indifference directed towards China's underwhelming accounts, there is little (if any) dialogue on why a state-sponsor might deliberately maintain low-invested trolls in their information warfare arsenals. In examining the data/model from this perspective, China is not the only country with its IQR resting on the 0 line but is joined by Ecuador, Honduras, Tanzania, and Thailand. While these four are statistically different from China they are not as substantively different from one another as they are from the high-investment countries.

After being quantitatively cued to look at these five more closely, the data simultaneously raises the question of what these five might have in common. It is not region, language, or culture – there are two in Asia, two in Latin America, and one in Africa. Nor is it foreign versus domestic information campaigns – China's is a foreign-directed campaign on the basis that the *great firewall of China* will not allow lay Chinese citizens to *tweet* and Twitter (n.d.) reports the other four as being domestic. However there are qualitative linkages that could potentially be gleaned about these trolls by associating them with the campaigns they were waging.

The Twitter Information Operation Archive and the Twitter Safety's official account (@TwitterSafety) collaborate to release brief public statements on why trolls were de-platformed and placed into the Information Operations Archive. From a thematic content analysis perspective, the following excerpts from Twitter's public statements reveal a possible trend:

- **China**: Twitter described this activity as "spammy;" using tactics that both "amplified CCP narratives" and sought to "artificially inflate impression metrics."
- **Ecuador**: Twitter stated that the "tactics most commonly used were hashtag manipulation and retweet spam."
- **Honduras**: Twitter attributed these troll accounts to a government staffer and described their purpose as "retweeting the President's account."
- **Tanzania**: Twitter stated that these troll accounts were "utilized to file bad faith reports" about political opponents with the Twitter terms of service managers.
- **Thailand**: Twitter indicated that these trolls engaged in coordinated behaviors "targeting prominent political opposition" and "amplifying pro-Royal Thai Army and pro-government" sentiments.

Ultimately, the word choice and descriptions of these campaigns appear to place an emphasis not on influencing people directly but rather on manipulating the Twitter algorithm and the information environment by saturating the platform with spam, flagging legitimate accounts as problematic in order to silence them/mitigate their reach, or placing a proverbial *thumb on the scale* to impact the way something does (or does not) trend via algorithmic manipulation. From this, it stands to reason that an account may not need high overall investment to appear convincing to people if in fact it is instead aimed primarily at the algorithm in such a way that a low-investment account generally punches at the same weight as any other account.

Relationship Between Investment & Goal?

Assuming the possibility that *promotional* goals often have more to do with targeting people's cognition directly and *demotional* goals often have more to do with targeting the algorithms to indirectly make certain sets of ideas less salient in the information environment, then the findings from the first OLS model may offer some insights into the results of the second OLS model.

Obviously, the lower-investing countries likely having demotional goals are antithetical to the first iteration of the Investment~Goal model where demotion was greater than promotion – but I have already addressed that this is a result of China being qualitatively coded incorrectly by ESOC. Lower-investing countries correlating with demotional goals line up well when China is re-coded to demotional – yet I have also acknowledged that this could be misconstrued by some as p-hacking given the sheer

volume of Chinese trolls. This leaves us with the third iteration of the Investment~Goal OLS model where China was omitted.

This last iteration revealed essentially no substantive difference in overall mean investment score between promotional and demotional goals, but that should not be misconstrued as saying that this implies the promotional trolls and demotional trolls are fundamentally the same. The aggregated nature of the *Investment Score* variable allows for 64 different permutations along the six individual 0/1 components. Moreover, since the majority of trolls do not do any of the individual *Investment Score* components (with the exception of 57.7% of trolls following 10 accounts, see Appendix A), it becomes self-evident that an aggregated score of 3 or 4 demonstrates that decisions are being made at the account operator level (now what exactly those decisions are is an entirely different matter). As a result, when both the promotional and demotional trolls average "4s" for their investment it is paramount to remember that not all 4s are created equal. In fact, when examining how the *Investment Score*'s individual components (Appendix B) relate to promotional and demotional goals a pattern emerges (see Table 3.5).

	Descript.	**Location**	**URL**	**Fol. 10**	**Fol. 50**	**Fol. 100**
(Constant)	0.655	0.458	0.098	0.977	0.851	0.730
Demotion	0.055	0.024	-0.015	0.006	0.089	0.150
Promotion	0.114	0.098	-0.013	-0.028	-0.047	-0.047
P-Value for Dif. Between Coefficients	<.01	<.01	.6168	<.01	<.01	<.01

Note: Columns are separate OLS regressions where the investment score components are treated as the dependent variable. These models calculate the probability of a troll being *Promotional* or *Demotional* based on having certain investment components (green annotates more likely/red annotates less likely). P-Values are the linear hypothesis tests that the coefficients are statistically different from each other.

Table 3.5: Investment Components Relative to Goal

Description and *Location* are statistically significant predictors of having *promotional* goals while *Follow10*, *Follow50*, and *Follow100* are statistically significant predictors of having *demotional* goals. These statistically significant predictors align with what one would intuitively expect if *promotional* tends to correlate with targeting people and *demotional* tends to correlate with targeting algorithms. *Description* and *Location* are components cognitively processed by social media users when deciding to engage with the profile while following a significant number of accounts increases the potential for broader algorithmic impact.

China as the Benchmark for Economy of Force

The principal takeaway here is that the notion of *Investment* needs to be taken seriously and understood as a series of choices on the part of the state-sponsor; choices that do not take place in a vacuum but rather that are made deliberately in a way that is economically efficient (i.e., accomplishing the goal with the least amount of unnecessary expenditures and fringe costs). Just because an account is a "6" does not mean it is highly effective and vice-versa, just because an account is a "0" does not mean it is ineffective. This vantage point offers far more insights to the Chinese troll apparatus, revealing instead the possibility that they are not corporately inept but rather one of the most economically efficient troll armies ever created – in terms of what they accomplish.

This possibility is most effectively communicated through a tangible example rather than an abstract theorization – and for that I turn to observational data from two studies I co-authored over the course of my doctoral journey. First, in the fall of 2021 we (Linvill et. al. 2021) observed Chinese troll accounts flooding the hashtags associated

with social justice commentary denouncing the CCP's human rights violations against the Uyghurs in Xinjiang province (e.g., #Xinjiang, #XinjiangCotton). At face value, the flooding had the look and feel of quintessential propaganda trying to convince the world that everything was copesetic in China's western province (i.e., was promoting the idea that China is an innocent victim of geopolitical libel; see Figure 3.25).

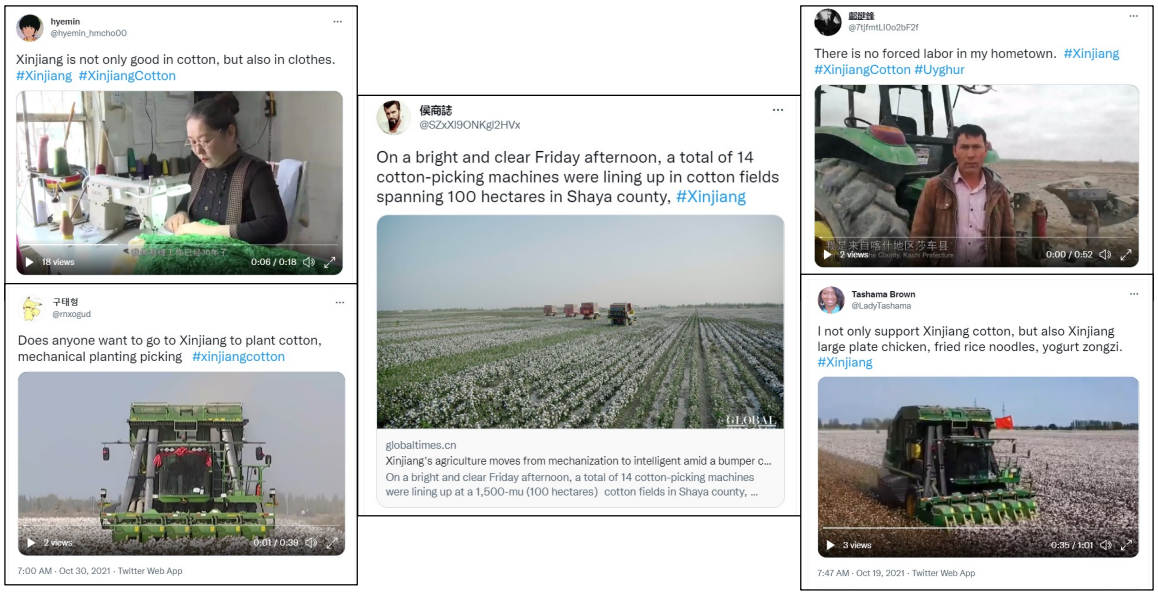

Figure 3.25: Example Pro-Xinjiang Tweets

The troll accounts pushing this content had many of the tell-tale markers of inauthenticity (e.g., amplifying one another's content verbatim [Figure 3.26], tweeting at all hours of the day and night). Upon closer examination, many were also otherwise unconvincing in terms of appearing realistic, amassing no followers or engagement (Figure 3.27) in ways similar to the Stanford Internet Observatory's critiques at the start of this chapter. Relating this to the research methodology above, the accounts in Figure

3.27 for “Gary Horton” and “Hitler 2” would both be scored as a “0” in the *Investment Score* dependent variable.[11]

Figure 3.26: In-Network Verbatim Retweeting

[11] In addition to having none of the six components contained in the *Investment Score*, neither of these accounts have banner pictures and both make use of low-effort, clipart-style profile pictures – further demonstrating that these types of Chinese accounts are essentially operating with default settings.

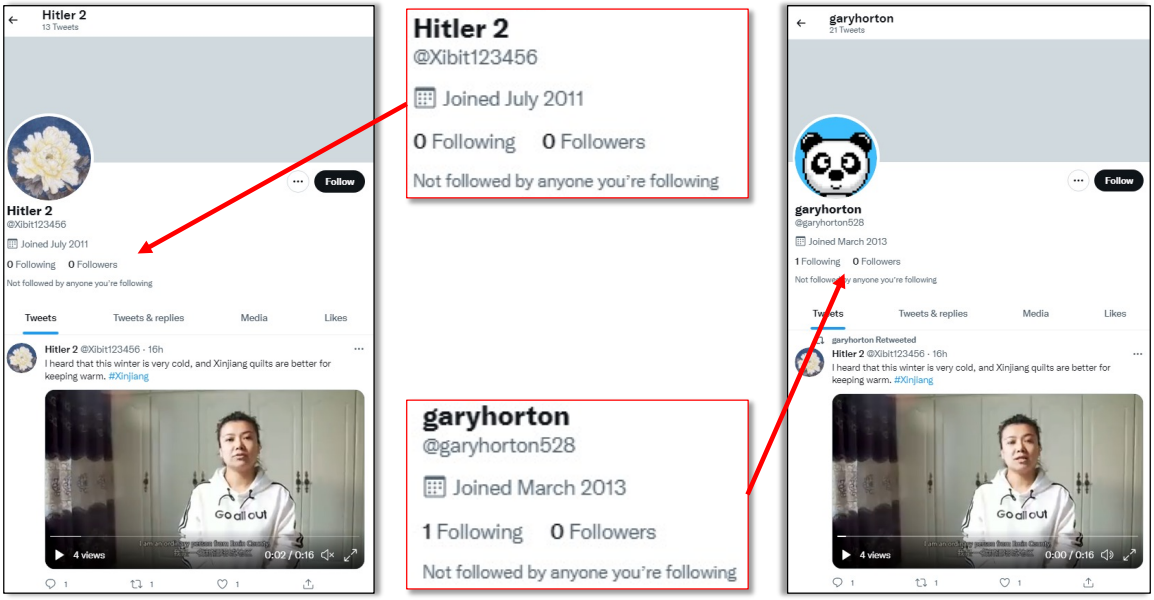

Figure 3.27: Two Example Chinese Troll Personas

But upon further probing, we realized that it did not necessarily matter that the trolls did not appear realistic, that the propaganda was not believable, or that the accounts were not receiving any engagement. To the contrary, the more we studied these accounts and the information operation(s) that they were waging – the more we realized that the interesting element was not what we were seeing but rather what we were not seeing. The more we explored the information operation(s) the more we realized that the original social justice commentary that started the hashtags in the first place was nowhere to be found. Worse still, even when we attempted to manually look for it via Twitter's search feature all we encountered was Chinese trolls and content because search algorithms in their simplest form yield that which there is the most of (Figure 3.28).

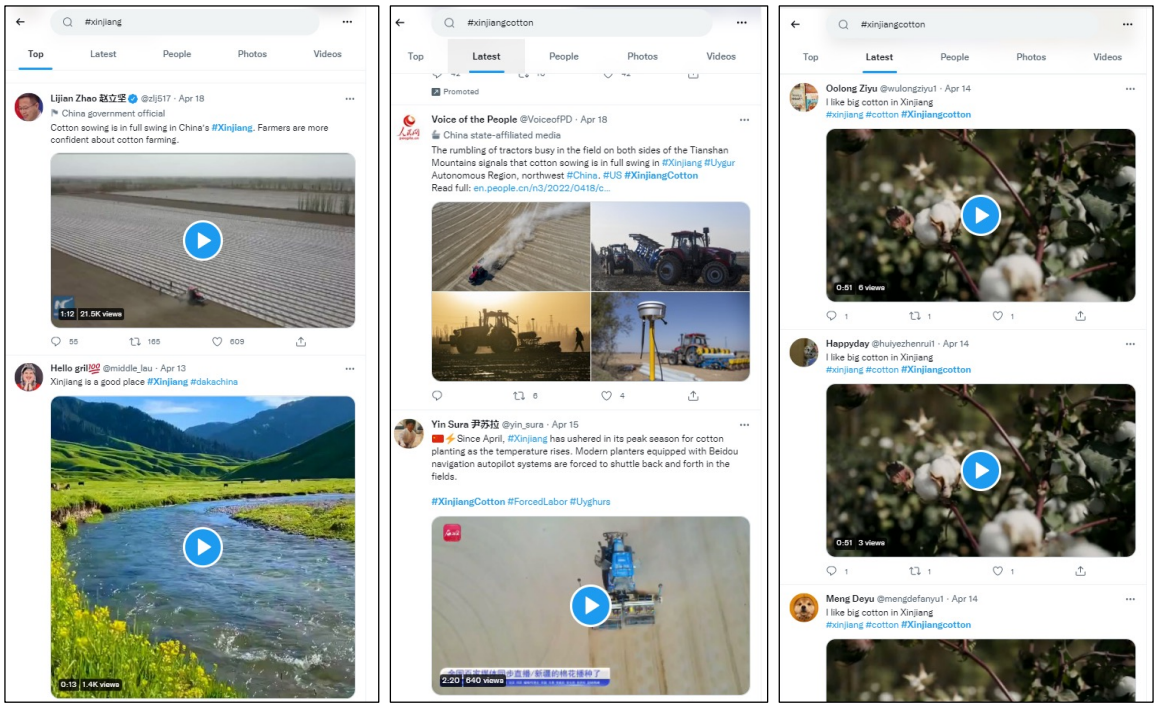

Figure 3.28: Manual Search for #Xinjiang & #XinjiangCotton

Worse still, China is perfecting its censorship via troll tradecraft and Xinjiang is not the only information operation where the CCP has successfully turned off conversations on social media that they identified as contrary to their national interest. In a different study conducted a year later, we (Warren et. al. 2023) found Chinese trolls silencing assertions that COVID-19 originated from a lab in Wuhan – and by extension silencing those who spread these claims, such as University of Pennsylvania virologist (and Chinese expatriate) Li-meng Yan. On Twitter, they created an entire army of trolls spoofing the doctor's actual account to make it impossible for social media users to find the real Li-Meng Yan (Figure 3.29). On Pinterest, the content attached to #limengyan was so graphically grotesque and disturbing (while also being generally unsophisticated) that the platform's trust and safety community guidelines upended the ability to search

for it (Figure 3.30)…which in the end means that Pinterest essentially did China's job for them because all they wanted in the first place was for Li-Meng Yan to be out of the public discourse. Simply put, *promotion* and *demotion* are not the same thing and thus should not be compared to one another as if they are not phenomenologically dissimilar.

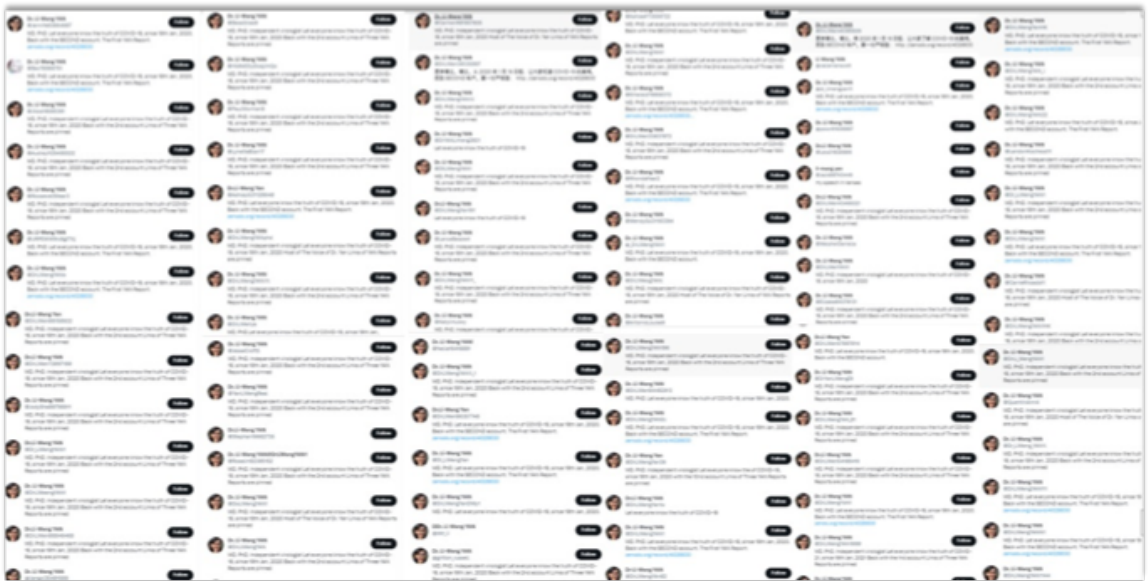

Figure 3.29: Troll Accounts Spoofing Li-Meng Yan
(Not Exhaustive)

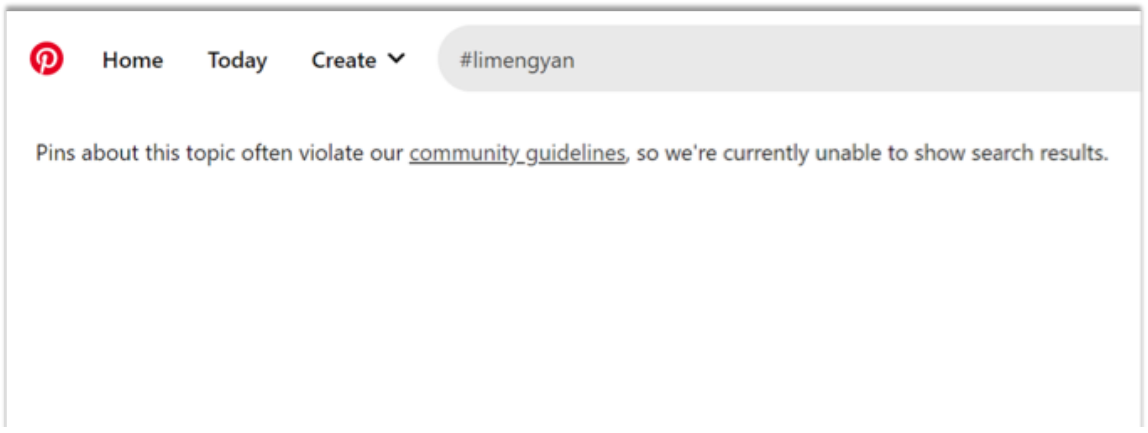

Figure 3.30: Chinese Trolls Weaponized Pinterest's
Community Guidelines

Way Forward

In the end, just as it is considered colloquially unwise to judge a book by its cover, so too should we reject the superficial tendency to judge a troll based on its appearance. Sometimes what is not seen is just as important (if not more so) than what is seen and as a result researchers and practitioners need to expand beyond case studies that predominantly focus on idea promotion and generating engagement with real users. Put more succinctly: The need for a taxonomy of state-sponsored trolls is dire as more and more geopolitical actors turn to information warfare as an inexpensive means of conducting the affairs of state at home and abroad.

For now, the primary takeaway from this chapter for the remainder of the dissertation is that the notion of account *investment* must be taken seriously! This takeaway is consistent with and complimentary to our (Linvill et. al. 2024) findings regarding trolls making location claims in the very first study using the entirety of the Twitter Information Operations Archive. While it is true that most state-sponsored trolls claim to be from *nowhere* (Appendix A), the trolls that do make location claims seem to correspond with troll operators weighing the *costs* (e.g., higher-risk of getting caught by Twitter when claimed location does not match the IP-address, requires more sociological expertise to keep up the persona's appearances) against the perceived *benefits* (e.g., increased potential for persuasion within a target audience, is supportive of a campaign's goal[s]). If such a cost-benefit approach to troll account curation is true, there is no reason to think that such an approach would not be equally applicable to the other components of the *investment score* dependent variable.

Bottom Line: The creation of a so-called troll farm is not a one-size-fits-all endeavor but rather has varying levels of start-up costs and overhead that the various state-sponsors must take into consideration relative to what goal(s) they seek to accomplish. *Promotion* versus *Demotion* is not a matter of right or wrong; they are two different goals that can each be done in different ways and thus it logically follows that trolls doing one may not look like trolls doing the other. This reality is particularly important when attempting to craft policy solutions to the problem of state-sponsored trolls on social media – which will be the topic of the next chapter.

CHAPTER FOUR

FRAMINGS VERSUS REALITY: THE GOVERNMENTAL RESPONSE TO TROLLS

In Act II of William Shakespeare's *The Tragedy of Julius Caesar* (first staged in 1599), Brutus receives numerous handwritten letters entreating him to partake in the conspiracy to depose the emperor. Designed to make Brutus believe that the Roman people preferred Pompey over Caesar, these letters all shared a major commonality: They are all fabrications that originate from Cassius. As one literary commentator offers, Cassius "undoubtedly dictates them to others so that the handwriting will seem different in each" (Delaney n.d.). The similarities to trolls are obvious, even to the point that one could imagine a modern adaptation of *Julius Caesar* that replaces letters with tweets. Consequently, the notion of deception through mediated social communications and manipulating behavior through fabricating the perception of grassroots support (known today as *astroturfing*) predates the 2016 U.S. presidential election by more than four centuries (perhaps even two millennia, if there is any historical legitimacy to the Shakespearean account) – in no small part because influencing people through informational means is as old as time itself.

While it is true that the sociological phenomenon exploited by state-sponsored trolls on social media is not unique to the current era, one must also acknowledge that the internet has lowered the proverbial cost of entry, simplified the targeting process, and exponentially "amplified the dissemination" of information reminiscent to the invention of the Guttenberg printing press, radio, and television (Posetti & Matthews 2018, 1;

Ellick, Westbrook, & Kessel 2018). It is precisely this juxtaposition between something being regarded as primarily old or new that must be balanced throughout when discussing the problem of state-sponsored trolls on social media.

The current chapter explores the governmental response to the problem of trolls through the question:

- RQ_3 = *What are Policies the U.S. Federal Government Uses to Solve the Problem of State-Sponsored Trolls on Social Media?*

This portion of the dissertation is exploratory and does not produce exhaustive findings because there is no lead federal agency and by extension there is no singular dataset through which to answer RQ_3, hence why the structure of this chapter will be different than its predecessors.

Governmental Response to Focusing Events

"There is nothing new under the sun." These words from Jewish antiquity are no less meaningful now than when Solomon (n.d.) first uttered them. Yet there is at times a temptation in the defense and foreign policy realms to treat the advent of major shifts in the geopolitical/technological landscape as *revolutionary* rather than *evolutionary* (i.e., treat it as something "new under the sun" vis-à-vis a contextual adaptation of an already existing phenomenon). Nuclear weapons and the resulting concepts of deterrence and mutually assured destruction undeniably changed how wars were fought, but their novelty overshadowed both the reality that the intellectual roots of *Realism* extend back to Thucydides (n.d.) and Clausewitz's (1873) assertion that "violence arms itself with the inventions of Art and Science in order to contend against violence" (i.e., war's nature does not change, regardless of new weaponry). Similarly, Beijing's advancements in

ballistic missile and air defense technologies have caused Western defense practitioners to invent new lexiconic terminology – namely, *anti-access*/*area denial* (A2/AD) – even though China's actions are ultimately a combined arms defense of the homeland that capitalizes on geographic/topographic advantages in ways evocative of the 300 Spartans at Thermopylae (Warren 2020a).

In some of my previous work exploring great power competition policy (Warren 2024), I argued that major national security and/or foreign policy focusing events normally trigger massive change in the policy (Table 4.1).

Focusing Event	Massive Change in Policy
Japan attacks Pearl Harbor	U.S. reverses policy of isolationism and becomes fully involved in the European and Pacific theaters of WWII.
U.S. emerges from WWII as the leading global superpower	National Security Act of 1947 lays the foundation for the National Security Council (NSC), Central Intelligence Agency (CIA), Department of Defense (DOD), & other military/foreign policy reorganizations.
Soviets launch Sputnik-1 as first artificial Earth satellite	The quest to put a man on the moon as national policy, the creation of the National Aeronautics & Space Administration (NASA), the National Defense Education Act of 1958.
U.S. military campaigns fail in Iran (1980) & Granada (1983)	Goldwater-Nichols Act of 1986 mandates joint operations as the standard for U.S. military preparedness.
Soviet Union collapses	Intelligence community (IC) reshuffled; foreign policy and joint force enterprises reduced in scope/scale.
Attacks of September 11, 2001	Department of Homeland Security (DHS) created, PATRIOT Act enacted, Global War on Terror initiated.

Table 4.1: Focusing Events & Policy Change[12]

Assuming there is validity to this argument, then one would reasonably expect the problem of state-sponsored trolls on social media to trigger widespread policy change and unity of effort within the whole of government. However, as we take an introspective

[12] Table reproduced by permission from: "On Competition: A Continuation of Policy by Misunderstood Means" (Warren 2024).

assessment of current conditions, the disjointed policy response since the 2016 elections raise the possibility that this focusing event was an anomalous theoretical outlier.

Adaptive Policy Solutions for State-Sponsored Trolls

As noted in Chapter 1, the warning signs that social media would become a digital battleground for 21st Century geopolitics were there prior to the 2016 elections – society and policymakers alike failed to heed them. This placed the U.S. on a reactive footing where the initial policy solutions were to co-opt or modify existing processes or entities.

The Mueller Probe

The first policy decision in response to the problem of state-sponsored trolls on social media was the appointment of Robert Mueller to investigate Russian interference. Mueller's (2019) report detailed not only how Prigozhin's IRA launched a vast array of troll activity across Facebook, Twitter, YouTube, Instagram, and Tumblr but also how these accounts generated engagement and amassed followers by pretending to be U.S. persons associated with various in-group/out-group dynamics (e.g., Black social justice, Tea Party, Muslims, LGBTQ). It described the use of troll accounts that had been automated and networked (a.k.a. *bot-nets*) to amplify the purposively curated accounts and raise the salience of ideas by giving the appearance of grassroots movements – even to the point of organizing real-world protests. In addition to the IRA Mueller also outlined how the GRU (i.e., Russian military intelligence) used trolls on social media as a delivery conduit for its hack-and-release cyberspace operations directed at the Democratic National Committee.

The beginning of the Mueller Report was essentially the first in-depth governmental report on the problem of state-sponsored trolls on social media. But as the report continued, it became increasingly less about the actual problem of trolls and more about the focusing event as it pertained to the implications for the Trump campaign specifically. Consequently the interpretation of the Mueller Report and the application of its findings were arguably skewed by the Trump-Russia collusion narrative, an accusation of questionable origin that was "amplified in a feedback loop by most US media" (BBC 2021). Although the final report did not take a position on *collusion* (finding instead no evidence of a *conspiratorial* agreement), the $32M investigation resulted in indictments for Russian companies and some Trump affiliates while simultaneously producing no sustainable solutions to the problem of trolls and fostering further partisan divides across the political spectrum (Breuninger 2019; Polantz 2019). Agnostic of the politics, the emphasis here is the policy mechanism was entirely focused on the Russians and elections while also not producing solutions that would make society sustainably safer.

Select Subcommittee on Intelligence Investigation

In parallel with Robert Mueller's investigation, the Select Subcommittee on Intelligence (SSCI) gathered information regarding all Russian election-related activities in 2016 and would publish a series of reports beginning in 2019. After Volume 1 examined offensive cyber activities directed at physical election infrastructure (e.g., voting machine security), Volume 2 turned attention to social media exploits and reached similar conclusions as Mueller regarding IRA tactics (e.g., impersonate real people, spread narratives, create division based on in-group/out-group) – noting also that

Prigozhin's managerial staff "made efforts to monitor and track the impact of its online efforts, through measurables such as comments, likes, reposts, changes in audience size, and other metrics" (27). The social media troll activity was about promoting an idea, namely "covert support of Russia's favored candidate" (3).

Yet once again as the investigation and response became more about the focusing event, so too did the governmental response and proposed solutions. Volume 3 cites a "heavily politicized environment" (2) as the foremost constraint upon taking action against the problem and thus the more immediate response to the focusing event was primarily imposing costs via geopolitical governmental responses (e.g., diplomatic warnings, cybersecurity measures, defending voting infrastructure, expelling Russian diplomats). Since the SSCI could not enact policy unilaterally, it made superficial recommendations for social media companies to be more transparent and for the IC to collect, monitor, and share data about new information warfare endeavors from hostile nations (i.e., prepare for the next information operation). It also asserted the problem of trolls and information operations was not necessarily the government's to solve but rather it would "ultimately need to be tackled by an informed and discerning population of citizens who are both alert to the threat and armed with the critical thinking skills necessary to protect against malicious influence" (Volume 2, 81). Aside from the utopianly untenable nature of a population of this kind, such an end state also hinges on the implicit assumption of idea promotion seeking to maliciously induce a behavioral response on the basis that critical thinking skills will not enable the proving of the negative to recognize idea demotion or censorship.

Department of State Global Engagement Center as "Lead" Agency

The State Department's Global Engagement Center (GEC) is technically designated as the lead…but in practice it is not that straight forward because the U.S. Department of State does not have the requisite domain awareness to detect malign activities or authority to command and control the interagency against the problem of state-sponsored trolls. These limitations are partly due to the GEC not being originally intended to address this particular issue.

The GEC's creation was initiated by President Obama on March 4, 2016 by Executive Order 13721 (i.e., before the focusing event). At that time, the GEC's mandate was a counterterrorism mission to "lead the coordination, integration, and synchronization of Government-wide communications activities directed at foreign audiences abroad in order to counter the messaging and diminish the influence of international terrorist organizations" (US Dept of State 2016). As indications of Russian interference emerged, Congress used the National Defense Authorization Act (NDAA) for FY17 to reactively expand the GEC's role to: "Lead, synchronize, and coordinate efforts of the federal government to recognize, understand, expose, and counter foreign state and non-state propaganda and disinformation efforts aimed at undermining U.S. national security interests" (Sec. 1287).

In the language of both the Executive Order and the NDAA, the focus is external to the United States. The GEC is directed to enable national security objectives by engaging in counter-messaging against adversaries and by promoting the truth abroad to reassure Allies and Partners (i.e., an inherently diplomatic mission). Yet in the 2019

NDAA (Sec. 1284), the fledgling GEC's purpose and functions were again modified. Table 4.2 contains the original GEC purpose/functions (left column) and the language in the FY19 NDAA that retroactively amended the language of the FY17 NDAA.

	NDAA (FY17) – Section 1287	NDAA (FY19) – Section 1284
Purpose	The purpose of the Center shall be to lead, synchronize, and coordinate efforts of the Federal Government to recognize, understand, expose, and counter foreign state and non-state propaganda and disinformation efforts aimed at undermining United States national security interests.	The purpose of the Center shall be to direct, lead, synchronize, integrate, and coordinate efforts of the Federal Government to recognize, understand, expose, and counter foreign state and foreign non-state propaganda and disinformation efforts aimed at undermining or influencing the policies, security, or stability of the United States and United States allies and partner nations.
GEC Function	(1) Integrate interagency and international efforts to track and evaluate counterfactual narratives abroad that threaten the national security interests of the United States and United States allies and partner nations.	(1) Direct, lead, synchronize, integrate, and coordinate interagency and international efforts to track and evaluate counterfactual narratives abroad that threaten the policies, security, or stability of the United States and United States allies and partner nations.
	(2) Analyze relevant information, data, analysis, and analytics from United States Government agencies, United States allies and partner nations, think tanks, academic institutions, civil society groups, and other nongovernmental organizations.	[No Change]
	(3) As needed, support the development and dissemination of fact-based narratives and analysis to counter propaganda and disinformation directed at the United States and United States allies and partner nations.	[No Change]
	(4) Identify current and emerging trends in foreign propaganda and disinformation in order to coordinate and shape the development of tactics, techniques, and procedures to expose and refute foreign misinformation and disinformation and proactively promote fact-based narratives and policies to audiences outside the United States.	(4) Identify current and emerging trends in foreign propaganda and disinformation in order to coordinate and shape the development of tactics, techniques, and procedures to expose and refute foreign propaganda and disinformation, and pro-actively support the promotion of credible, fact-based narratives and policies to audiences outside the United States.
	(5) Facilitate the use of a wide range of technologies and techniques by sharing expertise among Federal departments and agencies, seeking expertise from external sources, and implementing best practices.	[No Change]

Table 4.2: Original & Amended NDAA Language

	NDAA (FY17) – Section 1287	NDNAA (FY19) – Section 1284
GEC Function (Continued)	(6) Identify gaps in United States capabilities in areas relevant to the purpose of the Center and recommend necessary enhancements or changes.	[Redesignate as Paragraph (7)]
	(7) Identify the countries and populations most susceptible to propaganda and disinformation based on information provided by appropriate interagency entities.	Redesignate as Paragraph (8), Change to: "(8) Use information from appropriate interagency entities to identify the countries, geographic areas, and populations most susceptible to propaganda and disinformation, as well as the countries, geographic areas, and populations in which such propaganda and disinformation is likely to cause the most harm."
	(8) Administer the information access fund established pursuant to subsection (f).	[Redesignate as Paragraph (9)]
	(9) Coordinate with United States allies and partner nations in order to amplify the Center's efforts and avoid duplication.	[Redesignate as Paragraph (10)]
	(10) Maintain, collect, use, and disseminate records (as such term is defined in section 552a(a)(4) of title 5, United States Code) for research and data analysis of foreign state and non-state propaganda and disinformation efforts and communications related to public diplomacy efforts intended for foreign audiences. Such research and data analysis shall be reasonably tailored to meet the purposes of this paragraph and shall be carried out with due regard for privacy and civil liberties guidance and oversight.	[Redesignate as Paragraph (11)]
		[New Function] (6) Measure and evaluate the activities of the Center, including the outcomes of such activities, and implement mechanisms to ensure that the activities of the Center are updated to reflect the results of such measurement and evaluation.

Table 4.2 (Continued): Original & Amended NDAA Language

While the addition of "direct" and "integrate" to the purpose only further emphasizes that Congressional intent is for the GEC to be the colloquial quarterback in this policy arena (more on that later), the portion meriting discussion is the shift from "propaganda and disinformation efforts aimed at undermining United States national security interests" (FY17) to "propaganda and disinformation efforts aimed at undermining or influencing

the *policies*, security, or *stability* of the United States and United States allies and partner nations" (FY19 [emphasis added]). This original language is outwardly focused and perfectly aligned with the U.S. Department of State's diplomatic mission; the amended language introduces the potential for a conflict of interest by tasking the career civil servants in the foreign service to have involvement in domestic politics and processes.

After all, what constitutes a *threat to policies* when policies in a democratic society are byproducts of elections and constitutional procedures? Is it any threat to the current policy (i.e., the status quo) or a threat to the policy that the GEC's career civil servants deem the preferred policy position?[13] The Trump Administration differed from the Biden and Obama administrations regarding the Iranian Nuclear Deal – so when Iranian trolls sought to influence the reinstatement of said agreement, are they a "threat to policy" or not? This ambiguity becomes critically important in light of functions (1) and (4) where the GEC is meant to lead the identification and refutation of false narratives.

For now, the primary takeaway with regards to the GEC and the problem of state-sponsored trolls on social media is the organizational mandate from Congress places considerable emphasis on actively refuting the ideas promoted by adversaries and allows for potential GEC involvement in what is otherwise political processes. Moreover, the language of both of these policy elements seems intuitively connected to the 2016 election focusing event and is consistent with the framing trends identified in Chapter 2.

[13] This question is not raised out of conspiratorial concerns of a so-called *deep-state* but rather as a practical acknowledgement that the "long scholarly trail" of political science and public administration literature indicates the executive branch is more accurately understood as a plural "they" rather than a unilateral "it" (Rudalevige 2021, 18; see also Rao 2011; Allison & Zelikow 1999; Bose & Rudalevige 2020; Potter 2019). A tangible example of bottom-up policymaking in this domain will be discussed in the next section when the DHS unilaterally established a Disinformation Governance Board without a Congressional mandate.

Department of Defense as Unintended Mainline of Effort

Despite the presidential and congressional expectation that the GEC serve as the lead for countering information operations, the tangible outcomes in this policy arena have arguably emerged from the Department of the Defense (DOD) – particularly U.S. Cyber Command (USCYBERCOM) and the National Security Agency (NSA) – and the DOD's involvement shows no signs of stopping in the near future. Ironically, if someone had predicted 10 years ago that the United States military would have a long-term mandate to defend the integrity of elections at home, he or she would have been summarily laughed out of the room. Yet in 2020, one did not even have to look beyond the headlines of a governmental press release to encounter the claim: "DOD Has Enduring Role in Election Defense" (US Dept. of Defense 2020).

In fairness, the DOD being tapped to rapidly address atypical problems is not uncommon; to the contrary, Halperin (1972) finds that presidents commonly choose the military because its "tradition of discipline, efficiency and a clearly delineated chain of command increases the probability that precise orders will be observed and carried out with dispatch." In this case, Article II, the 2001 Authorization for the Use of Military Force, the Department of Homeland Security declaring election security as critical infrastructure, the FY19 NDAA, and National Security Presidential Memorandum 13 (NSPM-13) provided sufficient legal footing for the military to counter the threat of state-sponsored trolls targeting U.S. elections (Ney 2020) – but at the same time, authorizing the military to defend the integrity of U.S. elections had the potential to be irreconcilably

paradoxical with the notions of *posse comitatus* and broader civil liberties. Nevertheless, the DOD, in stereotypical fashion, seized the task and began moving out.

The majority of these governmental activities are shrouded in secrecy (and rightly so to preserve sources, methods, capabilities, etc.), but the DOD's activities in this policy domain are not entirely unknowable. The military's main line of effort for election defense is through the joint USCYBERCOM-NSA entity known as the Election Steering Group, whose primary objectives are to "generate insights on foreign adversaries who may interfere with or influence elections, bolster domestic defense by sharing information with interagency, industry, and allied partners, and impose costs on foreign actors who seek to undermine democratic processes" (CNMF 2022). The insights generated about adversaries by the partnership with the intelligence community's primary cryptologic agency then (presumably) tee up offensive cyberspace operations such as USCYBERCOM's takedown of a Russian troll farm (Barnes 2019).

From my practitioner experience, as mentioned previously I served as part of the 16th Air Force (Air Forces Cyber) stand-up (January 2020-July 2021) where we integrated the U.S. Air Force's intelligence and cyberspace capabilities under a unified command. One of our principal missions at that time was to defend the 2020 elections from foreign malign influence (Cohen 2019). Russia dominated many of our discussions, but at the time I simply assumed that was a byproduct of our command relationship as the operational cyber component to the geographic combatant command U.S. European Command (Matishak 2022) which is also dual-hatted as the DOD's global coordinating authority for the Russian problem set (Cavoli 2024, 18). However, in hindsight I do not

recall discussions within my purview where the other cyberspace components talked about other adversaries (e.g., China, Iran) in the same way we discussed Russia – especially not within the context of counter information operations or election defense. I also do not recall a time where we were taking our orders from the State Department's GEC (but the same cannot be said of instances where colleagues implied the GEC was behind the proverbial curve).

Department of Treasury Sanctions & Department of Justice Indictments

In an effort to impose costs upon troll sponsors and deter future information operations, the Department of Treasury (2021) directed sanctions at Yevgeniy Prigozhin for his role in financing the IRA troll farm's activities during the 2016 and 2020 elections. Similarly, the Department of Justice began to indict individuals involved in these exploits in the same way it would indict other foreign national violations of applicable U.S. laws. Russia (US Dept. of Justice 2018) and Iran (US Dept. of Justice 2021) both have indictments for interfering in American elections. China (US Dept. of Justice 2022) also has an indictment for using social media trolls, but the description of this criminal activity is different than the Russian and Iranian indictments. Despite making specific mention to trolls "targeting U.S. residents whose political views and actions are disfavored by the PRC government" and "covertly spreading propaganda to undermine confidence in our democratic processes," the press release for this indictment does not use the word "election"[14] at any point. In all of these cases, however, the

[14] This will be developed further in the Future Research section below, but for now the absence of an electoral reference stands as an outlier in the sense that the government seems to treat China differently than Iran and Russia and tends not to regard them as interfering in elections with their trolls.

general trend is reactive policy solutions to the problem of state-sponsored trolls on social media that emphasize combatting the promotion of ideas targeting specific populations in the operational context of elections and/or other political activities.

New Policy Solutions for State-Sponsored Trolls

In parallel with co-opting existing structures and processes to address the problem of state-sponsored trolls on social media, new policy mechanisms also began to percolate within the policymaking forums. As mentioned above, there is at times a temptation to treat major shifts in the geopolitical/technological landscape as *revolutionary* (i.e., treat it as something "new under the sun") rather than *evolutionary* – and this matter appears to have succumb to that temptation.

Regarding troll-waged information operations as a "new front" in the long history of great power competition (Jensen, Valeriano, & Manness 2020, 58), academics and policymakers alike saw these cyber-enabled malign actions as "an indicator of future operations that will target voting systems and the broader information environment in new and dangerous ways (US Cyber Solarium Commission 2020, 12).

> The power to hurt has become the power to hurt online. Just as the nuclear age heralded important changes to conceptualizing the use of force to achieve political objectives, the connectivity of the twenty-first century alters how rival states seek a position of relative advantage and coerce their adversaries (Jensen, Valeriano, & Manness 2020, 58-59).

And in the rush to label this as something completely new – "the *new threat* of using fake accounts to amplify divisive material and deceptively influence civic discourse" (Bills 2020; emphasis added); "the *new threat* of cognitive subversion" (Rosner & Siman-Tov 2018; emphasis added); or the "*new threat* developments to which the U.S. and [its] allies

are not well-equipped to respond" (Polyakova 2020; emphasis added) – the weightiness of the focusing event arguably drove framing the problem as primarily originating from Russia and primarily affecting elections (Chapter 2). Thus, it is unsurprising the proposed policy mechanisms centered around public education initiatives about the "new" problem and creating new governmental entities to combat it.

Public Education Initiatives

When scholars such as Joseph Nye (2019, 69) assert that "the defense of democracy in an age of cyber information war cannot rely on technology alone," it can result in policy mechanisms such as the US Cyberspace Solarium Commission's (2020, 69) recommendation that "digital literacy should be coupled with civics education explaining what democracy is, how individuals can hold their leadership accountable, and why democracy must be nurtured and protected." But does it not stand to reason that effective digital literacy training enabling people to recognize state-sponsored trolls in all contexts is inherently superior to reductively predisposing training recipients to apply their digital literacy through the lens of civics (especially given that state-sponsored trolls have near infinite applications and are only limited by operator creativity)? Similarly, at what point does federally sanctioned digital literacy coupled with civics education inadvertently become a form of political activities in and of itself contrary to the mandate for the civil service to be apolitical (e.g., Pendleton Act, Hatch Act)?

Nevertheless, digital literacy training fused with civics gained traction. In 2020, seven congressional representatives[15] formed the Congressional Task Force on Digital Citizenship focused on "promoting policies that encourage good digital citizenship," one of which being resources for "identifying misinformation and disinformation online and on social media" (Wexton n.d.). In 2022, bicameral legislation was introduced by two U.S. Senators and one Representative[16] to create grant programs promoting digital literacy to think critically and identify disinformation online (i.e., The Digital Citizenship and Media Literacy Act through the Department of Commerce and The Veterans Online Information and Cybersecurity Empowerment [VOICE] Act through the Department of Veterans Affairs [Smith 2022]).

The Department of Homeland Security (DHS) and its subordinate component the Cybersecurity and Infrastructure Security Agency (CISA) took a different approach to media literacy training by commissioning a series of graphic novels, the very first of which is entitled *Real Fake* and "demonstrates how threat actors capitalize on political and social issues (especially around election cycles) to stealthily plant doubt in the minds of targeted audiences and steer their opinion" (CISA n.d.). This graphic novel is replete with all of Chapter 2's framing trends. It hinges on election interference, it portrays Russia as the state-sponsor antagonist of the story (Figure 4.1), and emphasizes that the goal of the troll farm is promoting ideas through deceptively realistic content that generates social media engagement metrics (Figure 4.2). *Real Fake* also ends with a note

[15] Rep. Jennifer Wexton (D-VA), Rep. Don Beyer (D-VA), Rep. David Cicilline (D-RI), Rep. Yvette Clarke (D-NY), Rep. Bill Foster (D-IL), Rep. Bill Keating (D-MA), Rep. Zoe Lofgren (D-CA). There were no Republican members at the Task Force's inception and remain no Republican members at present.
[16] Sen. Michael Bennet (D-CO), Sen. Amy Klobuchar (D-MN), and Rep. Elissa Slotkin (D-MI).

from CISA stating “Disinformation is an existential threat to the United States, our democratic way of life, and the infrastructure on which it relies” which only further reinforces the contextualization that state-sponsored trolls and the information operations they wage are primarily an election-related problem.[17]

Figure 4.1: *Real Fake* Excerpts Portraying Russia as Antagonist

[17] The reference to “infrastructure” in this quote is likely a reference to elections given DHS’s designation of elections as critical infrastructure in 2017. Furthermore, the graphic novels are featured content of CISA’s Counter Foreign Influence Task Force (n.d.) that states its overview as: “CISA reduces risk to U.S. critical infrastructure by building resilience to foreign influence operations and disinformation. Through these efforts, CISA helps the American people understand the scope and scale of these activities targeting election infrastructure and enables them to take action to mitigate associated risks.”

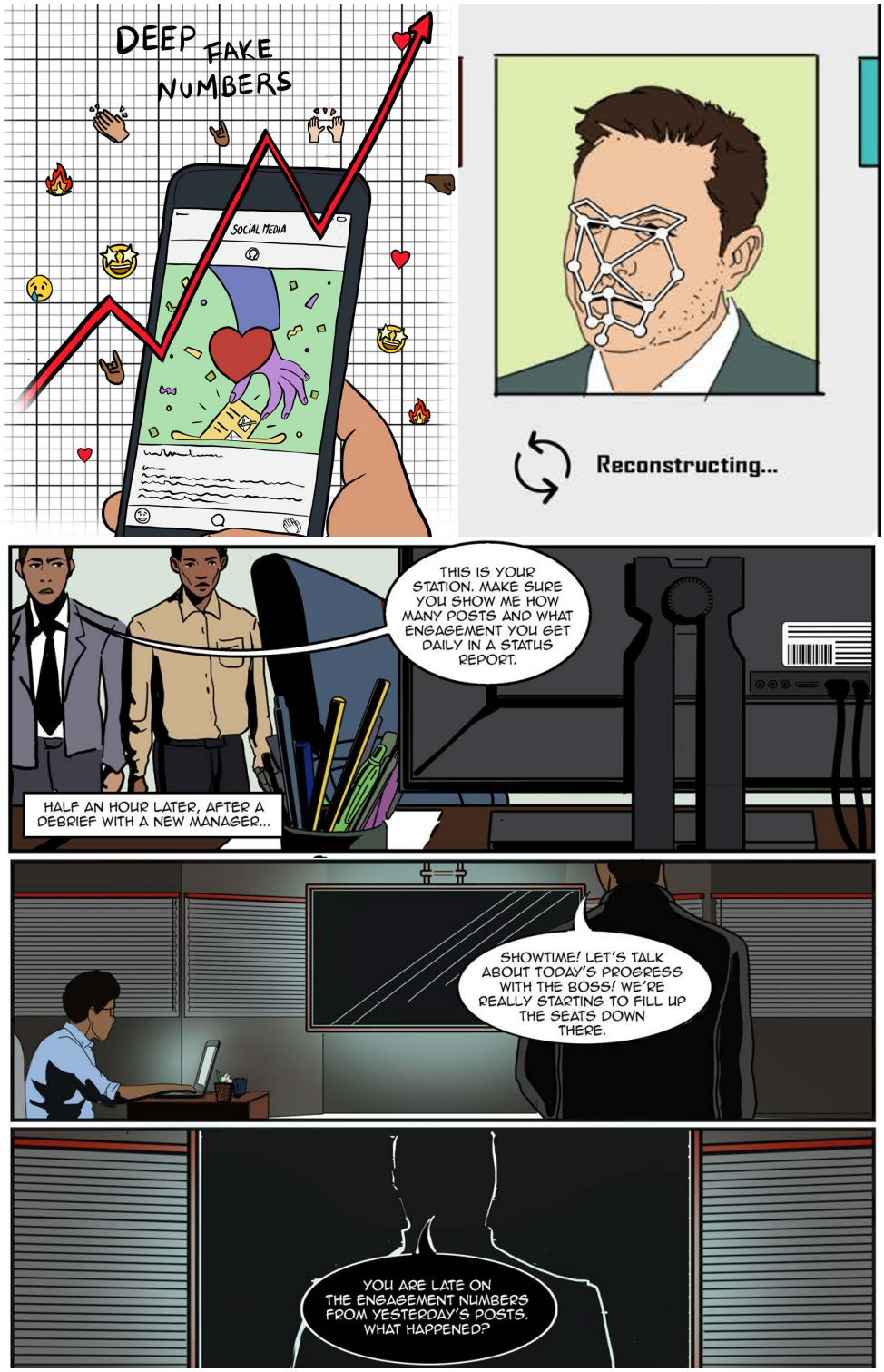

Figure 4.2: *Real Fake* Excerpts Emphasizing Realistic Content & Engagement

New Governmental Entities

As discussed in the previous section, the initial governmental response to the problem of state-sponsored trolls was to co-opt and modify existing entities (e.g., expanding the GEC's mission, mobilizing the military). Other responses took the form of new *ad hoc* entities such as the fall 2017 stand up of the Federal Bureau of Investigation's (FBI) Foreign Influence Task Force[18] which spanned the Counterterrorism, Cyber, Counterintelligence, and Criminal divisions. But unlike many of the aforementioned governmental responses to national security and/or foreign policy focusing events (e.g., announcing the creation of DHS less than one year after 9/11), the Russian activities of 2016 did not induce rapid large-scale change within the U.S. interagency.

In April 2022, six years after the focusing event, DHS stood up a Disinformation Governance Board helmed by Russian information operations researcher Nina Jankowicz – but this entity only lasted three weeks before being "paused" and ultimately dissolved in the wake of First Amendment and civil liberty concerns (Bond 2022). The termination of the Disinformation Governance Board occurred at approximately the same time the DHS Inspector General highlighted that the department did not have a unified strategy for countering disinformation.[19] A year later, the DOD would activate the Influence and

[18] The task force emphasizes "false personas and fabricated stories on social media platforms" and further perpetuates the previously encountered problem-framing trends by stating: "The goal of these foreign influence operations directed against the United States is to spread disinformation, sow discord, and, ultimately, *undermine confidence in our democratic institutions and values*" (FBI n.d., emphasis added).

[19] This Inspector General report also raises questions regarding overlapping missions on the part of DHS and GEC: "CISA and I&A [Intelligence & Analysis] also with the U.S. Department of State's (State Department) Global Engagement Center on countering disinformation. According to a State Department

Perception Management Office to combat disinformation (Holly 2023; Klippenstein 2023b) and the Irregular Warfare Center[20] as an organization with some equity stake holdings in the information operations arena.

In September 2022 the Director of National Intelligence (DNI) activated the Foreign Malign Influence Center (FMIC), six years after the focusing event and three years after the Legislative Branch established its existence through an amendment of the National Security Act of 1947 (DeVine 2023). Describing the FMIC's origin, one headline reads: "The Government Created a New Disinformation Office to Oversee All the Other Ones" (Klippenstein 2023a). While there is an undeniable layer of cynicism in the phrasing, the journalist also has a point. The FMIC (n.d.) has a mission of "defending America's democratic institutions" and does so in three areas:

- *Mission Management* – integrates government agencies, facilitates information sharing, and develops strategies to mitigate FMI.
- *Partner Engagement* – builds relationships with external partners, including government, civil society, and the private sector to share information and mitigate FMI activities.
- *Analytic Integration* – advances strategic analysis on the FMI problem set, synchronizes analytic efforts across the Intelligence Community, and provides comprehensive assessments to decisionmakers.

official, when the Global Engagement Center identifies disinformation campaigns abroad, it shares its analysis and reports with CISA and I&A to improve DHS' understanding of adversarial tactics, techniques, and procedures in spreading disinformation. The official added that another joint effort between CISA and the State Department involved working on Harmony Square, an online game that teaches players to recognize disinformation" (DHS 2022, 9)

[20] I attended an IWC symposium in spring 2024 where one of the world's foremost leading experts on state-sponsored social media trolls presented under the Chatham House Rule. While speaking on information operations, this expert stated: "I cannot talk about China enough…[*paused for dramatic effect, leaned into the microphone*]…China! China! China!" The moderator laughed, then without hesitation immediately shifted the conversation back to Russia and never returned to the topic of Chinese information operations.

This language substantively overlaps the NDAA mandate for the State Department's GEC, thereby lending credence to the redundancy criticism contained within the headline. In addition to the language of "defending America's democratic institutions" the FMIC's "process for notifying the public" is only applicable to election interference – further blurring the proverbial line between interagency *cooperation* and interagency *redundancy* given the ever-increasing number of entities involved in preserving election integrity (not to mention begging the question: What is the process for notifying the public in a non-election context?). Moreover, the overall intent for the FMIC is ambiguous given that it is a stopgap measure slated to sunset at the discretion of the DNI on December 31, 2028, and upon termination the "intelligence community would retain responsibility for assessing and providing warning of the threat of foreign efforts to interfere with U.S. elections" (DeVine 2023, 1). But the institutional efficacy of these measures notwithstanding, the broader trend within these new policy solutions once again places considerable emphasis on countering the promotion of ideas (e.g., "disinformation") targeting specific populations in the operational context of elections and/or other political activities.

Free-Market Policy Solutions for State-Sponsored Trolls

Finally, any discussion of policy solutions in the United States would be incomplete without acknowledging attempts to induce the free-market to solve the problem – and state-sponsored trolls are no exception, especially since social media straddles the divide between *private* and *public*. For instance, because social media platforms are privately owned companies one might argue that trolls should be primarily

addressed in accordance with the terms of service agreement between the platform and the user. On the other hand, since these platforms provide a public good by serving as a digital town square then the government has potential standing to ensure compliance with the First Amendment and equality in access similar to public utilities and common carrier regulations. Regardless of whether social media is viewed as primarily private or public issue, the market and governmental pressures seem to induce policy choices similar to those of the previous sections – actively combatting and refuting the promotion of illicit ideas in the context of elections and/or political functions (e.g., Congress pressured social media companies to "protect the census from disinformation" [Macagnone 2019]).

A key element of the public-private dichotomy surrounding the issue of state-sponsored trolls on social media is 47 U.S. Code § 230 (more commonly referred to as *Section 230*) which insulates platforms from liability with regards to the statements or actions of users (i.e., "No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider" [Legal Information Institute n.d.]). Once the "legal backbone of the internet," this policy has come under fire from all three governmental branches and both parties (Morrison 2023). Then-candidate Joe Biden stated that "Section 230 should be revoked, immediately should be revoked" (NYT Editorial Board 2020). President Trump expressed concerns that Section 230's "Good Samaritan" protections allowed for censorship of conservative voices (Villasenor 2020) while also providing no incentives

for the platforms to deal with the problem of state-sponsored trolls – sentiments also explicitly contained in his *Executive Order on Preventing Online Censorship*.[21]

The threat of removing the liability protections afforded by Section 230 is theoretically designed to induce firms within the marketplace to respond based on their own self-interest. The argument contends that by holding social media platforms liable,

> Their self-interest will lead them to remove such content, lest they be held liable for knowingly promoting terrorism…[but] not just to speech that could conceivably lead to terrorism, but also all kinds of speech that could potentially lead to harm, like challenging officially approved medical practices or alleging corruption, sexual or other physical assault, or racism (Funk 2023).

Whether or not it was in response to the threat of Section 230 nullification specifically, governmental pressures in general, or the broader adverse impact state-sponsored trolls could have on their corporate bottom lines, the point of emphasis here is that the social media platforms have responded to external forces.

All the parent-companies of social media platforms claim they have taken major steps to protect elections (Scola 2020). Meta sought to help Americans register to vote and protect election integrity from foreign influence via its "voter interference policies:"

> *Since 2016* [emphasis added], we've made substantial investments, built more teams and have worked with experts and policymakers to focus our efforts in the right places. We worked on more than 200 elections around the globe since then, learning from each, and now have more than 35,000 people across the company working on safety and security issues (Rosen 2020).

Twitter took similar steps with its "Civic Integrity Policy:"

> Twitter plays a critical role around the globe by empowering democratic conversation, driving civic participation, facilitating meaningful political debate,

[21] "At the same time online platforms are invoking inconsistent, irrational, and groundless justifications to censor or otherwise restrict Americans' speech here at home, several online platforms are profiting from and promoting the aggression and disinformation spread by foreign governments like China."

> and enabling people to hold those in power accountable. But we know that this cannot be achieved unless the integrity of this critical dialogue on Twitter is protected from attempts – both foreign and domestic – to undermine it. Today, we're announcing additional, significant product and enforcement updates that will increase context and encourage more thoughtful consideration before Tweets are amplified (Gadde & Beykpour 2020).

Yet with each of these types of policy decisions, the emphasis continued to drift away from the problem of state-sponsored trolls and towards policing the shared content (i.e., the ideas promoted) within the context of elections and other political functions (e.g., Google and Facebook launched fact-checking grants of $6.5M and $1M respectively, which in turn partly fueled the 200% increase in fact-checking organizations during the Trump Administration [Fischer 2020]).

Problem Framing & Policy Solutions

Based upon the survey of policy solutions in the sections above, the governmental trends in response to trolls brings this dissertation full-circle: *What is the Relationship Between How the Problem of State-Sponsored Trolls on Social Media is Framed & the Way U.S. Federal Policy Attempts to Solve It?* Chapter 1 identified trends in the academic literature that appear to portray trolls as a monolithic phenomenon centering upon having a believable account capable of promoting ideas. Chapter 2 subsequently identified these same trends seem to be demonstrably observable in the news reporting about state-sponsored trolls on social media as well, finding consistent emphasis on elections, Russia, and promoting ideas. The homogeneity within the discourse of Chapters 1 and 2 stand in stark contrast to the findings of Chapter 3, revealing a more heterogeneous reality than the academic literature and newspapers otherwise suggests.

In an ideal world, the policy mechanisms would share greater commonality with Chapter 3 in that the solutions should take into account the extremely multifaceted nature of the problem (i.e., reality). Regrettably for sound policymaking, the governmental solutions seem to correspond more with the incomplete representations contained in Chapters 1 and 2. Thus, for the purposes of this dissertation, the overarching inferential answer to the primary research question is that there does appear to be *a corresponding relationship between the way in which the problem of state-sponsored trolls on social media is framed and the way U.S. federal policy attempts to solve it.*

At face value, this seems intuitive. Focusing events in national security and foreign policy contexts often impact how problems and their corresponding policy solutions are framed (Birkland 1997; Warren 2024) – thus it is to be expected that the Russian efforts to polarize and divide Americans during the 2016 elections would have a sizable impact on the societal understanding of state-sponsored trolls in the same way the Barbary States or 9/11 shaped the understanding of *piracy* and *terrorism* respectively without being the only type of *pirates* or *terrorists*. Ironically, even the cited excerpt from Shakespeare in the present chapter's introduction involves a regime change at the head of state level. But superficially equivocating state-sponsored trolls with targeting the electoral politics of free societies skews the problem as smaller than it really is (e.g., offensive cyber operations and criminal indictments are only pursuing one genre of trolls). The same can be said of making default comparisons to Russia[22] and promoting

[22] While Chapter 3 demonstrates that much of what China does with its trolls is fundamentally different than what Russia does, the FMIC (2024, 3) officially compares China to the Russian "playbook" in ways nearly identical to the news surveyed in Chapter 2: "Beijing's growing efforts to actively exploit perceived

ideas, it further skews and distorts the problem in ways that are counterproductive to developing effective policy solutions.

The "imperfect understanding" (Rowling 2001, xv) then snowballs when media literacy training initiatives become a repeatedly pursued policy to counter state-sponsored trolls on social media. Media literacy can only be as effective as the empirical foundations underpinning the curriculum. Based on the monolithic framing tendencies identified in Chapter 1 and Chapter 2, digital literacy combined with critical thinking and civics education may theoretically be effective in helping people recognize idea promotion in the context of elections – but how could it teach someone to recognize idea demotion or to identify an information operation not targeting democratic processes? Moreover, how does one go about learning to recognize that a truthful idea is not on their social media feed because it has been censored via algorithmic manipulation (i.e., how does one prove the negative)? Pragmatically speaking: If the problem of state-sponsored trolls on social media is currently framed in an incomplete manner by the community of experts and within policymaking forums, how can digital literacy training for lay citizens be relied upon as a primary line of effort to overcome the challenge?

Similar efficacy concerns can be raised with combatting disinformation and counterfactual narratives with fact-checking style interventions. For one, it concedes the problem as endemic by not focusing initial policy actions on keeping state-sponsored trolls from gaining access to social media platforms at the onset. Second, it is only

---

U.S. societal divisions using its online personas move it closer to Moscow's playbook for influence operations."

effective against direct idea promotion and offers little-to-no tractability against indirect promotion or any form of demotion. Lastly, and arguably most importantly, the implementation of these types of policies (whether by governmental agencies, academia, or the platforms themselves) have potentially polarized faith in institutions and democratic processes even further – the very things the aforementioned policies repeatedly claimed to defend.

In an environment where informational threats exceed the capacity to respond (i.e., it is impossible to counter every disingenuous narrative), bureaucratic prioritization of tasks is influenced by organizational culture and the internal composition of the agencies involved (Wilson 1989; Wood & Waterman 1994; Potter 2019). When this bureaucratic dynamic converged with policing information in the context of elections, that which was meant to be apolitical became political. As noted by the U.S. House of Representatives (2023, 1) Committee on the Judiciary and the Select Subcommittee on the Weaponization of the Federal Government:

> Following the 2016 presidential election, a sensationalized narrative emerged that foreign "disinformation" affected the integrity of the election. These claims, fueled by left-wing election denialism about the legitimacy of President Trump's victory, sparked a new focus on the role of social media platforms in spreading such information. "Disinformation" think tanks and "experts," government task forces, and university centers were formed, all to study and combat the alleged rise in alleged mis- and disinformation. As the House Committee on the Judiciary and the Select Subcommittee on the Weaponization of the Federal Government have shown previously, these efforts to combat so-called foreign influence and misinformation quickly mutated to include domestic – that is, American – speech.

While this assertion does contain intrinsic levels of theatricality and political grandstanding, what must not be overlooked is the underlying core assertions are not without merit. For instance, there is no compelling legal or moral argument to defend the

decision to de-platform Donald Trump while simultaneously allowing the Taliban and Russian Ministry of Defense after the 2022 invasion of Ukraine to both maintain official accounts (Concha 2021; Warren 2022). Similarly, Reddit removed millions of pro-Trump postings (Copland & Davis 2020) while Facebook and Twitter algorithmically suppressed the Hunter Biden laptop scandal (WSJ 2022) – which proved to be real despite being repeatedly denounced as "Russian disinformation" (Robertson 2024; Jacques 2024).

Even if one assumes noble intent on the part of all involved with these particular circumstances (i.e., it is inconsistent implementation rather than maliciously selective enforcement), disparate outcomes along partisan fault lines do nothing to restore faith and trust in institutions. To the contrary, disparities in outcomes only raise further questions about an institution's impartiality and overall legitimacy in serving the public good (not to mention concerns of First Amendment adherence). At a time where faith and trust in institutions are at historic lows (Saad 2023), "liberal institutions need both to deliver better results and to be widely seen and understood to do so" if the public's trust is to be regained (Malloch-Brown 2023). The same can be said of the social media platforms themselves considering concerns about platform moderation have been linked to those on the right migrating to alternative platforms while those on the left remain on mainstream platforms (Pew Research Center 2022; Klein 2020) – thereby creating a plurality of echo chambers rather than singularly meaningful discourse and more venues for state-sponsored trolls to engage in malign activities.

Areas for Future Research

Based upon this dissertation's findings, there are numerous areas for future research examining the problem of state-sponsored trolls on social media and the broader body of policy literature. These areas include, but are not limited to: additional problem framings; generating better data and analysis for the empirical study of state-sponsored trolls on social media; and broadening existing policy theory to better explain focusing events in the context of foreign policy and national security.

Additional Problem Framing Research

In terms of framing the problem of state-sponsored trolls, more findings could be derived from international news reporting as well as entertainment and pop culture. Additionally, more substantive content analysis beyond simple term frequency would also provide deeper and more robust insights into the discourse-driven problem framing.

International News Reporting

A similar methodology to Chapter 2 could be applied to international news reporting (e.g., *BBC News*, *The Sydney Morning Herald*, *The Times of India, The Telegraph*, *Al Jazeera*) to see how the problem is framed in other international or regional contexts. Similarities and differences in research findings could highlight potential opportunities for not only intergovernmental collaboration with strategic allies and partners but also assist in developing a more effective policymaking community of interest that avoids previous mistakes and expedites the adoption of best practices.

Entertainment & Pop Culture

In experimenting with the Boolean logic used to construct the newspaper dataset,

I encountered articles that made recommendations regarding what readers should watch on television. Expecting these to be false positives for the DreamWorks animated film *Trolls* (which some were), I also found on-topic references to the HBO documentary *Agents of Chaos* (released September 2020) which explores the Russian interference in the 2016 U.S. elections. Even though the final Boolean omitted articles referencing this documentary, their existence points to another qualitative problem framing research domain for future researchers: Entertainment and Pop Culture.

"Many of the words and images generated and marketed by the 'pop culture' industry attempt to reflect the realities of American life and frequently help shape those realities" (Gitelson, Dudley, & Dubnick, 2014). Considering the exportation of American entertainment globally, it is not irrational to think pop culture's influence in shaping realities does not stop at our shores. Consequently it only seems rational to expect that the problem of state-sponsored trolls on social media would also find its way onto our various screens – which it has, even in my own anecdotal television consumption.

Sometimes the references are somewhat subtle. For instance, in S4:E20 of NBC/Universal's *Superstore* (aired May 9, 2019), two employees create fake twitter accounts pretending to be disgruntled customers to try and trick corporate into giving them more resources.[23] Similarly, the CBS series *The Good Fight* (premiered February 19, 2017) begins with Diane Lockhart (played by Christine Baranski) watching in horror

[23] Interestingly enough, the dialogue in this episode makes it sound as if detecting trolls is easy. After the accounts were revealed to be controlled by someone in the store, other employees began to comment: "They were so obvious about it too! All the fake customer accounts were created on the same day and they each only tweeted once!"

as Trump takes the oath of office while the season's remaining nine episodes attaches key elements of the plot to social media trolls, fake news, and the alt-Right being the source of society's worst problems. Lastly, in S2:E3 of the CBS series *Yellowstone* (aired July 10, 2019), Jamie Dutton declares, "I have a donor, willing to fund my campaign" – to which his sister Beth responds: "Let me guess, he's Russian and wants you to do a Facebook blast?"

Other times the references are overt and undeniable. A principal storyline component for season seven of Showtime's *Homeland* (premiered February 11, 2018) is a social media troll farm run by a character resembling Steve Bannon and a compromised U.S. President-Elect…all the while having episodes entitled: "Active Measures" (Episode 5); "Lies, Amplifiers, Fucking Twitter" (Episode 8); and "Useful Idiot" (Episode 9). In that same vein, S2:E15 of the ABC show *Quantico* (aired March 27, 2017) hinges on a private military company hiring internet trolls to disseminate fake news. In parallel with the plot vehicle of social media trolls, this episode introduces the character Henry Roarke by name as a member of a conspiracy to destroy America (Spoiler alert: Over the course of the six episodes the character actually appears in, Roarke uses collusive means to become President; declares that he will prioritize border security policy, a Muslim registry, and purging the rigged system; and ultimately commits suicide after being exposed for having corrupt ties to Russia).

In response to *Quantico* specifically, one Hollywood commentator noted that this season differed from its predecessor and had "shift[ed] course to follow a ripped-from-the-headlines approach" (Strause 2017). But this can be said in varying degrees of all the

aforementioned, particularly since not only do they offer some pseudo-replication of this chapter's findings (e.g., framings emphasizing Russia, elections) but also these shows coincide with the apex of the newspaper article dataset (i.e., 2017-2020) despite the reality that producing television programing is more time/resource intensive than publishing an article. Simply put, television shows such as these, the mere existence of an off-Broadway play *Russian Troll Farm: A Workplace Comedy*, etc. demonstrate that pop culture offers robust and vibrant potential for additional data to explore how the problem of state-sponsored trolls in social media is framed in societal discourse.

Partisan Framings

More research into potential partisan bias in the framing of state-sponsored trolls is also needed. Experientially, it has often seemed that Russian trolls operating in the context of U.S. elections have been consistently described as pro-Trump/Republicans – yet there at times seems to be a reticence to call Iranian or Chinese trolls as pro-Biden/Democrat, even though the Intelligence Community assessed that this was in fact the preference of Tehran and Beijing (CNN 2020). For instance, one journalist states:

> The influence operations in these countries, however, do not all share Russia's demonstrated preference for Trump and other Republicans. The Iranians, for example, typically oppose Trump in their disinformation messaging, criticizing his decision to pull the United States out of the 2015 nuclear deal with Iran and administration policy on other issues (Barnes 2018).

Similarly, another writes:

> Chinese social media accounts are not happy with President Donald Trump. A network of accounts on multiple platforms has been criticizing Trump and broadcasting more positive images of Democratic presidential candidate Joe Biden, as part of an apparent campaign to rebuke the White House (Stone 2020).

In both these examples, there seems to be a framing decision to juxtapose the Russians as actively campaigning on behalf of specific candidates…whereas the Iranians and Chinese are not campaigning for Democrats so much as expressing disagreement about policy. Put a different way, in both of these examples the framing of the sentence centers upon Trump (i.e., the Iranians are not pro-Biden but rather they "oppose Trump;" China's positive publicity for then-candidate Joe Biden are not election related activities but rather an effort to "rebuke the White House" occupied by Trump).

This framing trend can also be observed in governmental documents. President Biden appointed Avril Haines as the DNI and she was sworn in January 21, 2021. Less than two months later, she declassified a report entitled "Foreign Threats to the 2020 US Federal Elections." After concluding there were no indications that a foreign actor changed any technical aspects of the voting process (Key Judgment 1), the National Intelligence Council (2021, *i*) goes on to list three additional assessments:

- Key Judgment 2: We assess that Russian President Putin authorized, and a range of Russian government organizations conducted, influence operations aimed at denigrating President Biden's candidacy and the Democratic Party, supporting former President Trump, undermining public confidence in the electoral process, and exacerbating sociopolitical divisions in the US. Unlike in 2016, we did not see persistent Russian cyber efforts to gain access to election infrastructure. WE have high confidence in our assessment; Russian state and proxy actors who all serve the Kremlin's interests worked to affect US public perceptions in a consistent manner. A key element of Moscow's strategy this election cycle was its use of proxies linked to Russian intelligence to push influence narratives – including misleading or unsubstantiated allegations against President Biden – to US media organizations, US officials, and prominent US individuals, including some close to former President Trump and his administration.

- Key Judgment 3: We assess that Iran carried out a multi-pronged covert influence campaign intended to undercut former President Trump's reelection prospects – though without directly promoting his rivals – undermine public confidence in the electoral process and US institutions, and sow division and exacerbate societal

tensions in the US. We have high confidence in this assessment. We assess that Supreme Leader Khamenei authorized the campaign and Iran's military and intelligence services implemented it using overt and covert messaging and cyber operations.

- Key Judgment 4: We assess that China did not deploy interference efforts and considered but did not deploy influence efforts intended to change the outcome of the US Presidential election. We have high confidence in this judgment. China sought stability in its relationships with the United States, did not view either election outcome as being advantageous enough for China to risk getting caught meddling, and assessed its traditional influence tools – primarily targeted economic measures and lobbying – would be sufficient to meets its goal of shaping US China policy regardless of the winner. The NIO for Cyber assesses, however, that China did take some steps to try to undermine former President Trump's reelection.

Once again the issue is framed as Russia actively campaigning for Trump/Republicans while Iran and China are described as being anti-Trump rather than pro-Biden, a distinction without a difference for an election in a two-party system guaranteed to produce one of two outcomes.

In my experience as a state-sponsored troll researcher, this potential state-sponsored troll framing tendency has often felt as though it places those right-of-center on the political spectrum in the role of antagonist (which so do the pop culture references above from *The Good Fight*, *Homeland*, and *Quantico*). Anecdotally, these past three years of my doctoral journey have included many opportunities to teach undergraduates as well as present to my graduate school peers – and in both settings I have routinely conducted the same informal poll. I begin by asking those in the room to raise their hand if they had heard that Russia wanted a Trump presidency, to which those in attendance unanimously raise their hands. Yet when I ask them to keep their hands raised if they had also heard that Iran and China wanted a Biden presidency, without fail nearly every hand

in the room slowly returns to its owner's side. My teaching point in response to these results is always to emphasize that this knowledge differential is problematic *because no despot on any side of any ocean should get any say in who sits behind the Resolute Desk* (i.e., our politics are not the problem, state-sponsored trolls are).

When I requested that ChatGPT attempt to identify potential framing differences along partisan fault lines in the news article dataset from Chapter 2 – it rejected my requests and terminated the LLMs calculations, expressing concerns that the answer would violate OpenAI's terms of service. This came at the same time critics and watchdogs are expressing concerns of partisan bias in LLMs and other forms of generative AI (Baum & Villasenor 2023; Heikkilä 2023; De Vynck 2023). Since playing politics with national security is never the most advantageous course of action, if partisan framings of trolls do in fact exist it is paramount that free societies acknowledge and remedy the issue in order to more effectively drive policy discussions.

Effectiveness Framings

As mentioned in Chapter 1, the closest thing to taxonomy currently in existence is a rudimentary bifurcation of state-sponsored trolls being either *effective* or *ineffective* based on superficial observations regarding a profile's sophistication (Warren, Linvill, & Warren 2023). Consequently, I have offered evidence throughout this dissertation to support my contention that the efficacy of trolls should not be evaluated based on their appearance but rather on whether or not they accomplish their objectives. Since said goals may or may not be to assert active, direct influence upon unsuspecting users,

evaluation criteria should not succumb to the countability bias of relying on social media engagement metrics.

Although Chapter 2 offered evidence that there is a potential overemphasis on a believable profiles generating engagement and Chapter 3 offered why that particular problem framing can lead many observers to completely misunderstand the information operations waged by CCP-sponsored trolls, content analysis techniques such as term frequency are unable to capture the tone, sentiment, etc. necessary to understand the way(s) in which effectiveness is discussed. Developing a coding strategy to manually examine how effectiveness is conceptually operationalized and explained within the newspaper dataset would offer more holistic insights into the phenomenon of trolls. In the event the manual review yields similar conclusions, we would have further confidence in the assessment of a monolithic and incomplete framing of the problem that needs to be rectified in support of future policy development. Conversely, if additional effectiveness grading scales are identified then this could further refine future work exploring the relationship between the troll persona investment and operational goals. Either outcome would advance the empirical understanding of the problem of state-sponsored trolls on social media and further establish the foundations upon which a taxonomy could be built.

Empirical Troll Research Needs Better Data

Chapter 3 demonstrated that countries invest in their trolls differently and that there is a relationship between troll investment and an information operation's goal(s). However, Chapter 3's single-greatest obstacle to more generalizable and concrete

findings is the absence of better and more robust datasets than the Twitter Information Operation Archive and the ESOC Trends in Online Influence Efforts. Researchers and the platforms themselves need to collaborate in order to make more data available if a taxonomy of state-sponsored trolls on social media is ever to be actualized (or to even move the research community beyond analysis of campaigns on Twitter).

Part of this particular issue could be overcome by developing standards for attribution within the state-sponsored troll research community of interest. At present, there are essentially two options for researchers. The first is to perform case study work on an individual information operation and attribute the campaign to a state-sponsor by citing media or expert reporting on the same campaign; this is similar to the ESOC approach but a specific example would be the Clemson University Media Forensics Hub report "The Five-Year Spam" (Warren et. al. 2023) where we examined Chinese trolls and attributed their sponsorship to the CCP through citations to Mandiant and other external (primarily cyber) experts about the same campaign. The other option is to simply accept what the platforms make available, which relegates researchers to the Twitter Information Operations Archive or disparate reports published by the platforms.

With regards to the latter option, this enables the platforms to set the *de facto* research agenda based entirely on what they do or do not make available[24]…begging the age-old question: *Who will watch the watchers?* There are no Western nations in the Twitter dataset of Chapter 3; the closest would be the Spaniard trolls, but those belong to

[24] Or in the case of Meta, not only do scholars have to apply for access to the platform's academic tools and data sandbox but Meta must also approve a researcher's findings prior to being submitted for publication.

the government in exile of the Catalonians. Meta and Twitter de-platformed a pro-U.S. information operation in 2022 (Frenkel & Hsu 2022), but only made select portions of the data available to Graphika and the Stanford Internet Observatory (2022, 2) for in-depth analysis. TikTok's first three quarterly reports on information operations taking place on its platform discussed 22 separate operations, one of which originated from Taiwan but none from China (Ryan 2023) – an omission that would later cease to be when a 2024 report was published (coincidentally after Congress passed measures to ban the platform due to CCP influence) that included details about a small, 16-account information operation emanating from China (Abbruzzese & Ingram 2024). A community of interest established methodology for identification, data generation, and attribution across platforms could introduce accountability and transparency to the problem of state-sponsored trolls on social media – both of which critically important at a time of waning faith and trust.

Problem Framing & Policy Theory

Connecting the implications of these findings back to the policy literature reveals the possibility of a theoretical contribution for future researchers to explore and test. Kingdon describes the problem, policy, and politics streams as mostly independent of one another, existing in parallel until circumstances cause them to become coupled. In the case of focusing events, exogenous forces create a window of opportunity for policymaking by shocking the streams into alignment and thereby bringing an issue to the forefront of the agenda. The 2016 election interference undeniably focused attention to the problem of state-sponsored trolls on social media and highlighted how current policy

was inadequate for the task of preventing future repeated issues, however the policymaking apparatus did not produce outputs commensurate with many of the previously encountered national security/foreign policy focusing events (Table 4.1). *But why does that seem to be the case?*

One possible explanation is the focusing event was enigmatic and its impacts intangibly undeterminable. The focusing events in Table 4.1 could be measured objectively in lives lost, dollars of damage, or something falling out of space at a time space was believed to have no manmade objects in orbit. But how does one measure *influence* or *persuasion* in a world of near-infinite exogeneity? Hillary Clinton claimed Russian trolls cost her the White House (Taylor 2017) and journalists interpreted the Mueller Report as proof Russia affected the vote (Bump 2019)…yet others have attempted to empirically connect state-sponsored trolls with changes in voting behavior and ultimately find no supporting evidence for a demonstrable linkage (Eady et. al 2023; Dolan 2019; Biddle 2023; Lawson 2019; Oremus 2019). While it is correct to say that these positions are mutually exclusive and cannot both be simultaneously true, it is impossible to determine with certainty which is the accurate representation of the focusing event's outcome. Thus, how can the focusing event be understood when it is suspended in a Schrödinger-esque state of being where it both exists and does not exist?

Another possible explanation, which is not irreconcilable with the previous, is the focusing event was interpreted and understood on a scale of political winners and losers as opposed to its actual, objective characteristics. Historically, national security and foreign policy matters were generally immune to the corrosiveness of politicization.

> Not so long ago, political experts assumed foreign policy and national security were above the public fray. They were the domain of diplomats, intelligence agents, and other career specialists. The average person on the street couldn't find most countries on a map, the thinking went – much less have a meaningful impact on the affairs of state (Irving 2023).

But when the Russian interference in the 2016 elections and matter of state-sponsored trolls on social media writ large became inextricably linked to domestic politics, the public entered the fray and with it the monolithic framings of the problem skewed reality.

Returning to Kingdon's streams metaphor, if a focusing event is theorized as aligning the streams into a window of opportunity – then perhaps it matters which direction the streams are theoretically traversed if the focusing event is to transcend beyond merely forcing an issue onto the agenda and translate into actual large-scale policymaking. The present chapter surveyed the disjointed, sluggish, and sporadic policy responses the U.S. federal government directed at the problem of state-sponsored trolls on social media. If the focusing event was in fact understood on a scale of political winners and losers as opposed to its actual characteristics, then the policy response could be conceptualized as traversing the *politics* stream first which preemptively reduced options in the *policy* stream and created *post facto* modifications to the collective understanding of the *problem* stream. Such a hypothetical would offer potential explanatory power for the disparity in policy outputs and governmental unity of effort in response to a focusing event such as 9/11 which centered upon the *problem* of transnational terrorism, then generated numerous *policy* options, and was ultimately synergized by bipartisan cooperation throughout the government (Figure 4.3).

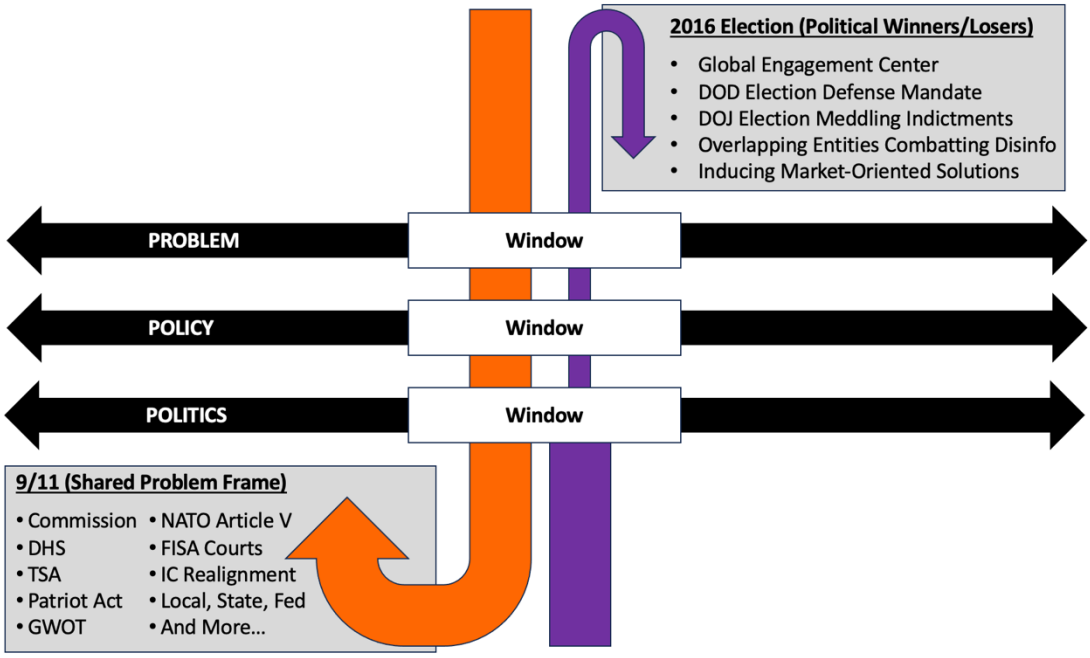

Figure 4.3: Possible Theoretical Expansion of Kingdon

Testing the potential legitimacy of an expended understanding of the Multiple Streams Framework in this manner would offer a possible bridge between the agenda setting literature and the scholarship surrounding policy formulation, policy learning, and policy implementation/evaluation. In the event the theoretical contribution offered in Figure 4.3 proves to be inconsistent with more methodologically rigorous probing, the resulting research would nonetheless provide empirical datapoints by which to examine the relationship between the nature of a national security/foreign policy focusing event and the resulting policy change or policy learning that does or does not occur (i.e., expands upon the work of Birkland into a different policy domain). Lastly, testing this theoretical contribution may also identify a gap in the existing models and frameworks that needs to be resolved by a new policymaking theory.

Conclusion

Ending this dissertation where it began, "imperfect understanding is often more dangerous than ignorance." As I postulated in Chapter 1, the U.S. has a collective understanding of state-sponsored trolls on social media that is imperfect – which in turn poses inherent dangers to the policymaking process and any attempt to address the problem with governmental interventions. This dissertation research sought to remedy elements of said imperfect understanding by first using Chapter 2 to identify that the problem is framed monolithically in media reporting, disproportionately emphasizing elections, Russia, and idea promotion through generating engagement. Chapter 3 then deep-dived the problem of state-sponsored trolls on social media, identifying a much more complex and heterogenous reality than is portrayed in the media while also bringing some order to the phenomenological chaos by finding a relationship between troll account investment and an information operation's goal(s). Unfortunately for the development of effective governmental interventions, the existing policy responses hold more in common with the incomplete sight-picture of Chapter 2 than the more holistic examination in Chapter 3.

These findings and the corresponding discussion beg one final question: *So what?* This dissertation has intentionally avoided making policy recommendations[25] in part because advocating normative policy position(s) exceed the scope of this study's exploration of the relationship between problem framing and solution development.

[25] I have written about what policies might be used to more effectively combat the problem of state-sponsored trolls on social media in a more appropriate venue (see Warren 2022).

However, there is an additional reason this dissertation has avoided taking a policy position and that is to make a demonstrable effort to generate apolitical insights in a manner consistent with the core values of the intelligence community.

As I outlined in Chapter 1, my overarching research design sought to weave together three strands in pursuing answers to my primary research question: Policy theory, the empirical study of trolls, and personal elements from my professional career. The first two strands are self-evident throughout this dissertation and thus it is now time to address the third: This dissertation should serve as a cautionary tale for intelligence analysts to better recognize the implications of their role as professional problem-framers for foreign policy and national security policymaking.

"Analytic objectivity and sound intelligence tradecraft ensure our nation's leaders receive unbiased and accurate intelligence to inform their decisions" (Office of the Director of National Intelligence n.d.). Put more succinctly, the IC's problem framings inform decisionmakers' policies. From a systems-thinking perspective, it stands to reason that valid inputs (i.e., problem framings) from the IC directly increase the probability of good outputs (i.e., effective policy) – but does not guarantee the quality of outputs because policymakers can always deviate, ignore, politicize, etc. the problem after the fact. Yet conversely, invalid (or incomplete) inputs from the IC essentially guarantee ineffective policymaking because there is an inherent incongruity and misalignment between problems and solutions. It is precisely this latter scenario that the IC must insulate itself from through an unwavering commitment to its mantra: *Speak truth to power*.

With regards to state-sponsored trolls on social media, there is some preliminary evidence that might indicate partisanship played a part in mischaracterizing the problem during efforts to craft policy solutions. At the same time, there are also events such as an intelligence briefer overstating the assessment of Russian interference in the 2020 elections along political fault lines to Congress (Diamond, Tapper, & Cohen 2020). Agnostic of any attempts to attribute the preponderance of blame for this particular set of circumstances, the broader takeaway is the critical importance of understanding a problem before trying to craft substantive solutions – and intelligence analyst are on the frontlines of problem framing for foreign policy and national security policymaking. Though this dissertation is unlikely in and of itself to resolve the current policy issues surrounding the problem of state-sponsored trolls on social media, its value as a guided debrief through the pitfalls of this particular problem has historicity value for those looking at future emerging threats. After all, those who do not learn the lessons of history are doomed to repeat them.

APPENDICES

APPENDIX A

INVESTMENT SCORE INDIVIDUAL COMPONENT REGRESSIONS

The novel *Investment Score* dependent variable in Chapter 3 is an aggregated measure produced by identifying the following attributes about each individual troll account (N = 87,437):

- The troll persona has a profile description/biographical data (Y/N).
- The troll persona has claimed a geographic location (Y/N).
- The troll persona has provided a personal URL (Y/N).
- The troll persona has at least 10 other Twitter accounts it follows (Y/N).
- The troll persona has at least 50 other Twitter accounts it follows (Y/N).
- The troll persona has at least 100 other Twitter accounts it follows (Y/N).

Each component is quantified based on present (1) versus not-present (0) criteria and then combined to yield an *Investment Score* on a scale from 0-6. Table A.1 below portrays the distribution of how many trolls were scored across each of the seven possible *Investment Score* options relative to state-sponsorship.

Because accounts sharing the same score are not necessarily the same due to there being multiple ways to achieve a score (i.e., there are 64 different permutations along the six individual 0/1 components), the efficacy of the *Investment Score* as a dependent variable hinges upon there being no dominant pattern within the individual components. Thus in order to demonstrate the robustness and potential explanatory power of the *Investment Score* as a dependent variable, it is important to understand the variance of each individual component.

	0	1	2	2	4	5	6	Total
Armenia	--	--	3	7	5	16	4	35
Bangladesh	--	3	2	5	3	1	1	15
China	24,857	3,207	1,289	596	429	711	61	31,150
Cuba	19	19	44	47	189	189	19	526
Ecuador	300	247	146	126	85	111	4	1,019
Egypt	466	212	250	504	477	591	41	2,541
Honduras	823	848	797	359	148	101	28	3,104
Indonesia	103	101	135	134	180	116	26	795
Iran	740	411	720	1,095	1,775	2,034	250	7,025
Mexico	10	10	46	73	107	29	1	276
Russia	246	255	286	675	1,170	2,102	221	4,955
S. Arabia	645	956	1,155	2,218	2,753	3,627	235	11,589
Serbia	812	2,006	1,524	2,699	1,011	474	32	8,558
Spain	82	27	42	78	89	68	2	388
Tanzania	86	68	83	24	6	1	--	268
Thailand	474	300	111	35	6	--	--	926
Turkey	904	547	607	1,517	1,710	1,674	381	7,340
Uganda	18	16	25	57	157	137	8	418
UAE	470	427	584	819	825	1,074	49	4,248
Venezuela	221	160	322	429	360	440	329	2,261

Table A.1: Investment Score Distribution by Country

OLS Regressions by Investment Score Component

In order to examine the variance of the components individually, the following six OLS regression models were used to analyze the data relative to the state-sponsors of trolls in the dataset. To measure the use of *Description* by country:

$$\text{Description}_i = \beta_0 + \beta_1\text{Armenia}_i + \beta_2\text{Bangladesh}_i + \beta_3\text{Cuba}_i + \beta_4\text{Ecuador}_i + \beta_5\text{Egypt}_i + \beta_6\text{Honduras}_i + \beta_7\text{Indonesia}_i + \beta_8\text{Iran}_i + \beta_9\text{Mexico}_i + \beta_{10}\text{Russia}_i + \beta_{11}\text{SaudiArabia}_i + \beta_{12}\text{Serbia}_i + \beta_{13}\text{Spain}_i + \beta_{14}\text{Tanzania}_i + \beta_{15}\text{Thailand}_i + \beta_{16}\text{Turkey}_i + \beta_{17}\text{Uganda}_i + \beta_{18}\text{UAE}_i + \beta_{19}\text{Venezuela}_i + u_i$$

(Equation 3)

To measure the use of *Location* by country:

$$\text{Location}_i = \beta_0 + \beta_1\text{Armenia}_i + \beta_2\text{Bangladesh}_i + \beta_3\text{Cuba}_i + \beta_4\text{Ecuador}_i + \beta_5\text{Egypt}_i + \beta_6\text{Honduras}_i + \beta_7\text{Indonesia}_i + \beta_8\text{Iran}_i + \beta_9\text{Mexico}_i + \beta_{10}\text{Russia}_i + \beta_{11}\text{SaudiArabia}_i + \beta_{12}\text{Serbia}_i + \beta_{13}\text{Spain}_i + \beta_{14}\text{Tanzania}_i + \beta_{15}\text{Thailand}_i + \beta_{16}\text{Turkey}_i + \beta_{17}\text{Uganda}_i + \beta_{18}\text{UAE}_i + \beta_{19}\text{Venezuela}_i + u_i$$

(Equation 4)

To measure the use of *Profile URL* by country:

$$\text{ProfileURL}_i = \beta_0 + \beta_1\text{Armenia}_i + \beta_2\text{Bangladesh}_i + \beta_3\text{Cuba}_i + \beta_4\text{Ecuador}_i + \beta_5\text{Egypt}_i + \beta_6\text{Honduras}_i + \beta_7\text{Indonesia}_i + \beta_8\text{Iran}_i + \beta_9\text{Mexico}_i + \beta_{10}\text{Russia}_i + \beta_{11}\text{SaudiArabia}_i + \beta_{12}\text{Serbia}_i + \beta_{13}\text{Spain}_i + \beta_{14}\text{Tanzania}_i + \beta_{15}\text{Thailand}_i + \beta_{16}\text{Turkey}_i + \beta_{17}\text{Uganda}_i + \beta_{18}\text{UAE}_i + \beta_{19}\text{Venezuela}_i + u_i$$

(Equation 5)

To measure the use of *Follow10* by country:

$$\text{Follow10}_i = \beta_0 + \beta_1\text{Armenia}_i + \beta_2\text{Bangladesh}_i + \beta_3\text{Cuba}_i + \beta_4\text{Ecuador}_i + \beta_5\text{Egypt}_i + \beta_6\text{Honduras}_i + \beta_7\text{Indonesia}_i + \beta_8\text{Iran}_i + \beta_9\text{Mexico}_i + \beta_{10}\text{Russia}_i + \beta_{11}\text{SaudiArabia}_i + \beta_{12}\text{Serbia}_i + \beta_{13}\text{Spain}_i + \beta_{14}\text{Tanzania}_i + \beta_{15}\text{Thailand}_i + \beta_{16}\text{Turkey}_i + \beta_{17}\text{Uganda}_i + \beta_{18}\text{UAE}_i + \beta_{19}\text{Venezuela}_i + u_i$$

(Equation 6)

To measure the use of *Follow50* by country:

$$\text{Follow50}_i = \beta_0 + \beta_1\text{Armenia}_i + \beta_2\text{Bangladesh}_i + \beta_3\text{Cuba}_i + \beta_4\text{Ecuador}_i + \beta_5\text{Egypt}_i + \beta_6\text{Honduras}_i + \beta_7\text{Indonesia}_i + \beta_8\text{Iran}_i + \beta_9\text{Mexico}_i + \beta_{10}\text{Russia}_i + \beta_{11}\text{SaudiArabia}_i + \beta_{12}\text{Serbia}_i + \beta_{13}\text{Spain}_i + \beta_{14}\text{Tanzania}_i + \beta_{15}\text{Thailand}_i + \beta_{16}\text{Turkey}_i + \beta_{17}\text{Uganda}_i + \beta_{18}\text{UAE}_i + \beta_{19}\text{Venezuela}_i + u_i$$

(Equation 7)

To measure the use of *Follow100* by country:

$$\text{Follow100}_i = \beta_0 + \beta_1\text{Armenia}_i + \beta_2\text{Bangladesh}_i + \beta_3\text{Cuba}_i + \beta_4\text{Ecuador}_i + \beta_5\text{Egypt}_i + \beta_6\text{Honduras}_i + \beta_7\text{Indonesia}_i + \beta_8\text{Iran}_i + \beta_9\text{Mexico}_i + \beta_{10}\text{Russia}_i + \beta_{11}\text{SaudiArabia}_i + \beta_{12}\text{Serbia}_i + \beta_{13}\text{Spain}_i + \beta_{14}\text{Tanzania}_i + \beta_{15}\text{Thailand}_i + \beta_{16}\text{Turkey}_i + \beta_{17}\text{Uganda}_i + \beta_{18}\text{UAE}_i + \beta_{19}\text{Venezuela}_i + u_i$$

(Equation 8)

Results

Equations 3-8 each individually produce an analysis of variance where China is

the reference category and the country coefficients equal how much more/less than China the corresponding country uses a given component. To convert these datapoints to actual means, the country coefficients were added to the constant and are presented in Table A.2. Because adding the mean of each component together ultimately yields the country's overall mean *Investment Score*, the table is arranged from low to high investors.

	Descript. Mean	Location Mean	URL Mean	Fol. 10 Mean	Fol. 50 Mean	Fol. 100 Mean	Investment Mean
China	0.108	0.079	0.004	0.129	0.058	0.045	0.424
Thailand	0.265	0.025	0.000	0.328	0.073	0.012	0.703
Tanzania	0.049	0.369	0.000	0.649	0.149	0.034	1.250
Honduras	0.509	0.176	0.013	0.590	0.168	0.085	1.541
Ecuador	0.244	0.261	0.007	0.666	0.393	0.231	1.802
Serbia	0.218	0.160	0.006	0.888	0.604	0.432	2.309
Spain*	0.665	0.405	0.026	0.719	0.554	0.345	2.714
Indonesia	0.698	0.428	0.067	0.694	0.502	0.415	2.804
Egypt	0.529	0.464	0.036	0.772	0.622	0.463	2.886
Bangladesh	0.667	0.733	0.333	0.800	0.333	0.133	3.000
UAE	0.706	0.459	0.016	0.808	0.613	0.462	3.064
Turkey	0.559	0.391	0.095	0.837	0.725	0.638	3.244
Mexico	0.928	0.199	0.011	0.953	0.717	0.453	3.261
Iran	0.697	0.479	0.063	0.817	0.722	0.626	3.403
Venezuela	0.756	0.506	0.237	0.793	0.593	0.521	3.408
Saudi Arabia	0.737	0.510	0.039	0.885	0.721	0.601	3.493
Uganda	0.888	0.452	0.022	0.933	0.833	0.696	3.823
Russia	0.665	0.705	0.076	0.905	0.822	0.737	3.909
Cuba	0.840	0.475	0.044	0.951	0.833	0.779	3.922
Armenia	1.000	0.914	0.457	0.771	0.600	0.571	4.314
Sample Mean	0.413	0.294	0.034	0.577	0.434	0.354	2.107

Note: Columns are the results of separate OLS regressions where individual country coefficients have been added to the constant (i.e., China) in order to compare actual means rather than statistical significance. Since each component is quantified on a 0-1 scale, each component's mean can be understood as what percentage of a country's trolls possess that particular component and are color-coded in a tripartite schema where red is less one-third of trolls (i.e., 0-.332); yellow is between one-third and two-thirds (i.e., .333-.665); and green is two-thirds or more (i.e., .666-1).

Table A.2: Investment Score Component Means by Country

Takeaways

If variations in overall account investment were being driven by fluctuations in one or two components of the aggregated *Investment Score* variable (i.e., if *Investment Score* was autocorrelated with one or more of its components), then we should encounter at least one monochromatically green column – but we do not. In fact, every column except for *URL* contains two or more green, yellow, and red mean values.

Although it is true that most trolls (57.7%) follow at least 10 other Twitter accounts, it is important to realize that most trolls do not possess any of the other components contained within the *Investment Score* variable. From this it stands to reason that accounts scoring with multiple markers likely correspond with deliberate choices being made by the troll operators at the time of creation (i.e., activating certain profile attributes under certain operational conditions). While the nature of these choices is certainly up for debate (e.g., $RQ_{2.2}$ in Chapter 3), the fact that choices are being made in the curation of these personas should be regarded as somewhat axiomatic given that these accounts do not emerge *ex nihilo* but rather from a malevolent actor that has been tasked to produce a certain effect online.

Lastly, this table also reveals that different countries take different paths to their investment scores. For instance, Bangladesh (3.000) and UAE (3.064) have similar mean *Investment Scores* but they achieve them through deferent choices. Bangladeshi trolls make greater use of *Location* and *URL*, but UAE trolls follow more accounts. Variations such as these present future researchers with opportunities to explore the ways in which *Investment Score* manifests across various state sponsors and how certain components

may or may not relate to what trolls are trying to accomplish (e.g., See Appendix B for an examination of a potential relationship between individual *Investment Score* components and the goal of the information operations the state-sponsored trolls are waging).

APPENDIX B

INVESTMENT SCORE CAMPAIGN GOAL REGRESSIONS

Once again, the novel *Investment Score* dependent variable is an aggregated measure produced by identifying the following attributes about the 16,800 non-Chinese trolls that could be associated with an ESOC information operation:

- The troll persona has a profile description/biographical data (Y/N).
- The troll persona has claimed a geographic location (Y/N).
- The troll persona has provided a personal URL (Y/N).
- The troll persona has at least 10 other Twitter accounts it follows (Y/N).
- The troll persona has at least 50 other Twitter accounts it follows (Y/N).
- The troll persona has at least 100 other Twitter accounts it follows (Y/N).

Each component is quantified based on present (1) versus not-present (0) criteria and then combined to yield an *Investment Score* on a scale from 0-6.

OLS Regressions by Investment Score Component

In order to examine the variance of the individual components relative to *promotional* and *demotional* goals, the six independent OLS regression models were used. To measure the use of *Description* by goal:

$$\text{Description}_i = \beta_0 + \beta_1\text{Promotion}_i + \beta_2\text{Demotion}_i + u_i \quad \text{(Equation 9)}$$

To measure the use of *Location* by goal:

$$\text{Location}_i = \beta_0 + \beta_1\text{Promotion}_i + \beta_2\text{Demotion}_i + u_i \quad \text{(Equation 10)}$$

To measure the use of *Profile URL* by goal:

$$\text{ProfileURL}_i = \beta_0 + \beta_1\text{Promotion}_i + \beta_2\text{Demotion}_i + u_i \quad \text{(Equation 11)}$$

To measure the use of *Follow10* by goal:

$$\text{Follow10}_i = \beta_0 + \beta_1\text{Promotion}_i + \beta_2\text{Demotion}_i + u_i \quad \text{(Equation 12)}$$

To measure the use of *Follow50* by goal:

$$\text{Follow50}_i = \beta_0 + \beta_1\text{Promotion}_i + \beta_2\text{Demotion}_i + u_i \quad \text{(Equation 13)}$$

To measure the use of *Follow100* by goal:

$$\text{Follow100}_i = \beta_0 + \beta_1\text{Promotion}_i + \beta_2\text{Demotion}_i + u_i \quad \text{(Equation 14)}$$

Results

Equations 9-14 each produce an analysis of variance where the intercept is those accounts that could not be sorted into either *promotion* or *demotion* as the reference category and the coefficients equal how much more/less than those accounts an *Investment Score* component predicts whether or not a troll is attempting to *promote* or *demote*. Internal to each column, the green/red designation identifies whether that particular investment score component is a better/worse predictor of being a *demotional* or *promotional* troll. With the exception of URL, there is a statistically significant difference between the *demotion* and *promotion* coefficients for the other five components of the *Investment Score* variable.

	Descript.	Location	URL	Fol. 10	Fol. 50	Fol. 100
(Intercept)	0.655	0.458	0.098	0.977	0.851	0.730
Demotion	0.055	0.024	-0.015	0.006	0.089	0.150
Promotion	0.114	0.098	-0.013	-0.028	-0.047	-0.047
P-Value for Dif. Between Coefficients	<.01	<.01	.6168	<.01	<.01	<.01

Note: Columns are separate OLS regressions where the investment score components are treated as the dependent variable. The Intercept are those trolls unable to be coded as *Promotional* or *Demotional* due to data limitations. These models calculate how much more/less likely a troll a troll is *Promotional* or *Demotional* based on having certain *Investment Score* components (green annotates more likely/red annotates less likely). The P-Values are the linear hypothesis test results that the coefficients are statistically different from each other.

Table B.1: Investment Components Relative to Goal

Takeaways

Just as in Chapter 3, the principal takeaway is that the notion of account *investment* needs to be taken seriously! Account curation is a direct result of choices made by the troll operator and said choices are not made in a vacuum or by a flip of a coin (see Appendix A). The results of Equations 9-14 offer supporting evidence that certain choices may lend themselves to certain goals which would further suggest that operationalizing trolls to conduct information operations may at times include attempts to optimize resource efficiency. If this proves to be true, then it provides all the more reason to adamantly reject one-size-fits-all evaluation strategies for state-sponsored trolls on social media.

REFERENCES

Abbruzzese, Jason, and David Ingram. 2024. "TikTok Says it Disrupted 15 Influence Operations This Year – Including one from China." *NBC News*, May 24. https://www.nbcnews.com/tech/tech-news/tiktok-says-disrupted-15-influence-operations-year-one-china-rcna153831.

Akca, Davut, Suleyman Ozeren, Ismail Onat, Suat Cubukcu. 2021. "Political Astroturfing in Twitterscape: The Role of Troll Armies in Turkey's Democratic Backsliding." Orion Policy Institute, July 28. https://www.orionpolicy.org/research/44/political-astroturfing-in-twitterscape-the-role-of-troll-armies-in-turkeys-dem.

Akers, Shawn. 2014. "'Law', 'Policy', and the 'Middle East': A Brief discussion of Terms" (video of lecture, PPOG 640, Liberty University Helms School of Government). http://www.apple.com/itunes.

Al-Rawi, Ahmed, and Anis Rahman. 2020. "Manufacturing Rage: The Russian Internet Research Agency's Political Astroturfing on Social Media" *First Monday* 25, no. 9 (September). http://dx.doi.org/10.5210/fm.v25i9.10801.

Allison, Graham, and Philip Zelikow. 1999. *Essence of Decision: Explaining the Cuban Missile Crisis*. 2nd ed. New York: Longman.

Anderson, James E. 2015. Public Policymaking: An Introduction. 8th ed. Stamford, CT: Cengage Learning.

Applebaum, Anne. 2014. "How to Fight the Internet's State-Sponsored Trolls." *The Washington Post*, November 30, A.17.

Aristotle. 350 B.C. *Rhetoric (Book I)*. Translated by W. Rhys Roberts. MIT Internet Classics Archive. http://classics.mit.edu/Aristotle/rhetoric.1.i.html.

Baca, Marie. 2019. "Facebook Makes Small Tweaks after Anti-Conservative Bias Report. They're Unlikely to Make the Issue go Away." *The Washington Post*, August 20. https://www.washingtonpost.com/technology/2019/08/20/facebook-makes-small-tweaks-following-anti-conservative-bias-report-theyre-unlikely-make-issue-go-away.

Barnes, Julian E. 2018. "U.S. Assessment of Election Interference Says the Russians Are Still at It." New York Times, Late Edition (East Coast), December 22: A.12.

_____. "Cyber Command Operation Took Down Russian Troll Farm for Midterm Elections." *The New York Times*, February 26. https://www.nytimes.com/2019/02/26/us/politics/us-cyber-command-russia.html.

Bastos, M., D. Marcea, F. Goveia. 2021. "Guy Next Door and Implausibly Attractive Young Women: The Visual Frames of Social Media Propaganda." *New Media & Society* (June 30). https://doi.org/10.1177%2F14614448211026580.

Baum, Jeremy, and John Villasenor. 2023. "The Politics of AI: ChatGPT and Political Bias." *Brookings Institute*, May 8. https://www.brookings.edu/articles/the-politics-of-ai-chatgpt-and-political-bias.

Baumgartner, Frank R., and Bryan D. Jones. 2009. *Agendas and Instability in American Politics*. 2nd ed. Chicago: University of Chicago Press.

BBC. 2021. "Trump-Russia Steele Dossier Analyst Charged with Lying to FBI," *British Broadcasting Corporation*, November 5. https://www.bbc.com/news/world-us-canada-59168626.

Beckerman, Gal. 2015. "Vlad's Got His Eye on You; Snowden's Leaks Have Made Putin's Job Easier by Offering Him the Chance to Argue that Russian Data is Unsafe in American Hands." *Wall Street Journal*, September 9.

Biddle, Sam. 2023. "Those Russian Twitter Bots Didn't Do S#!% in 2016, Says New Study." *The Intercept*, January 10. https://theintercept.com/2023/01/10/russia-twitter-bots-trump-election.

Bills, Christian. 2020. "The Internet Research Agency: Spreading Disinformation." *Small Wars Journal*, October 30. https://smallwarsjournal.com/jrnl/art/internet-research-agency-spreading-disinformation.

Birkland, Thomas. 1997. *After Disaster: Agenda Setting, Public Policy, and Focusing Events*. Washington, D.C.: Georgetown University Press.

_____. 2006. *Lessons of Disaster: Policy Change after Catastrophic Events*. Washington, D.C.: Georgetown University Press.

Bond, Shannon. 2022. "She Joined DHS to Fight Disinformation. She Says She Was Halted by…Disinformation." *NPR*, May 21. https://www.npr.org/2022/05/21/1100438703/dhs-disinformation-board-nina-jankowicz.

Bose, Meena, and Andrew Rudalevige, eds. 2020. *Executive Policymaking: The Role of the OMB in the Presidency*. Washington, D.C.: Brookings Institution Press.

Breuninger, Kevin. 2019. “Robert Mueller’s Russia Probe Cost Nearly $32 Million in Total, Justice Department says.” *CNBC*, August 2. https://www.cnbc.com/2019/08/02/robert-muellers-russia-probe-cost-nearly-32-million-in-total-doj.html.

Broad, William J. 2020. “Putin’s Long War Against American Science.” *The New York Times,* April 14. https://www.nytimes.com/2020/04/13/science/putin-russia-disinformation-health-coronavirus.html.

Bump, Phillip. 2019. “Actually, the Mueller Report Showed that Russia did Affect the Vote.” *The Washington Post*, April 19. https://www.washingtonpost.com/politics/2019/04/19/actually-mueller-report-showed-that-russia-did-affect-vote.

Buttelmann, David, and Robert Böhm. 2014. “The Ontogeny of the Motivation That Underlies In-Group Bias.” *Psychological Science* 25, no. 4 (April): 921–27. https://doi.org/10.1177/0956797613516802.

Carpenter, Ted Galen. 2019. “The Myth That Won’t Die: Donald Trump as Russia’s Puppet.” *CATO Institute*, March 25. https://www.cato.org/commentary/myth-wont-die-donald-trump-russias-puppet.

Cavoli, Christopher G. 2024. “United States House Armed Services Committee Statement of General Christopher G. Cavoli, United States Army United States European Command.” April 10. https://armedservices.house.gov/sites/republicans.armedservices.house.gov/files/USEUCOM%20GEN%20Cavoli%20CPS_HASC_2024.pdf

Cialdini, Robert B. 2006. *Influence: The Psychology of Persuasion,* revised ed. New York: Harper Business Books.

Cima, Lorenzo, Lorenzo Mannocci, Marco Avvenuti, Maurizio Tesconi, and Stefano Cresci. 2024. “Coordinated Behavior in Information Operations on Twitter.” *IEEE Access*, vol. 12 (May): 61568-61585. doi: 10.1109/ACCESS.2024.3393482.

CISA. n.d. “Resilience Series Graphic Novels.” Cybersecurity & Infrastructure Security Agency. https://www.cisa.gov/topics/election-security/foreign-influence-operations-and-disinformation/resilience-series-graphic-novels.

CISA Counter Foreign Influence Task Force. n.d. “Foreign Influence Operations and Disinformation.” Cybersecurity & Infrastructure Security Agency. https://www.cisa.gov/topics/election-security/foreign-influence-operations-and-disinformation.

Clausewitz, Carl von. 1873. *On War*, translated by J.J. Graham. London. https://www.clausewitz.com/readings/OnWar1873/BK1ch01.html.

Clegg, Nick. 2021. “In Response to Oversight Board, Trump Suspended for Two Years; Will Only Be Reinstated if Conditions Permit.” *Meta*, June 4, 2021. https://about.fb.com/news/2021/06/facebook-response-to-oversight-board-recommendations-trump.

CNN. 2020. Intelligence Community's Top Election Official's Statement Warning of Threats from China, Russia and Iran.” *CNN*, August 7. https://www.cnn.com/2020/08/07/politics/ncsc-director-2020-election-threat/index.html.

Cohen, Rachel S. 2019. “16th Air Force Sets Sights on Election Security, Integrated Air Defenses.” *Air & Space Forces Magazine*, October 14. https://www.airandspaceforces.com/16th-air-force-sets-sights-on-election-security-integrated-air-defenses.

Coleman, Keith. 2021. “Introducing Birdwatch, a Community-based Approach to Misinformation.” *Twitter Blog*, January 25. https://blog.twitter.com/en_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation.

Concha, Joe. 2021. “‘Strikingly Sophisticated’ Taliban Thrive on Twitter while Trump Still Banned.” *The Hill*, August 20. https://thehill.com/opinion/technology/568701-strikingly-sophisticated-taliban-thrive-on-twitter-while-trump-still.

Cook, Christine L., Simon Y.-C. Tang, and Jih-Hsuan Tammy Lin. 2023. “Comparing Shades of Darkness: Trolling Victims Experiences on Social Media vs. Online Gaming.” *Frontiers in Psychology* 14 (August). https://doi.org/10.3389/fpsyg.2023.1163244.

Copland, Simon, and Jenny Davis. 2020. “Reddit Removes Millions of pro-Trump Posts. But Advertisers, not Values, Rule the Day.” *The Conversation*, July 1. https://theconversation.com/reddit-removes-millions-of-pro-trump-posts-but-advertisers-not-values-rule-the-day-141703.

Cranmer, Gregory A., Darren Linvill, Hudson Smith, Bryan Denham, Joseph Bober, Kevin Nutt, and William Seaton. 2024. “Social Media Trolls as Faux Third-Party Agents of Image Repair: China’s Disinformation Campaign and Statecraft in the Daryl More Affair.” *Journal of Applied Communication Research* 52, 1 (January): 5-26. https://doi.org/10.1080/00909882.2023.2282508.

Cullison, Alan. 2011. “Web Problems Plague Russia Critics.” *Wall Street Journal*, 2011.

Cyber National Mission Force (CNMF). 2022. “How U.S. Cyber Command, NSA are Defending Midterm Elections: One Team, One Fight.” CNMF Public Affairs Office, August 25. https://www.16af.af.mil/Newsroom/Article/3139869/how-us-cyber-command-nsa-are-defending-midterms-elections-one-team-one-fight.

Delaney, William. n.d. “Julius Caesar: Act II – Scene 1.” Owl Eyes. https://www.owleyes.org/text/julius-caesar/read/act-ii-scene-i#root-71635-108.

DeVine, Michael E.. 2023. “The Intelligence Community’s Foreign Malign Influence Center (FMIC).” Congressional Research Service, August 9. https://crsreports.congress.gov/product/pdf/IF/IF12470.

De Vynnck, Gerrit. 2023. “ChatGPT Leans Liberal, Research Shows.” *Washington Post*, August 16. https://www.washingtonpost.com/technology/2023/08/16/chatgpt-ai-political-bias-research.

Diamond, Jeremy, Jake Tapper, Zachary Cohen. 2020. “US Intelligence Briefer Appears to have Overstated Assessment of 2020 Russian Interference.” *CNN*, February 23. https://www.cnn.com/2020/02/23/politics/intelligence-briefer-russian-interference-trump-sanders/index.html.

Diaz-Plaja, Ruben-Erick, and Joshua Polchar. 2023. “Don’t Fight the Future, Decide It!.” *Nato Review*, December 13. https://www.nato.int/docu/review/articles/2023/12/13/dont-fight-the-future-decide-it/index.html.

DiResta, Renee., C. Miller, V. Molter, J. Pomfret, and G. Tiffert. 2022. “Telling China’s Story: The Chinese Communist Party’s Campaign to Shape Global Narratives.” Stanford Digital Repository, Stanford Internet Observatory, Freeman Spogli Institute for International Studies. https://purl.stanford.edu/pf306sw8941.

DiResta, R., K. Shaffer, B. Ruppel, D. Sullivan, R. Matney, R. Fox, J. Albright, & B. Johnson. 2017. “The Tactics & Tropes of the Internet Research Agency.” *New Knowledge*, October. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1003&context=senatedocs.

Dolan, Eric W. 2019. “Study Finds No Evidence Russian Troll Campaign Impacted Polarization Among U.S. Twitter Users.” *PsyPost*, November 30. https://www.psypost.org/study-finds-no-evidence-russian-troll-campaign-impacted-polarization-among-u-s-twitter-users.

Donath, Judith S. 1999. “Identity and Deception in the Virtual Community.” In *Communities in Cyberspace*, edited by Marc A. Smith and Peter Kollock, 29-59. New York: Routledge.

Downs, Anthony. 1972. “Up and Down with Ecology – The Issue-Attention Cycle.” *The Public Interest* 28 (Summer 1972): 38-50. https://www.nationalaffairs.com /storage/app/uploads/public/58e/1a4/b56/58e1a4b56d25f917699992.pdf.

Duszak, Anna. 2002. *Us and Others : Social Identities Across Languages, Discourses and Cultures*. Amsterdam: John Benjamins Publishing Company.

Dwoskin, Elizabeth, and Tony Romm. 2018a. “Facebook Disrupts New Troll Operation.” *The Washington Post*, August 1: A.1.

_____. 2018b. “Facebook Purges Accounts; Platform Says it Culled Hundreds of Publishers Based on Behavior, not Political Stances.” *Los Angeles Times*, October 12: C.1.

Eady, Gregory, Tom Paskhalis, Jan Zilinsky, Richard Bonneau, Jonathan Nagler, and Josha A. Tucker. 2023. “Exposure to the Russian Internet Research Agency Foreign Influence Campaign on Twitter in the 2016 US Election and its Relationship to Attitudes and Voting Behavior.” *Nature Communications* 14 (January). https://www.nature.com/articles/s41467-022-35576-9#MOESM2.

Elgin, Ben. 2019. “Twitter Revises Data on Russian Trolls and Their 2017 Activity.” *Bloomberg*, February. 20. https://www.bloomberg.com/news/articles/2019-02 -20/twitter-revises-data-on-russian-trolls-and-their-2017-activity#xj4y7vzkg.

Ellick, Adam B., Adam Westbrook, and Jonah M. Kessel. 2018. “Meet the KGB Spies Who Invented Fake News.” *The New York Times*, November 12. https://www .nytimes.com/video/opinion/100000006210828/russia-disinformation-fake -news.html.

Epstein, Robert. 2019. “Why Google Poses a Serious Threat to Democracy, and How to End That Threat.” U.S. Senate Judiciary Subcommittee on the Constitution, Tuesday, June 16, 2:30 p.m.. https://www.judiciary.senate.gov/imo/media/doc /Epstein%20Testimony.pdf.

Fecher, Leland, Tyler Reich, Jack Taylor, and Patrick Warren. 2022. “Oh the Places You’ll Guo: The Tactics and Impact of Chinese Multilingual Narrative Flooding Campaign Through Political Cartoons.” Clemson University Media Forensics Hub, June. https://www.clemson.edu/centers-institutes/watt/hub/documents/ci -guo-influence-operation-2022.html.

Feinberg, Ashely. 2014. “The Birth of the Internet Troll.” *Gizmodo*, October 30. https://gizmodo.com/the-first-internet-troll-1652485292.

Fischer, Sara. 2020. “Fact-Checking Goes Mainstream in Trump Era.” *Axios*, October 13. https://www.axios.com/2020/10/13/fact-checking-trump-media.

Foreign Malign Influence Center. n.d. “How We Work.” Office of the Director of National Intelligence. https://www.dni.gov/index.php/fmic-how-we-work.

_____. n.d. “Overview of the Process for the Executive Branch to Notify the Public and Others Regarding Foreign Malign Influence and Interference Operations Targeting U.S. Elections.” Office of the Director of National Intelligence. https://www.dni.gov/files/FMIC/documents/Overview_of_the_Process_for_the_Executive_Branch_to_Notify_the_Public.pdf.

_____. 2024. “Introduction.” FMI Primer, Volume 1: April. https://www.dni.gov/files/FMIC/documents/products/04-25-24_Report_FMI-Primer-Public-Release.pdf.

Freelon, D., M. Bossetta, C. Wells, J. Lukito, Y. Xia, & K. Adams. 2020. “Black Trolls Matter: Racial and Ideological Asymmetries in Social Media Disinformation.” *Social Science Computer Review* 40, no. 3 (April). https://doi.org/10.1177/0894439320914853.

Frenkel, Sheera, and Tiffany Hsu. 2022. “Facebook, Twitter, and Others Remove Pro-U.S. Influence Campaign.” *The New York Times*, August 24. https://www.nytimes.com/2022/08/24/technology/facebook-twitter-influence-campaign.html.

Funk, Allie. 2023. “Q&A: Section 230 is at the Supreme Court. Here’s Why that Matters for Free Expression.” Freedom House, May 18. https://freedomhouse.org/article/qa-section-230-supreme-court-heres-why-matters-free-expression.

Gadde, Vijaya, and Kayvon Beykpour. 2020. “Additional Steps We’re Taking Ahead of the 2020 US Election.” X Blog, October 9. https://blog.x.com/en_us/topics/company/2020/2020-election-changes.

Galeotti, Mark. 2016. “Beware of Hackers, Not Assassins: Commentary.” *The New York Times*, January 23, A.19.

Gant, Jim. 2014. *One Tribe At A Time: The Paper That Changed the War in Afghanistan*. New York: Black Irish Entertainment LLC.

Geddes, John. 2016. “The Persuasion Triad – Aristotle Still Teaches.” Interaction Design Foundation. https://www.interaction-design.org/literature/article/the-persuasion-triad-aristotle-still-teaches.

Gerasimov, Valery. 2013. “The Value of Science Is in the Foresight: New Challenges Demand Rethinking the Forms and Methods of Carrying out Combat Operations.” In *Military Review*, translated by Robert Coalson, January-February 2016. https://www.armyupress.army.mil/Portals/7/military-review/Archives/English/MilitaryReview_20160228_art008.pdf.

Gilbert, David. 2024. “Why China is So Bad at Disinformation,” *Wired*, April 29. https://www.wired.com/story/china-bad-at-disinformation.

Gitelson, Alan R., Robert L. Dudley, and Melvin J. Dubnick. 2014. “Popular Culture & Politics.” Oxford University Press’s American Government Myths and Realities: 2014 Election Edition. https://global.oup.com/us/companion.websites/9780199374229/stud/ch1/pcp.

Goodnow, Frank J. 1900. *Politics and Administration: A Study in Government*. New York: The MacMillan Company.

Graphika, and Stanford Internet Observatory. 2022. “Unheard Voice: Evaluating Five Years of Pro-Western Covert Influence Operations.” Stanford Internet Observatory Cyber Policy Center, August 24. https://stacks.stanford.edu/file/druid:nj914nx9540/unheard-voice-tt.pdf.

Grossman, Shelby, Emily Tianshi, David Thiel, Renée DiResta. 2022. “My Heart Belongs to Kashmir: An Analysis of a Pro-Indian Army Covert Influence Operation on Twitter.” Stanford Internet Observatory, September 21. https://stacks.stanford.edu/file/druid:zs105tw7107/20220921%20India%20takedown.pdf.

Halevy, Nir, Gary Bornstein, and Lilach Sagiv. 2008. “‘In-Group Love’ and ‘Out-Group Hate’ as Motives for Individual Participation in Intergroup Conflict: A New Game Paradigm.” *Psychological Science* 19, no. 4, (April): 405–11. https://doi.org/10.1111/j.1467-9280.2008.02100.x.

Halperin, Morion H. 1972. “The President and the Military.” *Foreign Affairs*, January. https://www.foreignaffairs.com/articles/1972-01-01/president-and-military.

Harold, Scott W., Harold Beauchamp-Mustafaga, and Jeffrey W. Hornung. 2021. “Chinese Disinformation Efforts on Social Media.” RAND Corporation. https://www.rand.org/content/dam/rand/pubs/research_reports/RR4300/RR4373z3/RAND_RR4373z3.pdf.

Haugh, Timothy, Nicholas Hall, and Eugene Fan. 2020. "16th Air Force and Convergence for the Information War." *The Cyber Defense Review* 5, no. 2 (Summer 2020): 29-43. https://cyberdefensereview.army.mil/The-Journal /Publications/.

Heikkilä, Melissa. 2023. "AI Language Models are Rife with Different Political Biases." *MIT Technology Review*, August 7. https://www.technologyreview.com/2023/08 /07/1077324/ai-language-models-are-rife-with-political-biases.

Holly, James. 2023. "Influence and Perception Management Office." *NSI*, October 3. https://nsiteam.com/smaspeakerseries_03october2023.

Hopp, Toby, Patrick Ferrucci, and Chris Vargo. "Why Do People Share Ideologically Extreme, False, and Misleading Content on Social Media? A Self-Report and Trace Data–Based Analysis of Countermedia Content Dissemination on Facebook and Twitter," *Human Communication Research* 46, no. 4 (October 2020): 357-384, https://doi.org/10.1093/hcr/hqz022.

Horwitz, Jeff, Sam Schechner, and Deepa Seetharaman. 2020. "Facebook Imposes Limits on Election Content, Bans 'Stop the Steal' Group." *The Wall Street Journal*, November 5. https://www.wsj.com/articles/facebook-takes-down-group -organizing-protests-of-vote-counting-11604603908.

Hotz, Robert Lee. 2011. "Decoding Our Chatter; Want to Monitor an Earthquake, Track Political Activity or Predict the Ups and Downs of the Stock Market? Researchers Have Found a Bonanza of Real-Time Data in the Torrential Flow of Twitter Feeds." *Wall Street Journal*, October 1.

Huang, Echo. 2019. "Why China Isn't as Skillful at Disinformation as Russia." *Quartz*, September 19. https://qz.com/1699144/why-chinas-social-media-propaganda-isnt -as-good-as-russias.

Hundley, Parker, Anthony Jones, Darren Linvill, Lian Norris, Helen Schmidt, and Jayson Warren. 2022. "Spicy Memes from Spicy Panda." Clemson University Media Forensics Hub, November. https://www.clemson.edu/centers-institutes/watt/hub /images/spicy-memes-spicy-panda.pdf.

Huntington, Samuel P. 1996. *The Clash of Civilizations and the Remaking of World Order*. New York: Simon & Schuster.

Irving, Doug. 2023. "Truth Decay is Putting U.S. National Security at Risk." RAND Corporation, June 28. https://www.rand.org/pubs/articles/2023/truth-decay-is -putting-us-national-security-at-risk.html.

Isaac, Mike, and Sheera Frenkel. 2018."Facebook Takes Down Iran-Based Network Amid Disinformation." *New York Times*, Late Edition (East Coast), October 27: B.5.

Jacques, Ingrid. 2024. "Trump Right About Hunter's 'Laptop from Hell,' Though Biden Claimed Russian Disinformation." *USA Today*, June 6. https://www.usatoday.com /story/opinion/columnist/2024/06/06/hunter-biden-trial-laptop-trump /73982808007.

Jamison, Amelia M., David A. Broniatowski, and Sandra Crouse Quinn. 2019. "Malicious Actors on Twitter: A Guide for Public Health Researchers." *American Journal of Public Health*, April 10. https://ajph-aphapublications-org.libproxy .clemson.edu/doi/full/10.2105/AJPH.2019.304969.

Jensen, Benjamin, Brandon Valeriano, & Ryan Maness. 2020. "Fancy Bears and Digital Trolls: Cyber Strategy with a Russian Twist." In *Military Strategy in the 21st Century*, edited by Kersti Larsdotter, 58-80. England: Routledge.

Jonsson, Oscar. 2019. *The Russian Understanding of War: Blurring the Lines Between War and Peace*. Washington, D.C.: Georgetown University Press.

Keller, Franziska, David Schoch, Sebastian Stier, and JungHwan Yang. 2017. "How to Manipulate Social Media: Analyzing Political Astroturfing Using Ground Truth Data from South Korea." *Proceedings of the International AAAI Conference on Web and Social Media* 11 no. 1 (May): 564-567. http://dx.doi.org/10.1609/icwsm .v11i1.14941.

Kingdon, John. 2011. *Agendas, Alternatives, and Public Policies*, 2nd ed. Glenview, IL: Longman.

Klein, Ofra. 2020. "Does Censoring the Radical Right on Social Media Work?" Open Democracy, December 7. https://www.opendemocracy.net/en/global-extremes /does-censoring-radical-right-social-media-work.

Klippenstein, Ken. 2023a. "The Government Created a New Disinformation Office to Oversee All the Other Ones." *The Intercept*, May 5. https://theintercept.com/2023 /05/05/foreign-malign-influence-center-disinformation.

_____. 2023b. "Inside the Pentagon's New 'Perception Management' Office to Counter Disinformation." *The Intercept*, May 17. https://theintercept.com/2023/05/17 /pentagon-perception-management-office.

Krutka, Daniel G., & Spencer P. Greenhalgh. 2021. "You Can Tell a Lot About a Person by Reading Their Bio: Lessons from Inauthentic Twitter Accounts' Activity in #Edchat." *Journal of Research on Technology in Education* (August). https://doi.org/10.1080/15391523.2021.1962454.

Kurlantzick, Joshua. 2020. "How China Ramped Up Disinformation Efforts During the Pandemic." Council on Foreign Relations, September 10. https://www.cfr.org/in-brief/how-china-ramped-disinformation-efforts-during-pandemic.

Lapowsky, Issie. 2018. "Iran's New Facebook Trolls Are Using Russia's Playbook." October 26. https://www.wired.com/story/iran-facebook-trolls-using-russia-playbook/.

Lawson, Sean. 2019. "What if Russian Disinformation Isn't As Effective As We Thought?" *Forbes*, December 6. https://www.forbes.com/sites/seanlawson/2019/12/06/what-if-russian-disinformation-isnt-as-effective-as-we-thought.

Legal Information Institute. n.d. "47 U.S. Code § 230 - Protection for Private Blocking and Screening of Offensive Material." Cornell Law School Legal Information Institute. https://www.law.cornell.edu/uscode/text/47/230.

Linvill, Darren, and Patrick Warren. 2019a. "Russian Trolls Can Be Surprisingly Subtle, and Often Fun to Read." *The Washington Post*, March 8. https://www.washingtonpost.com/outlook/russian-trolls-can-be-surprisingly-subtle-and-often-fun-to-read/2019/03/08/677f8ec2-413c-11e9-9361-301ffb5bd5e6_story.html.

_____. 2019b. "That Uplifting Tweet You Just Shared? A Russian Troll Sent It." *Rolling Stone*, November 25. https://www.rollingstone.com/politics/politics-features/russia-troll-2020-election-interference-twitter-916482.

_____. 2020. "Troll Factories: Manufacturing Specialized Disinformation on Twitter." *Political Communication* 37, no. 4 (February): 447–467. https://doi.org/10.1080/10584609.2020.1718257.

_____. 2021a. "The Real Target of Authoritarian Disinformation Autocrats Care More About Domestic Control Than Influence Abroad." *Foreign Affairs*, March 24. https://www.foreignaffairs.com/articles/russian-federation/2021-03-24/real-target-authoritarian-disinformation.

_____. 2021b. "Understanding the Pro-China Propaganda and Disinformation Tool Set in Xinjiang." *Lawfare Blog*, December 1. https://www.lawfareblog.com/understanding-pro-china-propaganda-and-disinformation-tool-set-xinjiang.

_____. 2025. “Paths to Influence: How Coordinated Influence Operations Affect the Prominence of Ideas.” In *Corporate Cancel Culture and Brand Boycotts: The Dark Side of Social Media for Brands*, edited by Angeline Close Scheinbaum, pp. forthcoming. UK: Routledge/Psychology Press.

Linvill, Darren, Jayson Warren, Patrick Warren, and David White. 2024. “Where Do Trolls Say They Are?: Understanding the Use of Place Claims in Nation State Influence Operations.” *International Journal of Public Opinion Research* 36, no. 3 (Autumn). https://doi.org/10.1093/ijpor/edae022.

Linvill, Darren, Patrick Warren, Steven Sheffield, Jayson Warren, Beau Brierre, Grant Cole, Jonathan Heijjer, Tyler Reich, Grant Saunders, and Jack Taylor. 2021. “The Xinjiang Nylon Influence Operation.” Clemson University Media Forensics Hub, November. https://www.clemson.edu/centers-institutes/watt/hub/images/ci-xinjiang-nylon.pdf.

Linvill, Darren, Patrick Warren, and David White. 2022. “Russian Social Media Promotion of Sputnik V in Latin America.” Carnegie Endowment for International Peace, March. https://carnegieendowment.org/files/Linvill_Warren_Sputnik_V_Latin_America_final_1.pdf.

Lutz, Eric. 2020. “Alexander Vindman Confirms Trump is 100% a Putin Puppet.” *Vanity Fair*, September 14. https://www.vanityfair.com/news/2020/09/alexander-vindman-confirms-trump-is-100-a-putin-puppet.

Macagnone, Michael. 2019. “Facebook, Other Social Media Sites Pressured to Protect Census.” *Roll Call*, November 8. https://rollcall.com/2019/11/08/facebook-other-social-media-sites-pressured-to-protect-census.

Malloch-Brown, Mark. 2023. “How to Rebuild Trust in Institutions: Results, Results, Results.” World Economic Forum, December 18. https://www.weforum.org/agenda/2023/12/how-to-rebuild-trust-in-philanthropy-results-results-results.

Mangan, Dan. 2020. “Postal Service Data Shows Poor Mail-in Ballot Delivery Rate in Key Swing States, Judge Suggests Postmaster General DeJoy Might Have to Testify.” *CNBC*, November 6. https://www.cnbc.com/2020/11/04/2020-presidential-election-postal-data-shows-ballot-delivery-rate.html.

Markman, Art. 2019. “Comparison is Crucial for Explanation.” *Psychology Today*, January 24. https://www.psychologytoday.com/us/blog/ulterior-motives/201901/comparison-is-crucial-explanation.

Marshall, Lisa. 2020. “Who Shares the Most Fake News? New Study Sheds Light.” *CU Boulder Today*, June 17. https://www.colorado.edu/today/2020/06/17/who-shares-most-fake-news-new-study-sheds-light.

Martin, A. D., Shapiro, N. J., Ilhardt, J. 2020/2023. “Trends in Online Influence Efforts, v4.” Empirical Studies of Conflict, Princeton University. https://esoc.princeton.edu/publications/trends-online-influence-efforts.

Matishak, Martin. 2022. “One-on-One With the Air Force’s Cyber Chief.” *The Record*, April 17. https://therecord.media/one-on-one-with-the-air-forces-cyber-chief.

Mello Jr., John P. 2022. “A Third of US Social Media Users Creating Fake Accounts.” *Tech News World*, August 10. https://www.technewsworld.com/story/a-third-of-us-social-media-users-creating-fake-accounts-176987.html.

Mikkelson, David. 2018. “Did CNN Purchase an Industrial-Sized Washing Machine to Spin News?”. Snopes. March 1. https://www.snopes.com/fact-check/cnn-washing-machine.

Morrison, Sara. 2023. “Section 230, the Internet Law That’s Under Threat, Explained.” *Vox*, February 23. https://www.vox.com/recode/2020/5/28/21273241/section-230-explained-supreme-court-social-media.

Mueller, Robert S. 2019. “Report on the Investigation into Russian Interference in the 2016 Presidential Election.” US Department of Justice, March. https://www.justice.gov/archives/sco/file/1373816/dl

Nakashima, Ellen. 2018. “White House Authorizes ‘Offensive Cyber Operations’ to Deter Foreign Adversaries.” *The Washington Post*, September 20. https://www.washingtonpost.com/world/national-security/trump-authorizes-offensive-cyber-operations-to-deter-foreign-adversaries-bolton-says/2018/09/20/b5880578-bd0b-11e8-b7d2-0773aa1e33da_story.html.

National Intelligence Council. 2021. “Foreign Threats to the 2020 US Federal Elections.” Office of the Director of National Intelligence, March 10. https://www.dni.gov/files/ODNI/documents/assessments/ICA-declass-16MAR21.pdf.

NDI. “Disinformation, Social Media, and Electoral Integrity.” National Democratic Institute, n.d. https://www.ndi.org/disinformation-social-media-and-electoral-integrity.

New York Times Editorial Board. “Joe Biden: Former Vice President of the United States.” *The New York Times*, January 17. https://www.nytimes.com/interactive/2020/01/17/opinion/joe-biden-nytimes-interview.html.

Ney, Paul C. Jr. 2020. “DOD General Counsel Remarks at U.S. Cyber Command Legal Conference.” U.S. Department of Defense, March 2. https://www.defense.gov/News/Speeches/Speech/Article/2099378/dod-general-counsel-remarks-at-us-cyber-command-legal-conference.

Nimmo, Ben, Camille François, C. Shawn Eib, and Léa Ronzaud. 2020. “Spamoflauge Dragon Goes to America: Pro-Chinese Inauthentic Network Debuts English-Language Videos.” *Graphika*, August. https://public-assets.graphika.com/reports/graphika_report_spamouflage_dragon_goes_to_america.pdf.

Nimmo, Ben, and Eric Huchins. 2023. “Phase-Based Tactical Analysis of Online Operations.” Carnegie Endowment for International Peace. March 16. https://carnegieendowment.org/2023/03/16/phase-based-tactical-analysis-of-online-operations-pub-89275.

Nunberg, Geoff. 2019. ‘Disinformation’ is the Word of the Year – And a Sign of What’s to come. *NPR*, December 30. https://www.npr.org/2019/12/30/790144099/disinformation-is-the-word-of-the-year-and-a-sign-of-what-s-to-come.

Nye, Joseph S. 2019. “Protecting Democracy in an Era of Cyber Information War.” Belfer Center for Science and International Affairs, Harvard Kennedy School, February. https://www.belfercenter.org/sites/default/files/files/publication/ProtectingDemocracy.pdf.

Office of the Director of National Intelligence. n.d. “Objectivity.” Office of the Director of National Intelligence. https://www.dni.gov/index.php/how-we-work/objectivity.

Oremus, Will. 2019. “Russian Trolls Aren’t Actually Persuading Americans on Twitter, Study Finds.” *Medium*, November 25. https://onezero.medium.com/russian-trolls-arent-actually-persuading-americans-on-twitter-study-finds-d8fd6bcacaba

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action.* New York: Cambridge University Press.

Oxford English Dictionary, s.v. “troll (v.), additional sense,” March 2024, https://doi.org/10.1093/OED/3681985253.

Pew Research Center. 2022. “The Role of Alternative Social Media in the News and Information Environment.” Pew Research Center, October 6. https://www.courthousenews.com/wp-content/uploads/2022/10/pew-alternative-social-media-study.pdf.

Polantz, Katelyn. 2019. “Mueller Investigation Cost $32 Million, Justice Department Says.” *CNN*, August 2. https://www.cnn.com/2019/08/02/politics/mueller-report-cost/index.html.

Polyakova, Alina. 2020. “The Global Engagement Center: Leading the United States Government’s Fight Against Global Disinformation Threat.” Hearing before the U.S. Senate Subcommittee on State Department and USAID Management, International Operations, and Bilateral International Development of the Committee on Foreign Relations, 116th Congress, March 5. https://www.govinfo.gov/content/pkg/CHRG-116shrg41862/html/CHRG-116shrg41862.htm.

Porpitakpan, C. 2006. “The Persuasiveness of Source Credibility: A Critical Review of Five Decades’ Evidence.” *Journal of Applied Social Psychology* 34, no. 2 (July): 243–281. https://doi.org/10.1111/j.1559-1816.2004.tb02547.x.

Posetti, Julie, and Alice Matthews. 2018. “A Short Guide to the History of ’Fake News’ and Disinformation: A Learning Module for Journalists and Journalism Educators.” International Center for Journalists, July. https://www.icfj.org/sites/default/files/2018-07/A%20Short%20Guide%20to%20History%20of%20Fake%20News%20and%20Disinformation_ICFJ%20Final.pdf.

Potter, Rachel Augustine. 2019. *Bending the Rules: Procedural Politicking in the Bureaucracy*. Chicago, IL: University of Chicago Press.

Pressman, Jeffrey L., and Aaron B. Wildavsky. 1979. *Implementation: How Great Expectations in Washington Are Dashed in Oakland: or, Why It’s Amazing That Federal Programs Work at All, This Being a Saga of the Economic Development Administration as Told by Two Sympathetic Observers Who Seek to Build Morals on a Foundation of Ruined Hopes*. 2nd ed. Berkeley: University of California Press.

Rapp, Christof. 2022. “Aristotle’s Rhetoric.” Stanford Encyclopedia of Philosophy, March 15. https://plato.stanford.edu/entries/aristotle-rhetoric.

Rao, Naomi. 2011. “Public Choice and International Law Compliance: The Executive Branch is a ‘They,’ Not an ‘It.’” *Minnesota Law Review* 96, no. 1 (November 2011): 194-277. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1984224.

Rid, Thomas. 2020. *Active Measures: The Secret History of Disinformation and Political Warfare*. New York: Farrar, Straus and Giroux.

Robertson, Katie. 2024. “Hunter Biden’s Laptop, Revealed by New York Post, Comes Back to Haunt Him.” *The New York Times*, June 11. https://www.nytimes.com/2024/06/11/business/media/hunter-biden-laptop-new-york-post.html.

Rosen, Guy. 2020. “Preparing for Election Day.” Meta, October 7. https://about.fb.com/news/2020/10/preparing-for-election-day.

Rowling, J.K. 2001. *Fantastic Beasts & Where to Find Them*. New York: Scholastic.

Rudalevige, Andrew. 2021. *By Executive Order: Bureaucratic Management and the Limits of Presidential Power*. Princeton, New Jersey: Princeton University Press.

Ryan, Fergus. 2023. “Can TikTok Alone Tackle CCP-linked Information Ops?.” *Australian Strategic Policy Institute*, September 18. https://www.aspistrategist.org.au/can-tiktok-alone-tackle-ccp-linked-information-ops.

Saad, Lydia. 2023. “Historically Low Faith in U.S. Institutions Continues.” Gallup, July 6. https://news.gallup.com/poll/508169/historically-low-faith-institutions-continues.aspx.

Sarno, Dawn, Patrick Warren, Jeff Black, Jayson Warren, and Jack Taylor. 2024. “Generalized Inauthenticity: A Scoping Study of Phishing, Fake News, and Trolls.” Clemson University Media Forensics Hub Working Paper.

Sanger, David E. and Steven Lee Myers. 2023. “China Sows Disinformation About Hawaii Fires Using New Techniques.” *New York Times,* September 12. https://www.nytimes.com/2023/09/11/us/politics/china-disinformation-ai.html.

Satariano, Adam. 2021. “Inside a Pro-Huawei Influence Campaign.” *The New York Times*, January 29. https://www.nytimes.com/2021/01/29/technology/commercial-disinformation-huawei-belgium.html.

Sbraccia, Steve. 2022. “Be Aware of Fake Social Media Accounts: More than Billion Were Ousted in 2021.” *CBS 17*, January 10, 2022. https://www.cbs17.com/news/investigators/be-aware-of-fake-social-media-accounts-more-than-1-billion-were-ousted-in-2021/.

Schattschneider, E.E. 1935. *Politics, Pressures, and the Tariff*. New York: Atherton.

_____. 1960. *The Semi-Sovereign People: A Realist’s View of Democracy in America*. New York: Holt, Rinehart and Winston.

Scola, Nancy. 2020. “Social Media Has Met the Enemy, and It’s Us.” *Politico*, November 3. https://www.politico.com/news/2020/11/03/social-media-misinformation-during-election-433906.

Scott, Mark. 2024. “China’s Kremlin-Style Disinformation Playbook.” *Politico*, January 11. https://www.politico.eu/newsletter/digital-bridge/chinas-kremlin-style-disinformation-playbook.

Shkabatur, Jennifer. 2019. “The Global Commons of Data,” *Stanford Technical Law Review* 22: 354-411. https://law.stanford.edu/wp-content/uploads/2019/09/Shkabatur_Global-Commons_20190830-1.pdf.

Silverman, Craig, and Jeff Kao. 2022. “In the Ukraine Conflict, Fake Fact-Checks Are Being Used to Spread Disinformation.” *ProPublica*, March 8. https://www.propublica.org/article/in-the-ukraine-conflict-fake-fact-checks-are-being-used-to-spread-disinformation.

Skowronski, Jason. 2019. “Trolls and Bots are Disrupting Social Media — Here’s how AI can Stop Them (Part 1).” *Towards Data Science*, July 23, 2019. https://towardsdatascience.com/trolls-and-bots-are-disrupting-social-media-heres-how-ai-can-stop-them-d9b969336a06.

Smith, Tina. 2022. “Klobuchar, Bennet, Slotkin Introduce Bicameral Legislation to Strengthen Media Literacy Education and Improve Personal Cybersecurity.” Office of the U.S. Senator for Minnesota Tina Smith, July 8. https://www.smith.senate.gov/klobuchar-bennet-slotkin-introduce-bicameral-legislation-to-strengthen-media-literacy-education-and-improve-personal-cybersecurity.

Solomon. n.d. Ecclesiastes 1:9. English Standard Version.

Stone, Jeff. 2020. “Chinese Accounts Blast Trump, with Help from AI-Generated Pictures.” *Cyberscoop*, August 13. https://cyberscoop.com/graphika-spamouflage-dragon-china.

Strause, Jackie. 2017. “‘Quantico’ Boss Explains the Show’s Full-Circle Season 2 Finale.” *The Hollywood Reporter*, May 15. https://www.hollywoodreporter.com/tv/tv-news/quantico-season-2-finale-explained-1004045.

Tabatabai, Ariane M. 2018. “A Brief History of Iranian Fake News: How Disinformation Campaigns Shaped the Islamic Republic.” *Foreign Affairs*, August 24. https://www.foreignaffairs.com/articles/middle-east/2018-08-24/brief-history-iranian-fake-news.

Taylor, Jessica. 2017. “Clinton Says She Was ‘Right” About ‘Vast Russia Conspiracy’; Investigations Ongoing.” *NPR*, June 1. https://www.npr.org/2017/06/01/530941011/clinton-says-she-was-right-about-vast-russia-conspiracy-investigations-ongoing.

Thompson, Marcelo. 2020. “Beyond Gatekeeping: The Normative Responsibility of Internet Intermediaries.” *Vanderbilt Journal of Entertainment and Technology Law* 18, no. 4: 783-848. https://scholarship.law.vanderbilt.edu/jetlaw/vol18/iss4/4.

Thucydides. n.d. *The History of the Peloponnesian War*, translated by Richard Crawley. Project Gutenburg. https://www.gutenberg.org/files/7142/7142-h/7142-h.htm.

Timberg, Craig, and Tony Romm. 2019. “Iranians, Others Take Russian Cue in Election Disinformation Efforts.” *The Washington Post*, July 26: A.1.

Troianovski, Anton. 2018. “A Former Russian Troll Speaks: ‘It Was Like Being in Orwell’s World.’” *The Washington Post*, February 17. https://www.washingtonpost.com/news/worldviews/wp/2018/02/17/a-former-russian-troll-speaks-it-was-like-being-in-orwells-world.

Trump, Donald J. 2020. “Executive Order on Preventing Online Censorship.” US Executive Office of the President, May 28. https://trumpwhitehouse.archives.gov/presidential-actions/executive-order-preventing-online-censorship.

Twitter. n.d. “Twitter Moderation Research Consortium.” Twitter Transparency Center. https://transparency.twitter.com/en/reports/moderation-research.html.

_____. n.d. @TwitterSafety. https://twitter.com/TwitterSafety.

US Congress. Senate. Committee on Armed Services. 2016. “National Defense Authorization Act for Fiscal Year 2017.” S.2943. 114th Congress. https://www.congress.gov/bill/114th-congress/senate-bill/2943.

_____. House. 2018. “National Defense Authorization Act for Fiscal Year 2019.” H.R.5515. 115th Congress. https://www.congress.gov/115/bills/hr5515/BILLS-115hr5515enr.pdf.

_____. Senate. Select Committee on Intelligence. 2020. “Report of the Select Committee on Intelligence United States Senate on Russian Active Measures Campaigns and Interference in the 2016 U.S. Election.” 116th Congress. https://www.intelligence.senate.gov/publications/report-select-committee-intelligence-united-states-senate-russian-active-measures.

_____. House. 2023. “The Weaponization of ‘Disinformation’ Pseudo-Experts and Bureaucrats: How the Federal Government Partnered with Universities to Censor Americans’ Political Speech.” Interim Staff Report of the Committee on the Judiciary and the Select Subcommittee on the Weaponization of the Federal Government, November 6. https://judiciary.house.gov/sites/evo-subsites/republicans-judiciary.house.gov/files/evo-media-document/EIP_Jira-Ticket-Staff-Report-11-7-23-Clean.pdf.

US Cyberspace Solarium Commission. 2020. March. https://www.solarium.gov.

US Department of Defense. 2020. “DOD Has Enduring Role in Election Defense.” DOD Public Affairs, February 10. https://www.defense.gov/News/News-Stories/Article/Article/2078716/dod-has-enduring-role-in-election-defense.

US Department of Homeland Security. 2022. “DHS Needs a Unified Strategy to Counter Disinformation Campaigns.” Office of the Inspector General, August 10. https://www.oig.dhs.gov/sites/default/files/assets/2022-08/OIG-22-58-Aug22.pdf.

US Department of Justice. 2018. “Grand Jury Indicts Thirteen Russian Individuals and Three Russian Companies for Scheme to Interfere in the United States Political System.” Office of Public Affairs, February 16. https://www.justice.gov/opa/pr/grand-jury-indicts-thirteen-russian-individuals-and-three-russian-companies-scheme-interfere.

_____. 2021. The United States Attorney’s Office: Southern District of New York. “U.S. Attorney Announces Charges Against Two Iranian Nationals for Cyber-Enabled Disinformation and Threat Campaign Designed to Interfere with the 2020 U.S. Presidential Election.” November 18. https://www.justice.gov/usao-sdny/pr/us-attorney-announces-charges-against-two-iranian-nationals-cyber-enabled.

_____. 2023. “40 Officers of China’s National Police Charged in Transnational Repression Schemes Targeting U.S. Residents.” Office of Public Affairs, April 17. https://www.justice.gov/opa/pr/40-officers-china-s-national-police-charged-transnational-repression-schemes-targeting-us

US Department of State. 2016. “Global Engagement Center [Archived Content].” Under Secretary for Public Diplomacy https://2009-2017.state.gov/r/gec/#:~:text=The%20Global%20Engagement%20Center%20is,Obama%20on%20March%2014%2C%202016

_____. 2022. Global Engagement Center. “PRC Efforts to Manipulate Global Public Opinion on Xinjiang.” August 24. https://www.state.gov/prc-efforts-to-manipulate-global-public-opinion-on-xinjiang.

_____. n.d. “Global Engagement Center,” Under Secretary for Public Diplomacy. https://www.state.gov/bureaus-offices/under-secretary-for-public-diplomacy-and-public-affairs/global-engagement-center.

US Department of the Treasury. 2021. “Treasury Escalates Sanctions Against the Russian Government’s Attempts to Influence U.S. Elections.” U.S. Department of the Treasury, April 15. https://home.treasury.gov/news/press-releases/jy0126.

US Executive Office of the President. 2017. National Security Council. *National Security Strategy*, December. https://trumpwhitehouse.archives.gov/wp-content/uploads/2017/12/NSS-Final-12-18-2017-0905.pdf, 27.

Valeriano, Brandon, Benhamin Jensen, and Ryan C. Maness. 2020. *Cyber Strategy: The Evolving Character of Power and Coercion*. New York: Oxford University Press.

Villasenor, Maya. 2020. “Revisiting Section 230: The Implications are International.” *Council on Foreign Relations*, November 16. https://www.cfr.org/blog/revisiting-section-230-implications-are-international.

Volz, Dustin. 2021. “Pro-China Online Network Used Fake Accounts to Urge Asian-Americans to Attend Protests, Researchers Say.” *The Wall Street Journal*, September 8. https://www.wsj.com/articles/pro-china-online-network-used-fake-accounts-to-urge-asian-americans-to-attend-protests-researchers-say-11631109601.

Wang, Amy B. 2019. “‘Like I Said: A Puppet’: Hillary Clinton Doubles Down on Trump and Russia.” *The Washington Post*, January 14. https://www.washingtonpost.com/politics/2019/01/14/like-i-said-puppet-hillary-clinton-doubles-down-trump-russia.

Wanless, Alicia, and Laura Walters. 2020. “How Journalists Become an Unwitting Cog in the Influence Machine.” *Carnegie Endowment for International Peace*, October 13. https://carnegieendowment.org/2020/10/13/how-journalists-become-unwitting-cog-in-influence-machine-pub-82923.

Warren, Jayson. 2017. *The Failure of Westphalia: A Constructivist Examination of Western and Middle Eastern Relations*. Scotts Valley, CA: CreateSpace.

_____. 2020a. “If A2/AD Will Blot Out the Sun…Then We Will Fight in the Shade: 300 Spartans and Information Warfare in the Twenty-first Century,” *Wild Blue Yonder* (June): 100-105. https://www.airuniversity.af.edu/Wild-Blue-Yonder/Article-Display/Article/2208624/if-a2ad-will-blot-out-the-sun-then-we-will-fight-in-the-shade-300-spartans-and.

_____. 2020b. “Not All Wars are Violent: Identifying Faulty Assumptions for the Information War,” *Air and Space Power Journal* 34, no. 4 (Winter 2020): 76-90. https://www.airuniversity.af.edu/Portals/10/ASPJ/journals/Volume-34_Issue-4/F-Warren.pdf.

_____. 2022. “Informed Consent: The Marketplace of Ideas as an Imperative for Free and Open Elections.” The Hoover Institution, November. https://www.hoover.org/hoover-institution-summer-policy-boot-camp-directors-award-0.

_____. 2024. “On Competition: A Continuation of Policy by Misunderstood Means.” In *Great Power Cyber Competition: Competing and Winning in the Information Environment*, edited by David V Gioe and Margaret W. Smith, 25-43. New York: Routledge.

Warren, Jayson, Darren Linvill, and Patrick Warren. 2023. “Custer’s Last Tweet: Avoiding a Digital Little Bighorn in the Fight for Hearts and Minds.” Irregular Warfare Initiative, April 5. https://irregularwarfare.org/articles/custers-last-tweet-avoiding-a-digital-little-bighorn-in-the-fight-for-hearts-and-minds.

Warren, Patrick, Darren Linvill, Leland Fecher, Jayson Warren, Steven Sheffield, Jack Taylor, Alexa Gubanich, Parker Hundley, Josiah Lamont, Sarah Meadows, Josephine Rohrer, Molly Sutton, and Shay Easler. 2023. “The 5-Year Spam: Tracking a Persistent Chinese Influence Operation.” Clemson University Media Forensics Hub, February. https://www.clemson.edu/centers-institutes/watt/hub/images/5-year-spam.

Wells, Georgia, and Liza Lin. 2022. “Pro-China Twitter Accounts Flood Hashtag Critical of Beijing Winter Olympics.” *The Wall Street Journal*, February 8. https://www.wsj.com/articles/pro-china-twitter-accounts-flood-hashtag-critical-of-beijing-winter-olympics-11644343870

West, Darrel M. 2011. “Ten Ways Social Media Can Improve Campaign Engagement and Reinvigorate American Democracy.” *Brookings Institution*, June 28. https://www.brookings.edu/articles/ten-ways-social-media-can-improve-campaign-engagement-and-reinvigorate-american-democracy.

Wexton, Jennifer. n.d. “Congressional Task Force on Digital Citizenship.” US House of Representatives, Virginia’s 10th District. https://wexton.house.gov/about/congressional-task-force-on-digital-citizenship.htm.

Wild, Johanna. 2022. “This New Tool Lets you Analyse TikTok Hashtags.” Bellingcat, May 11. https://www.bellingcat.com/resources/how-tos/2022/05/11/this-new-tool-lets-you-analyse-tiktok-hashtags.

Wilson, James Q. 1973. *Political Organizations*. New York: Basic Books.
_____. 1989. *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books, Inc.

Wilson, Woodrow. 1887. "The Study of Administration," *Political Science Quarterly* 2, no. 2 (June): 197-222. https://www.jstor.org/stable/2139277.

Wood, B. Dan, and Richard W. Waterman. 1994. *Bureaucratic Dynamics: The Role of the Bureaucracy in a Democracy*. Boulder, CO: Westview Press, Inc.

WSJ. "Zuckerberg and Hunter Biden's Laptop." *The Wall Street Journal,* August 26, 2022. http://www.wsj.com/articles/mark-zuckerberg-and-hunter-bidens-laptop-joe-rogan-russian-disinformation-fbi-11661544602.

Zannettou, Savvas, Tristan Caulfield, William Setzer, Michael Sirivianos, Gianluca Stringhini, & Jeremy Blackburn. 2019. "Who Let The Trolls Out?: Towards Understanding State-Sponsored Trolls." *WebSci* 19 (June): 353-362. https://www.researchgate.net/publication/334152827_Who_Let_The_Trolls_Out_Towards_Understanding_State-Sponsored_Trolls.

Zorn, Eric. 2019. "What Are you Going to Do? A Russian Troll Just Won My Tweet of the Week Poll." *Chicago Tribune*, August 2. https://www.chicagotribune.com/columns/eric-zorn/ct-column-twitter-russian-trolls-tweet-poll-zorn-20190802-2wc6lyhibbdk5dcyv3lgodxory-story.html.