Georgetown Security Studies Review
Volume 11 | Issue 2
February 2024

An official publication of the Center for Security Studies at
Georgetown University’s Edmund A. Walsh School of Foreign Service

https://georgetownsecuritystudiesreview.org

Disclaimer
The views expressed in Georgetown Security Studies Review do not necessarily represent those of the editors or staff of GSSR, the Edmund A. Walsh School of Foreign Service, or Georgetown University. The editorial board of GSSR and its affiliated peer reviewers strive to verify the accuracy of all information contained in GSSR. However, the staffs of GSSR, the Edmund A. Walsh School of Foreign Service, and Georgetown University make no warranties or representations regarding the completeness or accuracy of information contained in GSSR, and they assume no legal liability or responsibility for the content of any work contained therein.

Copyright 2012–2024, Georgetown Security Studies Review. All rights reserved.
GEORGETOWN SECURITY STUDIES REVIEW

Published by the Center for Security Studies
at Georgetown University’s Edmund A. Walsh School of Foreign Service

EDITORIAL STAFF
Ona Spreenberg, Editor-in-Chief
Christopher White, Deputy Editor
Arjan van Tongerlo, AE for Europe & Central Asia
Anna Matilde Bassoli, AE for the Indo-Pacific
Katherine Wells, AE for the Middle East & North Africa
Zoë Kern, AE for Sub-Saharan Africa
Lauren Munster, AE for the Western Hemisphere
Mano Harada, AE for Environment & Health
Yebin Won, AE for Gender & Identity
Christian Trotti, AE for Defense
Thekla Ketcher, AE for Race & Security
Alexander Chang, AE for Technology
Becca Ebert, AE for Terrorism & Transnational Threats
Xiang Chi, AE for Geoeconomics and Resource Competition

PEER REVIEW STAFF

James McGee, Senior Peer Reviewer
Logan Duarte, Senior Peer Reviewer
Samanvya Hooda, Senior Peer Reviewer
Aaron Mikell, Senior Peer Reviewer
Samantha Olson, Peer Reviewer
Dylan Gooding, Peer Reviewer
Ryan Wisowaty, Peer Reviewer
Natalie Zakarian, Peer Reviewer
Rudy Antoncic, Peer Reviewer
Leo McMahon, Peer Reviewer
Bree Megivern, Peer Reviewer
Keven Hernandez, Peer Reviewer
Tyler Koteskey, Peer Reviewer
Lilli Abraham, Peer Reviewer
Apurva Ramakrishnan, Peer Reviewer
Evan Hydock, Peer Reviewer
Alexandra Seymour, Peer Reviewer

The Georgetown Security Studies Review is the official academic journal of Georgetown University’s Security Studies Program. Founded in 2012, the GSSR has also served as the official publication of the Center for Security Studies and publishes regular columns in its online Forum and occasional special edition reports.

Access the Georgetown Security Studies Review online at
https://georgetownsecuritystudiesreview.org
Contact the Editor in Chief at gssr@georgetown.edu
# TABLE OF CONTENTS

**Of Lice and Men**
America Needs to Rethink Its National Security Paradigm  
*Gina M Bennett*

**What Made War Inevitable**
Great-Power Competition and Civil War in Thucydides’s History of the Peloponnesian War  
*Christian Trotti*

**Arms, Ideology, and Alignment**
Analyzing U.S.-Soviet Realignment During the Ogaden War  
*Matthew Fiorelli*

**Reining in the Iranian Nuclear Threat**
The Unviability of a JCPOA Revival and the Need for a New Game Plan  
*Logan Duarte*

**The Double-Edged Sword of Diplomatic Immunity**
The ICJ and the Case Studies of Germany v. Italy and Mohammed bin Salman  
*Apurva Ramakrishnan*

**Russian Influence and Disinformation Operations in the Balkans**

**Malaya to Vietnam**
The British Counterinsurgency Model and Its Replication Challenges  
*Steven Hu*

**Tightening the Screws**
Examining the Efficacy of U.S. Sanctions Against Russia Amid the Russo-Ukrainian War  
*Nick Lekkas*

**It’s All Connected**
The Impact of Russian Sanctions on Global Trade Relationships  
*Ashley Kelso*

**Decoding Beijing**
*Keven Hernandez*

**A Tale of Two Koreas**
How the Diverging Korean Language Will Challenge Future Unification  
*Olivia Schieber*

**The Decaying Superpower**
A Review of the Russian Navy  
*Bora Balçay*

**Sharing Secrets**
Why Do States Publicly Share Intelligence?  
*Peter Kerkhof*
Of Lice and Men: America Needs to Rethink Its National Security Paradigm

Gina M Bennett

Traditional security theories, such as variations of realism, continue to frame U.S. approaches to national security. Yet, their theoretical basis derives almost exclusively from the hunter-to-protector-defender survival roles as they evolved from pre-Neolithic to post-Neolithic eras. Because separate public and private spheres emerged from the agrarian revolution, post-Neolithic alphabetic writing largely excluded capturing the evolution of gatherer to caretaker-planner survival roles occurring within domiciles. Later, narratives from mythology, religion, philosophy, science, and politics devalued the criticality of societal security and well-being occurring in private spaces and, by omission, dismissed their relevance to public life. The lack of gatherer-caretaker-planner knowledge in conceptualizing traditional security theories is a key reason why democratic governance is insecure even while national sovereignty is safe. Not all existential threats emerge from tangible and external enemies, nor are solutions to them found through deterrence and military readiness.

Toward a Hunter-Gatherer Holistic Security Paradigm

A group of scientists published an article in the Jerusalem Journal of Archaeology in November 2022 about their discovery of an engraved ivory comb from 1700 BC—currently the oldest example of the earliest form of the alphabet.1 Inscribed on this delicate artifact were instructions for its use (may this tusk root out the lice of the hair and the beard). One of the archaeologists involved told The New York Times, “People kind of laugh when you tell them what the inscription actually says.”2 Perhaps they were hoping the discovery of the first known alphabetic writing would provide some profound insight into ancient history. Instead, it provides a glimpse into the history and evolution occurring within the domicile when only the activities in the public sphere were being recorded. Despite the amused reaction, this discovery is far more significant to contemporary societies than scientists or observers realize.

Before the invention of hygiene, it is very likely the domicile-confined caretaker of nearly 4,000 years ago observed the results of a lice infestation and recognized the existential threat it posed if left unaddressed. This post-Neolithic community came up with a solution clearly informed by patient observation and critical thinking. Then they took the time to painstakingly carve the purpose of the appropriately designed comb so anyone could see. Not only could this solution rid the domicile of the pest, but it also proves an effective strategy to prevent future infestations by using the comb routinely.

The proverbial saber-tooth tiger that the hunter-protector-defender was worried about would eventually grow full from eating humans. Lice would not. And it would be especially foolish to attempt to extinguish a lice infestation with a spear. So, while observers may laugh at this significant artifact’s banality, it clearly demonstrates a very different form of security approach. Creating security in this manner for everyone does not disadvantage anyone.

The example of the lice comb is surprisingly applicable to America’s security today. America is protecting itself against saber-tooth tigers but suffers from a long-festering “lice” infestation. Tiny bugs are carrying diseases across the country, such as political
apathy; partisan divisiveness; race, gender, and religious intolerance; hateful discourse; and ignorance of history and the Constitution.

According to Pew, 85% of Americans think the U.S. political system needs major changes. 3 58% are dissatisfied with the way democracy is working, 53% could not name the three branches of government, and 25% could not name even one. 4 Only 1% of Americans trust the federal government to do the right thing “about always.” Only 15% trust it to do the right thing most of the time. 5

The survey findings do not demonstrate the consent of the governed. The hunter-protector-defender approach of realism security theories may have kept America safe, but it has not secured America’s democracy. “There remains a clear and present danger to American democracy.” This oft-repeated statement by the Select Committee to Investigate the January 6th Attack on the United States Capitol should give pause to our continued adherence to realism security paradigms that clearly fall short of domestic stability. 6 A thousand Americans have been charged, some with seditious conspiracy, including the most powerful members of the Executive Branch and the former president himself.

Not since the Shays Rebellion in 1786-1787 has a vulnerability in the institutions of U.S. democracy been so shockingly exposed. 7 America’s founders, having overthrown such a tyrant in the form of King George III, were deeply distrustful of the seductive and corrupting nature of power. They painstakingly constructed a government that constrained the ability of any one person or institution to amass unchecked power. That construct is found in the three branches of government and the complex checks and balances created in their co-governing processes that more than half of America cannot even identify.

So, it’s not surprising that so many Americans are unaware of the full implication of the former U.S. president’s fixation on removing apolitical civil servants from across the Executive Branch and replacing them with political appointees loyal to him. 8 This one move alone would silence dissent and destroy transparency in governance. This simple act would turn America into an autocracy—a government in which one individual possesses all the decision-making power. No one “in the know” about the details and classified information related to government activities would be permitted to challenge the views of the person in power.

America’s democracy really can unravel that quickly. There are many countries around the world the U.S. foreign policy establishment has categorized as “autocracies” precisely for denying dissent within the ranks of government. Unlike autocratic, authoritarian, or totalitarian governments, America’s is not a winner-take-all setup. Elected officials in a representative democracy are accountable to their entire constituency, not just those who voted for them. Compromise, not tyranny, is the mechanism to end gridlock. America is at its best when it is governed by elected leaders and public servants alike who willingly submit to the scrutiny of their affairs and who choose self-restraint over personal gain. When America is at its best internally, its power of influence globally on both geopolitics and the global economy is far greater than when the government behaves erratically.

Ridding America of this lice infestation is a long-term mission, requiring patience and resilience. Even if the United States survives this current crisis, it certainly deserves a new security approach that would safeguard both
America’s domicile and its place in the global arena. This paper proposes a new security paradigm that incorporates a more balanced hunter-gatherer perspective to security studies theory to ensure national security approaches recognize the subtle dangers posing the greatest risk to democratic governance and prioritize appropriate solutions.

This paper is organized into two sections. First, the paper questions the utility of post-Neolithic historical archives as the basis for security studies. Findings from other scientific inquiries challenge the veracity and completeness of this history and indicate that the Neolithic division of society’s survival and security roles into public and private spheres led to the omission of critical knowledge about evolutionary approaches to societal threats, security, and prosperity. Drawing from the earliest known human survival roles and their evolution prior to the development of recorded history illustrates that the transparency of nomadic society enabled the development of shared values based on shared observations of Hunter-Gatherer (later referred to as hunter-protector-defender and gatherer-caretaker-planner) contributions to survival and well-being. Some roles aligned with anatomical advantages but were visible to all, thus, sex dimorphism and gender are referenced only due to history’s construction of them. Secondly, the paper analyzes foundational security studies' authors’ dismissal of the full spectrum of human activities. Instead, traditional security studies authors have subjectively heavily relied on documenting wars, violence, and external threats in their historical accounts, rendering security studies theories like realism inapt to solve other kinds of threats. The evident failure of U.S. national security approaches to have provided security for the Constitution and democratic governance should inspire seeking a better theory. This paper suggests leveraging experiences from prehistoric Hunter-Gatherer societies to enable a more holistic and balanced conceptualization of security that includes both the tangible safety of society and the security of its autonomy and prosperity.

The One Theory that Eluded Them All

One of international security studies most revered modern thought leaders, Kenneth Waltz, said in a 1998 interview:

I don’t think that anybody under the sun would deny the statement that if you could have a single theory that would comprehend both international and domestic, both political and economic matters, all in one theory, hey, that would be a lot better than a simple theory of international politics. However, nobody’s thought of how to do it. I’ve thought about that a lot. I can’t figure out how. Neither can anybody else so far.9

When Waltz said this, the world was grappling with the chaos resulting from the dissolution of the Soviet Union and the communist bloc. The peace dividend with the end of the Cold War was short-lived, leaving little time for new theorizing. Old, socio-political conflicts re-emerged, and the world came to be what former CIA Director Jim Woolsey characterized during his confirmation hearing as “a bewildering variety of poisonous snakes.”10 America turned to familiar realism-based security strategies once again, and it would seem that Waltz’ “one theory” proposal would remain an aspiration.

But does it have to be? What if a single security paradigm could enable a nation to holistically approach its security, balancing the domestic, international, political, and economic? Such a theory may not be as difficult to devise as so many have thought.
Rather than continually asking why one theory is so difficult to construct, a new approach would be to ask a different question: why is there a divide in the first place?

The answer to this question is frustratingly obvious, with substantial implications for security, law, politics, governance, and other disciplines. The failure to incorporate prehistory, combined with written history’s singular focus on the public sphere, resulted in the loss of a substantial body of knowledge that could be informing today’s security theories and strategies.

**Left Out of Recorded History…**

Theories in any discipline remain important because they create shortcuts for responding to a crisis or unfamiliar development without the delay of otherwise dissecting it in full. The discipline of security studies has constructed many theories for understanding the core drivers of power struggles and conflict to provide faster pathways to resolving them. Identifying specific ends, ways, and means—the subject of strategy—is made easier by having a sketch drawn already of the context and likely effective and ineffective ways and means to achieve the desired outcomes within that context.

It’s a bold task. Yet, security theories, like law, politics, philosophy, and other such disciplines, have drawn their conclusions from thousands of years of recorded history. By synthesizing data from this history and analyzing observations and insights, various theories have been constructed that the United States and other nations have leveraged to navigate an ever-changing set of threats to national and global stability.

Though drawing on thousands of years of written history may seem sufficient, this record is not as robust as one might think. For one thing, written history covers only about 0.8% of human history. Scientists have gathered abundant evidence of the existence of homo sapiens for 500,000 years, and yet alphabetic writing only dates back roughly 4,000 years. Thus, current theories about everything exclude as many as 496,000 years of human interaction and societal evolution.

What about the observations and insights of the other 99.2% that have not been part of informing theories? Exclusion of this knowledge should not be equated with a conclusion that the information is unimportant or irrelevant. Theorists simply have not known what was missing. But today, there is no excuse for ignoring at least some of this pre-history because other sciences, such as anthropology and neuroscience, have been filling in the gaps preceding documented time for decades.

Archaeologists and anthropologists have long noted that during the Paleolithic ages, early humans largely lived as nomadic forager-hunters and gatherers. As such, their domiciles were minimal and mobile, making privacy a near impossibility. Hunter and gatherer contributions to survival and human security could be viewed by all. This transparency over hundreds of thousands of years enabled shared observations and valuation of the importance of all to the sustainment of the human species.

Every capable person had a role to play in contributing to societal security amid harsh living conditions. Anatomical fit appears to have been a key determinant in the assignment of these roles. Strong, fast (typically male) bodies were better at hauling basic needs while relocating camp and fighting off bigger predators. Female bodies were the only ones that could produce children, resulting in their primary roles of caretaking, teaching, gathering resources, and strategic planning.
That said, roles were not assigned by birth sex or gendered in pre-history the way written history has portrayed them. If you could do a task well, it became your role because survival for all depended upon leveraging every person’s capabilities to the fullest. Based on anthropological studies and archeological findings of this pre-history, ongoing studies of current hunter-gatherer communities and research on primates most similar to homo sapiens, suggest modern sex and gender narratives that restrict roles based on birth sex did not apply to our hunter-gatherer forebears. 18, 19

This is an important fact too often lost in retroactive applications of gender narratives. The world today owes its existence to the 500,000-year survival of a society that did not apply stereotyped and limiting gender norms. That matters because it’s a much longer period of time - 125 times longer - than the 4,000 years or so societies have lived with them.

The 1975 publication of Toward an Anthropology of Women authored by a consortium of women anthropologists, sought to address the male bias in the field.20 Since their groundbreaking work, growing bodies of research and findings in the disciplines of anthropology, archaeology, neuroscience, and psychology, among others, are making new discoveries about old conclusions. Scholars and practitioners in these fields have been taking a fresh, objective approach in their analysis of history, pioneering approaches to new research, reducing the biases of latter-day patriarchal gender framing from understanding history as it was.

Recent studies debunk theories constructed on a long-held bias that only men were big game hunters, and that possessing the organs necessary for pregnancy must have limited the females’ ability to fulfill the broadest range of societal survival roles. 21 Researchers at the University of Washington and Seattle Pacific University found that women were not only habitual hunters but also engaged in more diverse methods for hunting by varying tools and approaches. Moreover, as an extension of their caretaker role, women even brought children along on hunting parties. 22

Similarly, researchers at the University of Seville and the University of Vienna overturned the long-held conclusion that the remains unearthed in Spain in 2008 from a burial site were those of a young male leader who had been dubbed the Ivory Man. 23 Based on the fact that there were displays of deference and respect towards the individual, archeologists assumed it was a man. With modern technology, scientists uncovered the presence of the X chromosome, and so now the remains are known as the Ivory Lady. The discovery suggests that women also held leadership positions in Europe’s Copper Age society between 2900 and 2650 BCE — before the invention of alphabetic writing.

The researchers noted that their findings argue for reconsidering “prevailing ideas about power, social complexity, and gender differences.” Yet, such conclusions in anthropological and archeological research rarely seem to resonate in political science. Because history, as recorded in writing, is incomplete, incorporating findings and new knowledge that fill in gaps and offer corrections should be a persistent mission of security studies. 24

…But Not Irrelevant or Unimportant

In a different realm of evolutionary science, neuroscientists have been using modern technology to uncover human instincts by scanning brain activity and responses to stimuli. What is clear from today’s science is that human instincts, molded by evolution,
still reflect the survival behaviors of the pre-Neolithic eras.

Neuropsychiatrist Dr. Louann Brizendine, M.D., takes these discoveries a step further and argues in *The Female Brain* that typical male and female brains function differently. Her premise is that male and female brains leverage their slight anatomical differences and respond differently both in how they perceive threats, and the reactions threats produce. Her arguments continue to spur discussion, inquiry, and debate in anatomical sciences. However, the core assertion that natural selection has honed survival instincts is not disputed. Fight, flight, and freeze have been described as nearly universal responses to threats, particularly perceived existential threats. Dr. Brizendine questions this universality based on her premise that male and female brains function differently and respond in unique ways to perceived threats, including how they first detect and assess whether something is a threat.

Her findings may require more research, but logic alone should lead to a similar conclusion. Suppose the typical male and female brains reflect hundreds of thousands of years of natural selection, 90% or more of which would have occurred during nomadic hunter-gatherer eras. Should we not expect descendants of hunters and gatherers to have inherited some distinctive survival instincts based on navigating different threat contexts and performing unique survival roles?

Hunter-protector-defenders benefit from having honed the skills and cognition for detecting dynamic threats and responding with decisive force. Yet not all existential threats are tangible, visible, or fast-moving. And violence simply cannot effectively neutralize so many of today’s diverse threats. Using a spear to attack a drought in hunter-gatherer contexts is like using a tank to tackle a pandemic, a battleship to counter malign influence, or armed drones to address political disenfranchisement and dysfunctional party politics. The wrong tools for the pacing threats.

Early gatherers would have had to navigate some predators, including hunters, in order to survive. However, gatherers experienced threats that emerged in more diverse ways and across a broader array of contexts, like lice infestations, infant illnesses, and droughts. To survive myriad threats, anthropologists tell us gatherers had to observe, collaborate, communicate their observations, strategize to create solutions, and patiently apply them to ensure they endure. These instincts helped gatherers detect more nuanced threats, rely less on ineffective violent responses to them, and create and share strategic solutions. Therefore, the post-Neolithic separation of survival roles into public and private spheres led to the exclusion of gatherer-honed instincts, skills, and knowledge about safety and security from the historical archives that informed modern-day security studies.

**The Divide that Distorted History**

Nomadic hunter-gatherer communities transitioned into settled societies with the discovery and eventual advancement of farming. No longer required to move to acquire food for survival, groups of humans could create early domiciles and farms. Put simply, hunters became protector-defenders of the farm and homestead, and gatherers became the primary caretakers and strategic planners for families. Both engaged in innovation and invention to improve the effectiveness and efficiency of their contributions to the family’s survival and well-being.

Though the Neolithic Revolution may have set the stage for modern civilization, it also produced an aberration in social structures with damaging consequences for societal
security. Agrarian settled life separated the roles of protector-defender from those of caretaker-planner into public view and private spaces, respectively. Both sets of roles remained equally critical to human evolution and survival, but this reality became less obvious due to the different natures of the public sphere and private space. The former was the locus for accumulating and sharing ideas, observations, and philosophies among more and more people gathering in public. The latter was shielded from collaboration and accumulation by protective ownership instincts that produced a desire for--and social, religious, and legal norms about--privacy.

This divide grew rigid with time, becoming a near-universal social norm, from which emerged the flawed assumption that what happens in the privacy of the domicile is not relevant to the public sphere. This belief strengthened over thousands of years to become “fact” by the time alphabetic writing emerged. The history recorded and studied by philosophers, warriors, political theorists, and other historical scribes, such as Thucydides, Cicero, Augustine, Aquinas, Machiavelli, Hobbes, Clausewitz, and Morgenthau, never questions this assumption, but rather offer narratives in further justification of it.

Given that disciplines other than political science are upending old conclusions about history, it’s imperative to consider the fallibility of the common works cited as foundational to security studies. The more ancient scribes dismissed the importance of the domicile with open disdain. For example, Thucydides, often considered the founder of “political realism,” deemed the activities of the domicile “frivolous.” In fact, he cast the entire societal backdrop as immaterial to the Peloponnesian War, and his theories derived from it. Imagine if this standard were applied to modern times. Would Americans agree that everything other than the Global War on Terror was unimportant for twenty years?

In addition to what Thucydides omitted in his recording of history, it is important also to scrutinize the accuracy of his direct and indirect observations of the Peloponnesian War. Thucydides admitted to including lengthy formal speeches that were literary “reconstructions” of what was said. He continued to write his history for years after the conclusion of the war in 404 BCE, presumably drawing on memory.

Since the 1970s, criminal investigative agencies and criminal psychologists have studied the risks of relying on eyewitness accounts to assign criminal culpability. Decades of scientific research have uncovered sources of inaccuracy based on an array of variables that affect and dilute memory. Eyewitness misidentifications are known to have played a role in 70 percent of at least 349 wrongful convictions in the United States overturned based on DNA evidence. Imagine if the basis upon which we have built national and international security theories is 70 percent wrong.

Thucydides is not alone in relying on direct observations and memory to record his texts. History that forms the foundational reading in most security studies programs, such as works by the list of authors noted above, did so as well. Research in other disciplines over the past 50 years should result in caveating the objectivity, reliability, and most importantly, completeness of such foundational texts.

Moving forward in time from Thucydides, Roman statesman Marcus Tullius Cicero constructed a concept of nation-state security that remains today and is rarely challenged for its irony or incompleteness. Cicero’s “the safety of the people shall be
the supreme law” has framed U.S. national security since its inception. Yet, Cicero condemned women who dared have interests outside their domestic confinement in his 56 BCE speech, *Pro Caelio*, as lacking in virtue and willing to use their sexuality to get ahead. This blatantly misogynistic narrative creates a compelling justification for his dismissal of the relevance of the experiences, knowledge, evolution, and thinking occurring within the domicile to theories and priorities for politics, governance, and security.

By the first century AD, Cicero’s definition of security as the safety of the state and the belief that women were inferior and their activities unimportant was further validated with the expansion of Christianity. In the first century BCE, St. Augustine offered the foundation for Just War theory, later advanced by 13th-century Thomas Aquinas. Like Cicero, they focused on physical threats and violence but excluded insights from other aspects of society or the other half of the population. Augustine said, “Woman was given to man, woman who was of small intelligence and who perhaps still lives more in accordance with the promptings of the inferior flesh than by superior reason.”

Aquinas was even more demeaning: “As regards the *individual nature*, woman is defective and misbegotten, for the active force in the male seed tends to the production of a perfect likeness in the masculine sex; while the production of woman comes from a defect in the active force or from some material indisposition, or even from some external influence.”

Had either philosopher asked a gatherer, they would have learned from women’s hundreds of thousands of years of experience the fruitlessness of engaging in a fight in which you do not have a reasonable chance of success. Or that gatherers had long devised ways to neutralize threats other than violence, which undercuts the conclusion that war is a last resort when you have not included all methods in the comparison.

These ancient scholars are mere examples of so many who hyper-prioritized war, violence, and external threats and made no attempt to balance their observations with the rest of human activity, priorities, and perceptions. These histories overwhelmingly assume that every other area of life — including the entire spectrum of human and economic interaction that the vast majority of people and societies were engaged in — was trivial when contrasted with the supreme importance of war and conflict. This matters because it distorts the definition of security and reduces it to tangible safety, and that does not reflect how so many people conceptualize security.

Modern scholars attempted to use more inclusive language to dismiss the relevance of the private sphere to their theories. They assumed that the hunter-protector-defender’s experiences and observations were universally shared, so when they spoke of “mankind,” they implied the universal “humankind.” As well intended as they may have been, the results are actually worse. They validate, in a more polite way, the still hubris-filled conclusion that the experiences of gatherer-caretaker-planners were the same as that of hunter-protector-defenders and that there are no applications in the public sphere to derive from the unique activities and methods developed in the private space for detecting and mitigating threats.

For example, 17th-century English philosopher Thomas Hobbes gave the world “social contract theory,” which sounds as if it might reflect a broader array of societal experiences but in the end, fails to give the interactions, knowledge gathered, or
priorities of the domicile equal consideration to those of the public sphere. Hobbes argues the need for law, order, and a sovereign by saying, “Government is necessary, not because man is naturally bad... but because man is by nature more individualistic than social.”

This description bears no resemblance to the role of mother and caretaker, which women were restricted to at the time, which are entirely social in nature. Moreover, as the one responsible for bearing, nursing, and raising children and caring for anyone unable to care for themselves, her role is demonstrative of the least individualistic nature you could find.

Perhaps Hobbes meant to argue that women did not need government? Or, that women should lead government because a man cannot be trusted with responsibility over more than himself?

In the 19th century, Max Weber defined a nation-state in a way that still dominates security theories and international relations: “a human community that (successfully) claims the monopoly of the legitimate use of physical force within a given territory.”

No thought to the alternative gatherer-caretaker-planner perspective that a human community that successfully shares resources, enables its population to reach their full potential to contribute to society, and creates justice for all within a given territory should be considered a nation-state.

Even more modern theorists derive their observations from and through a singular focus on violent threats, violent responses, and the history of the hunter-protector-defender’s dominance in the public sphere. In his book, Man, State, and War (note the title), Kenneth Waltz noted, “The locus of the important causes of war is found in the nature and behavior of man. Wars result from selfishness, from misdirected aggressive impulses, from stupidity.”

Taking all these theories together, one could logically conclude that the nature and behavior of women—of whom so many religious leaders, political theorists, and philosophers took great pains to describe as different from that of men—holds the antidote to selfishness, misdirected aggression, and stupidity. Another implication of the logic might be that the experiences, observations, insights, and skills that evolved within the gatherer-caretaker-planner’s private sphere could both reveal tactics and strategies to neutralize threats unresponsive to violence and offer ways to avoid conflicts in the first place.

The Neolithic revolution produced a great deal of human advancement, but it also created an artificial hard divide between public and private spheres, that in turn, disrupted hundreds of thousands of years of a more balanced and realistic concept of societal security. It is this divide that stumped Kenneth Waltz and others. Additionally, traditional security theories also are insufficient to guide today’s security approaches because even the 0.8% of history used to inform them derived from, at best, half of societal knowledge about security. It matters to national security because history’s extremely limited and disparaging coverage of gatherer-caretaker-planner’s roles cemented an assumption that because they were “un-hunter-like,” they were not useful to the security of the state.

Bridging Divides to Secure America

The lice infestation plaguing America will continue to spread without substantially changing how the country views its national security as complementary but separate from national safety. Knowledge of how democracy works and what it requires of
every American—regardless of age, party affiliation, residence, economic status, or any other demographic—is the most critical requirement for securing a functioning democracy. Civics education for all ages would not cost as much as building a single aircraft carrier.

Hunters and gatherers prioritized using limited resources differently, but they could each see the importance and value of the other’s contribution to societal safety, security, and well-being. The persistent devaluing of gatherer-caretaker-planner roles is eating at America’s ideals, integrity, and future.

Divided values between public and private spaces are evident most tangibly in how caretakers and educators are treated compared to the Presidents, manufacturers, inventors, legislators, judges, CEOs, entrepreneurs, scientists, and adult citizens that they create and train. Partial and full-time stay-at-home parents are among the most marginalized segments of the U.S. population. Their work is unpaid, unseen, and uncounted. Their needs for raising healthy and responsible children are systematically de-prioritized, even by politicians who claim to glorify them.

The emergence of private property out of the Neolithic Revolution offered the promise of improved safety for humans in fixed dwellings that could be more easily protected. Yet, it resulted in laws and norms that prioritize the threat context of the hunter in the public space over that of the gatherer in the domicile. Today, the physical safety of one in four women in America is at greatest risk in the privacy of their homes.

The divide between hunter and gatherer approaches to community survival also manifests itself in the divided politics and economic investments of agricultural and industrial sectors and communities. Rural-based workforces and communities in the United States have been disproportionately disadvantaged by globalization and technological advances. Bridging this divide requires deliberate and sustained efforts that may appear to industrial communities as inequitable. Gatherers-caretakers know that sometimes you have to provide more to some, and less to others, to produce equality.

De-valuing and dehumanizing the “other,” which began in ancient times with misogynistic narratives of women and their domestic concerns, has become the norm in public discourse. Yet, this kill-or-be-killed approach prevents security for all. Tolerant and respectful discourse, modeled by public leaders committed to demonstrating that divergent priorities can co-exist peacefully, would produce better outcomes and stability for all.

The public sphere’s preoccupation with the here and now can desensitize it to the need to plan for the future, and this is evident in the generational divide in American socio-political discourse today. Adaptive evolution has been key to human survival, requiring planning and adjusting to change rather than fighting it. Prioritizing healthy living and innovative use of resources with the same fervor as expanding the defense apparatus is evolution. Because every sector, from education to defense, needs Americans who are fit and mentally resilient.

A more balanced hunter-gatherer national security would require recognizing that many existential threats emerge more subtly and with more complexity than obvious acts of violence, making reflexive hunter responses the wrong tool for mitigating them. America needs a new security paradigm that leverages history and prehistory to provide a fuller understanding of how to detect and eliminate threats, whether foreign or domestic, violent or nonviolent, dynamic or stealthy, of known
origins or unattributable. Such a theory would, as Waltz put it, “be a lot better.”

About the Author: Gina M Bennett is a retired member of the CIA’s Senior Analytic Service, author of two “National Security Mom” books, and adjunct professor at Georgetown University, where she teaches “Hunter-Gatherer National Security” at the Edmund A. Walsh School of Foreign Service. She also serves on the Alturas Institute’s Board of Directors and as Strategic Advisor to Girl Security.


4 Schaeffer, Pew Research Center.


15 Rapp, Toward an Anthropology of Women.


17 Rapp, Toward an Anthropology of Women.

18 Quinn, “Anthropological Studies on Women’s Status.”


20 Dyble, et. al., “Human Behavior.”


24 Cintas-Peña, et. al., “Amelogenin Peptide Analyses Reveal Female Leadership in Copper Age.”


26 Scott, Gender and the Politics of History.


What Made War Inevitable: Great-Power Competition and Civil War in Thucydides’s History of the Peloponnesian War

Christian Trotti

In History of the Peloponnesian War, Thucydides famously asserts that “what made war inevitable was the growth of Athenian power and the fear which this caused in Sparta.” However, by only focusing on this parsimonious insight about a change in the balance of power at the systemic level of analysis, many scholars ignore Thucydides’s emphasis on factors at lower levels of analysis within states. Like other hegemonic wars, the history of the Peloponnesian War cannot be told without acknowledging the catalyzing role of smaller civil wars and proxy conflicts in addition to government policies and individual decisions that shaped them. Therefore, given the role of the Peloponnesian War as a case of intrinsic importance in the literature, this paper further investigates the war to establish a historical and analytical basis for exploring how contemporary proxy conflicts (e.g., the Russo-Ukrainian War) might affect strategic competition between the United States and China today.

Through case study and historical analysis of the Peloponnesian War, this paper finds that bipolar great-power competition, hegemonic interstate war, and civil war are intrinsically linked in Thucydides’s narrative. Therefore, this paper develops the following model for understanding the relationship among phenomena at different levels of analysis, and how these phenomena produce and accelerate hegemonic war. First, bipolar competition between two great powers at the systemic level of analysis alters the relative balance of power between smaller states and their populaces, thereby changing the incentive structures of enterprising individuals who subsequently mobilize factional violence and launch civil wars. Second, these civil wars at lower levels of analysis eventually exacerbate systemic volatility, driving the competing great powers to hegemonic interstate war once they affect vital national interests.

This paper begins by applying the contemporary civil war literature to eight case studies of civil war from early in History of the Peloponnesian War, thereby exploring how systemic factors like great-power competition could affect state-societal phenomena like civil wars. It then transitions to textual and historical analysis to reinterpret Thucydides’s power transition thesis, revealing how civil wars and proxy conflicts exacerbate tensions at the systemic level of analysis. Finally, the paper concludes with further avenues for research and policy implications for contemporary great-power competition.

Introduction

Although written nearly 2,500 years ago, describing a war between two ancient powers that no longer exist, Thucydides’s History of the Peloponnesian War continues to offer a wealth of insights for contemporary readers. Relating to the causes of war, Thucydides is well-known among international relations (IR) scholars for providing the analytical foundation of modern power transition theory, or the theory of hegemonic war. This explanation asserts that interstate war is caused by a change in relative power between two states, as the hegemon declines and a challenger rises. The earliest articulation of this theory is Thucydides’s central thesis that “what made war inevitable was the growth of Athenian power and the fear which this caused in Sparta.” Recently re-popularized in the policy discourse as the “Thucydides
Trap’ by political scientist Graham Allison, this theory is often used to explain a variety of great-power conflicts throughout history, including World Wars I and II. Today, scholars and policymakers regularly debate whether or not the United States and the People’s Republic of China (PRC) will fall into such a trap—with pessimists drawing upon History of the Peloponnesian War to argue that such tectonic conflicts are the almost inevitable products of timeless human instincts like fear, honor, and self-interest, leaving little room for contemporary policymakers to navigate these historical trends.

However, while often quoted, Thucydides is rarely understood. By only focusing on structural factors like the balance of power in the broader international system, many IR scholars ignore Thucydides’s emphasis on factors at lower levels of analysis within states (e.g., the character and policies of governments at the state-societal level, the clash of bureaucratic politics at the organizational level, or the decisions of specific leaders at the individual level). Specifically, History of the Peloponnesian War is often preoccupied with the plethora of civil wars that occurred within the various city-states of the ancient Greek world, exploring how these conflicts at the state-societal level of analysis interact with the systemic great-power competition between Athens and Sparta. Indeed, in an underappreciated quote, Thucydides asserts that “later, of course, practically the whole of the Hellenic world was convulsed, with rival parties in every state—democratic leaders trying to bring in the Athenians, and oligarchs trying to bring in the Spartans.” These leaders were often successful, playing the major powers against both each other and local enemies while creating new flashpoints and battlegrounds that shaped the course of the central conflict between Athens and Sparta. Indeed, the Peloponnesian War was not only a conflict between Athens and Sparta—it was a clash of competing alliance architectures, as the Athens-dominated Delian League confronted the Spartan-led Peloponnesian League. The alignment and behavior of every city-state were therefore critical to the overall preparations for and conduct of the war, and proxy conflicts abounded as Athens and Sparta continuously intervened in other city-states’ domestic politics.

Thus, the history of the Peloponnesian War—and many other hegemonic wars that featured a systemic power transition—cannot be told without acknowledging the catalyzing role of smaller civil wars and proxy conflicts in addition to government policies and individual decisions that shaped them. This is a glaring gap that scholars and policymakers can no longer afford to ignore. Given the role of History of the Peloponnesian War as a landmark text and a case of intrinsic importance in IR scholarship, any reinterpretation of Thucydides’s thesis would bear significant implications for understanding and extrapolating modern trends. Therefore, to establish a historical and analytical basis for exploring how contemporary proxy conflicts (e.g., the Russo-Ukrainian War) might affect strategic competition between the United States and China today, it is imperative to further investigate the Peloponnesian War and the related conflicts that surrounded it.

Upon further analysis of the Peloponnesian War, as well as prevailing civil war and IR scholarship, it is evident that bipolar great-power competition, hegemonic interstate war, and civil war are intrinsically linked. Thucydides’s narrative indicates a confluence of factors among the systemic, state-societal, organizational, and individual levels of analysis. Thus, based upon the Peloponnesian War, this paper develops the following model for understanding the
relationship among phenomena at different levels of analysis, and how these phenomena produce and accelerate hegemonic war:

First, bipolar competition between two great powers at the systemic level of analysis alters the relative balance of power between smaller states and their populaces, thereby changing the incentive structures of enterprising individuals who subsequently mobilize factional violence and launch civil wars. Second, these civil wars at lower levels of analysis eventually exacerbate systemic volatility, driving the competing great powers to hegemonic interstate war once they affect vital national interests.

In sum, systemic instability causes civil war at lower levels of analysis, which reverberates back to the systemic level. The dynamic phenomena of this model indicate a pervasion of instability across levels of analysis, represented metaphorically by convulsions that rattle the entire system and each of its constituent elements. Such pervasive instability had a major effect on the causes and early conduct of the Peloponnesian War.

To build the first half of the model, the paper first reviews insights from the contemporary civil war literature as a means to explore the mechanisms by which systemic factors like great-power transition could affect state-societal phenomena like civil wars. It then applies these frameworks to eight case studies of civil war from early in History of the Peloponnesian War. Of the eight, six civil wars occurred before the Peloponnesian War: Boeotia, Samos, Epidamnus, Macedon, Potidaea, and Plataea. The other two occurred early in the Peloponnesian War: Mytilene and Corcyra. Since Thucydides presents ancient Greek city-states as autonomous political units, most of the civil wars analyzed in this paper will be those within city-states, with two exceptions: Potidaea and Mytilene. These two city-states did not suffer a civil war within themselves; instead, they revolted against the broader Athenian Empire. Since the Athenians had consolidated their interstate alliance architecture, the Delian League, into a relatively centralized empire controlled by Athens, this empire can be considered a broader state. Thus, in revolting against it, Potidaea and Mytilene fought “anticolonial wars”—defined as civil wars by political scientists James Fearon and David Laitin, among others. Drawing from this case study analysis, the paper builds the first half of the model, articulating direct and indirect pathways for translating systemic competition to state-societal violence and exploring ‘transmission belts’ at the individual and organizational levels of analysis. This first half concludes by refuting potential counterarguments.

The paper will then transition toward building the second half of the model, exploring how civil wars and proxy conflicts exacerbate tensions at the systemic level of analysis. Through textual and historical analysis, this section first reinterprets Thucydides’s power transition thesis by revealing that the unique importance of alliances in the Hellenic world—and thus the Athenian encroachment on Spartan allies at the state-societal level—was a crucial factor in explaining the outbreak of the war. This section then extrapolates these lessons to construct the second half of the model, arguing that it is not the power transition by itself that causes hegemonic war, but rather the moment when this power transition affects certain policy objectives and national security interests set by lower levels of analysis, whether individual, organizational, or state-societal, as established by the first half of the model.

Upon completing the model, the paper concludes with further avenues for research.
and policy implications for contemporary great-power competition.

**Systemic to State-Societal: How Great-Power Competition Incentivizes Local Civil Wars and Proxy Conflicts**

As the first half of the overall model, this section seeks to explore how great-power competition at the systemic level of analysis acts as a source of instability that affects smaller states and provokes civil wars and proxy conflicts at the state-societal level. To do so, this section incorporates arguments from contemporary scholarship on the causes of civil war, thereby testing their relevance for eight case studies of civil war immediately before and during the Peloponnesian War.

*The Civil War Literature: Motivation vs. Opportunity*

The contemporary civil war literature is shaped by a debate between motivation and opportunity as the most relevant cause of civil wars. Regarding the former, it is common to explain conflicts by describing the grievances that compelled the belligerents to act. These motivations can be derived from different levels of analysis. For example, accounts of the Bolshevik and Cuban Revolutions often highlight the central role of figures like Vladimir Lenin and Fidel Castro, respectively. Their idiosyncratic political ambitions, distaste for the previous regime, and interpretations of Marxism shaped the causes and conduct of intrastate violence in these cases, thus demonstrating *individual* level of analysis. By contrast, scholars often attribute the American Civil War to conflicting motivations at the *state-societal* and *organizational* levels of analysis, as the Confederate States of America sought more expansive interpretations of state rights to permit the continuation and expansion of slavery. Each civil war, like each interstate war, is defined by its own unique motivations, thereby offering a compelling story—which makes this school of thought popular. However, the primary drawback is the difficulty in establishing these motivations as a primary cause or independent variable—since they are usually present even in the absence of conflict or violence. For example, the grievances that contributed to the Bolshevik Revolution, the Cuban Revolution, and the American Civil War had all been simmering for decades, but violence only occurred at a specific point in time. Why does the outbreak of violence not always correlate with the inception of belligerent motivations?

To explain this gap, a competing school of thought prioritizes opportunity as a primary causal factor in explaining civil war. These scholars posit that, while motivations and grievances endure for long periods of time, belligerents (usually insurgents) only initiate violence when the ideal opportunity arises. One important thread in this literature is the idea of “weak states”. James Fearon and David Laitin hypothesize that “financially, organizationally, and politically weak central governments render insurgency more feasible and attractive.” Therefore, changes in the relative balance of power between a state and its prospective rebels at the state-societal level of analysis—specifically when rebels are empowered relative to the state—often embody a key causal opportunity for the initiation of intrastate violence.

Accordingly, there are a variety of *internal*, state-societal factors that can weaken a state. For example, rough terrain and large populations may complicate efforts at policing or deterring rebels, as demonstrated by the French Indochina War and the Afghanistan War—although, like motivations, those factors are also largely constant and therefore do not always explain
the specific outbreak of violence. Better internal factors include economic or political crises that prevent a government from organizing sufficient resources to deter or combat rebels.

Within this opportunity-oriented school of thought, scholars are increasingly interested in proxy conflicts, where external factors like the balance of power at the systemic level of analysis can weaken smaller states and provide the opportunity for civil war. For example, Stathis Kalyvas and Laia Balcells apply this perspective to the Cold War by arguing that “[US and USSR] raised the military capacity of both states and rebels worldwide,” but “[t]his mutual improvement in military capacity nevertheless favored the rebels, a result of the rise and diffusion of a particularly robust version of the technology of insurgency.” Thus, in order to strengthen their respective blocs, the United States and the Soviet Union provided direct support to friendly rebels against regimes backed by their rival—or conversely, withdrew support for states perceived to be aligning more closely with that rival. The proxy nature of these conflicts often empowered local rebels to wage civil war against their respective states.

Regardless of whether the root factors are internal (state-societal) or external (systemic), this opportunity-oriented school of thought maintains that civil war at the state-societal level of analysis is caused by the intrinsic opportunity provided by a weakened state. However, this explanation is incomplete. Ultimately, the rebels and their leaders must decide to take advantage of such opportunities. Therefore, the literature also covers the causal mechanism, or ‘transmission belt,’ where a change in the relative balance of power between the state and the rebels at the state-societal level of analysis results in civil war. This necessitates the involvement of the individual and organizational levels of analysis. In the civil war literature, there is a phenomenon whereby individuals exploit opportunities in order to advance their own power or interests. These “political entrepreneurs—individuals who may not share the beliefs of extremists but who seek political office and power—may reflect the polarization of societies and, through their actions, propel this process further.” Thus, political entrepreneurs at the individual level of analysis use their rhetoric to mobilize factions at the organizational level of analysis, committing violence against their fellow countrymen and provoking civil war when confronted with the right opportunity.

Ultimately, the civil war literature provides an abundance of insights, specifically on how intrastate violence at the state-societal level of analysis may be caused by factors and variables across the levels-of-analysis framework. This literature is incrdbly useful in explaining the complex phenomena of the Peloponnesian War.

**Case Study Analysis**

A close analysis of the eight selected case studies of civil wars before and during the Peloponnesian War suggests that the opportunity school of thought better explains these conflicts. Moreover, even within this opportunity school of thought, the scholarship on external, systemic factors is most applicable. Accordingly, most of the civil wars in Thucydides’s analysis demonstrate how the systemic power transition from Sparta to Athens and the resulting great-power competition between these rivals weakened local city-states, thereby providing the opportunity for rebellion at the state-societal level of analysis.

This paper does not seek to discount the role of motivation throughout the civil wars in
History of the Peloponnesian War. Indeed, Thucydides himself asserts that a primary motivation for civil war at the individual level of analysis was the “love of power, operating through greed and through personal ambition.”\(^{15}\) However, such general greed is an enduring element of human nature, which is a constant, and “to explain anything one must consider more than human nature.”\(^{16}\) Thucydides also discusses state-societal motivations—in one passage recounting a pre-war debate in Sparta, the Athenian delegation acknowledges that many of their subject city-states harbor grievances against the Athenian Empire.\(^{17}\) However, for much of this period, there were no civil wars. Ultimately, “grievances can be found literally everywhere but civil wars are relatively rare, so the presence of grievances overpredicts the occurrence of civil war.”\(^{18}\) Therefore, a constant independent variable cannot explain a varying dependent variable.\(^{19}\) In the period before the Peloponnesian War, almost constant motivations like individual greed and state-societal grievance cannot explain the outbreak of civil war or the lack thereof. While these motivations are certainly important in shaping subsequent intrastate violence, such motivations are a necessary but insufficient cause of civil war.

By contrast, opportunity-oriented literature provides better explanations. For case studies Boeotia and Samos—both of which occurred during the Pentecontaetia period (429-435 BCE) prior to the Peloponnesian War—systemic factors clearly precipitated the causal opportunity for civil war. In Boeotia, civil war began when an exiled Boeotian party seized several Boeotian towns, thus battling the pro-Athens Boeotian regime.\(^{20}\) Thucydides observed that Boeotian exiles launched the civil war precisely when Athens was distracted by Spartan attempts to change ownership of the temple at Delphi.\(^{21}\) In other words, the Boeotian rebels attacked when they believed that Athens was more preoccupied with staving off Sparta’s influence and thus less able to support its friendly regime in Boeotia.

Meanwhile, civil war began in Samos when Athens decided to help Miletus in its war with Samos as well as support private Samos individuals who hoped to set up a different form of government. Accordingly, the Athenians ordered a contingent of their navy to establish a friendly democratic regime in Samos.\(^{22}\) In this case study, the Athenian desire to expand its sphere of influence relative to Sparta resulted in Athens exporting democracy to Samos, reflecting the importance of regime change and proxy conflict as a tool in their grand strategy. Through direct military aid to the pro-democracy faction within Samos, Athens provided an ideal opportunity for the state-societal phenomenon of civil war.

This systemic ‘opportunity’ argument is similarly relevant in the four civil war case studies that occurred immediately prior to the Peloponnesian War (435-431 BCE): Epidamnus, Macedon, Potidaea, and Plataea. Epidamnus, for example, “declined and lost most of her power” as a result of an interstate war with foreign enemies, which provided the opportunity for a democratic faction within Epidamnus to oust the aristocratic regime.\(^{23}\) These exiled aristocrats then joined the aforementioned foreign enemies in attacking Epidamnus, thereby forcing the new democratic regime to seek Corcyrean and then Corinthian aid.\(^{24}\) In sum, Epidamnus experienced civil war due to external war weakening the state, but this conflict was somewhat unique in that its root causes were largely isolated from the systemic instability wrought by the Athens-Sparta geopolitical rivalry.

By contrast, Athens and Sparta were more directly involved in creating opportunities
for civil war in Macedon, Potidaea, and Plataea. Specifically, the Macedonian civil war was caused by direct Athenian aid to rebel leaders Philip and Derdas against the reigning Perdiccas, likely as a means of curbing Perdiccas’s ability to intervene in important Peloponnesian affairs such as the battle for influence over Potidaea. This aid affected the relative balance of power between the rebels and the Macedonian state, forcing Perdiccas to seek external allies to shore up his compromised military position.\(^{25}\)

Meanwhile, the roots of the Potidaean revolt against Athens were more complicated. Potidaea was officially a colony of Corinth, but they were simultaneously “allies of Athens in the tribute-paying class,” and were thus part of the Athenian Empire.\(^{26}\)

Due to escalating tensions between Corinth (an ally of Sparta) and Athens, the latter cracked down on Potidaea before Corinth or other rival powers like Macedon could pry them from the Athenian Empire. As a result, Athens sought to “make… alterations in the existing state of affairs”—including by demanding Potidaea “to pull down the fortifications looking towards Pallene, to send hostages to Athens, to banish their Corinthian magistrates, and in future not to receive those who were sent out annually from Corinth to replace them”—thereby incurring the wrath of the Potidaeans.\(^{27}\)

Thus far, this analysis has focused on motivations, but as usual, these motivations were insufficient (albeit necessary). Civil war only began when the Spartans provided an opportunity for Potidaea to revolt; they promised to invade Attica, the bountiful and valuable Athenian countryside if Athens attacked Potidaea.\(^{28}\)

The promise of this direct military assistance (a result of the systemic bipolarity between Athens and Sparta) offered the prospect of changing the balance of power between Athens and Potidaea, thus allowing Potidaea to wage an anticolonial civil war against Athens.

Immediately before the Peloponnesian War, civil war also began in Plataea due to Thebes, which used its military to directly intervene on behalf of a Plataean faction trying to revolt against the pro-Athenian regime in Plataea.\(^{29}\)

The Thebans were incentivized to do this by the broader competition between Athens and Sparta, which they saw as rapidly accelerating toward a potential hegemonic interstate war. Again, systemic factors had state-societal repercussions.

Lastly, the two civil war case studies during the early Peloponnesian War demonstrate the opportunity provided by systemic factors. Mytilene had desired to revolt against the Athenian Empire since before the Peloponnesian War, but it waited because the Spartans “had not been willing to receive them into their alliance.”\(^{30}\)

This speaks to the importance of opportunity over motivation; while the motivation existed, they needed to wait until an opportunity presented itself at the systemic level. This occurred as Athens was increasingly “suffering from the plague and also from the full force of the [Peloponnesian] war.”\(^{31}\)

Also, it appeared that Mytilene would eventually be aided by the Spartans and Boeotians.\(^{32}\)

This opportunity allowed for the revolt of Mytilene in 428 BCE. Civil war also occurred in Corcyra when Corcyrean prisoners returned to their city to change Corcyra’s allegiance from pro-Athens to pro-Corinth (and thus pro-Sparta). This resulted in a civil war between the pro-Corinth oligarchs and the Pro-Athens democrats in 427 BCE.\(^{33}\)

Building the First Half of the Model: Indirect vs. Direct Pathways and Transmission Belts
Analysis of these civil war case studies demonstrates that bipolar great-power competition at the systemic level alters the relative balance of power at the state-societal level, weakening the city-state’s authority and providing the opportunity for the rebels to wage a potentially successful civil war. This paper adopts two pathways that outline alternative mechanisms by which the systemic level translates to the state-societal level: indirect vs. direct. In the indirect pathway, rebels perceive that their state is weaker due to the instability wrought by great-power competition, especially if their state is potentially threatened by external actors. This speaks to the importance of subjective perceptions of future state weakness among the rebels, in addition to objective estimations of its current strength or weakness. In this indirect category are Boeotia, Mytilene, and, to some extent, Corcyra (since the rebels exploited the course of the war to provoke anti-Athenian sentiment). This path corresponds to Fearon and Laitin’s emphasis on weak states.

On the other hand, in the direct pathway, one of the external powers intervenes directly on behalf of the rebels; it is incentivized to do so by the volatility of systemic bipolarity, which demands an expansion of one’s sphere of influence relative to the enemy bloc. The civil wars in Samos, Macedonia, Potidaea, Plataea, and, to some extent, Corcyra (it seems that Corinthian aid was implicit) fall into this direct category. This path corresponds to Kalyvas and Balcels’ argument about the role of direct superpower intervention in causing civil war.

Regardless, for both pathways, systemic great-power competition between Athens and Sparta set the conditions for state-societal civil war. Of course, it is important to note here that the actors at the systemic level—especially those that intervened and caused civil wars—were not always Athens and Sparta specifically; smaller powers like Corinth, Thebes, and Corcyra sometimes provoked or shaped civil wars in other states. Regardless, these smaller city-states were acting within a geopolitical context defined by competition between two competing blocs: the Delian League (increasingly consolidated into the Athenian Empire) and the Peloponnesian League. IR scholars refer to this systemic distribution of power as a bipolarity. Thus, while smaller city-states like Corinth, Thebes, and Corcyra retained the power to shape civil wars in other city-states—embodiment some tenets of a multipolar distribution of power—they did so to improve the strategic position of their broader bloc, which embodies bipolarity.

The synthesis of these elements reflects the “bi-multipolarity” of the system, which demonstrates the intersection of different causal factors that often complicate the application of IR frameworks. There are a few exceptions to these indirect and direct pathways. First, Epidamnus does not fall into either category. While this civil war was certainly caused by the weakness of the Epidamnian city-state due to the systemic-level factor of war with an outside enemy—thereby corresponding to the opportunity literature writ large—it was not the result of great-power competition. Thus, it does not reflect the overall model whereby systemic bipolarity shapes state-societal civil war; however, it still speaks to the importance of destabilizing systemic phenomena, since competition involving Corcyra, Corinth, Athens, and Sparta was important in escalating this civil war. The other exceptions are the civil wars in Ithome (helots vs. Spartans) and Egypt (Egyptians vs. Persians). Since Thucydides does not ascribe an explicit cause to them, they cannot fit into this model and were therefore not covered in the analysis above.
As covered in the literature review above, there is also an important ‘transition belt’ that needs to be accounted for in any model relating systemic to state-societal phenomena: the role of political entrepreneurs. Thucydides addresses these political entrepreneurs in his work. When he discusses Corcyra in Book Three, he expounds his general thesis on civil war, and he explains that “leaders of parties in the cities [who were fighting civil wars] had programmes which appeared admirable—on one side political equality for the masses [i.e. democracy], on the other the safe and sound government of the aristocracy—but in professing to serve the public interest they were seeking to win the prizes for themselves.”\(^{38}\) These political entrepreneurs preyed on a broader Thucydidean fear which transcended the levels of analysis; this was a direct result of the systemic “instability in the Greek world” which affected the relative balance of power within a state and propagated uncertainty among the populace.\(^{39}\) Thus, by organizing factions, the entrepreneurs developed an ideology that allowed them to court the Spartans if they were oligarchic or the Athenians if they were democratic; but they were only motivated by personal gain through regime change so that they would become the new authority.\(^{40}\) The systemic instability, a result of bipolar great power competition, allowed them to do this; according to Thucydides, “when each party could always count upon an alliance which would harm its opponents and at the same time strengthen its own position, it became a natural thing for anyone who wanted a change of government to call in help from outside.”\(^{41}\)

This mechanism completes the first half of the overall model of pervasive instability developed by this paper. Great-power competition at the systemic level shifts the relative balance of power between a smaller city-state and its prospective rebels at the state-societal level. Such shifts alter the incentive structures of political entrepreneurs at the individual level, providing them with the opportunity to prey on fear, to mobilize factions at the organizational level which potentially correspond to the regime type of an external great power, and to rebel against their own state. This subsequently results in the phenomenon of civil war at the state-societal level of analysis. Sometimes the process is indirect, as political entrepreneurs and their factions look for situations where their state is weakened by external factors; sometimes the process is direct, as political entrepreneurs and their factions receive direct aid from an outside power. Regardless, this model demonstrates the influence of volatility throughout the system. In summarizing the ripple of instability from the systemic level to the state-societal level, it can be said that “the very existence of the [great power] rivalry encouraged dissent” at lower levels of analysis.\(^{42}\)

**Counterarguments**

IR scholars and historians may disagree with this model as it relates to the Peloponnesian War. One potential counterargument could assert that contemporary civil war scholarship cannot be applied to the Peloponnesian War, since insurgencies of the Cold War and post-Cold War do not resemble ancient Greek civil wars. For the most part, Greek rebels in small city-states necessarily operated differently from most modern insurgents in nation-states — given that states were significantly smaller, the former did not have the luxury of retreating into the rough terrain of the hinterland and mobilizing forces away from the authorities, so insurgencies largely possessed a different character. Similarly, during antiquity, tactics for receiving external aid were limited, including opening the city gates to another
state’s army or receiving physical money or resources. Such tactics do not approximate the plethora of options presently available to great powers in providing aid to rebels, such as modern means of financing, weapons shipping, intelligence support, and other tools.

However, there are three reasons to refute this counterargument. First, the most relevant lesson from civil war literature is the importance of the weak state; the opportunity provided by an externally weakened state in causing a civil war was just as important in the Peloponnesian War as it was for the Cold War and beyond. Second, there are multiple similarities between contemporary insurgencies and ancient Greek rebellions. In connecting the great-power competition of the Cold War to the civil wars of that period, Kalyvas and Balcells argue that the United States and the Soviet Union provided “material support, revolutionary beliefs, and military doctrine” to insurgencies. These channels were also employed by ancient Greek insurgents. In the Hellenic world, material support for or against city-states significantly altered the relative power between the state and rebels and was crucial in causing civil war (especially through the direct path). Regarding revolutionary beliefs, ancient Greek political entrepreneurs mobilized people along faction lines, tying their identity (oligarchic or democratic) to the regime type of the state to which they were appealing (Sparta or Athens, respectively). The same can be said of communist and anti-communist rebels during the Cold War. In terms of military doctrine, while Hellenic insurgencies were certainly different from insurgencies in the Cold War, there is an important similarity: they were both highly political endeavors, involving more than just military force. Thus, great-power competition, whether between Athens and Sparta or between the United States and the Soviet Union, caused and affected civil wars in equivalent ways, resulting in similar insurgencies. Finally, some of the ancient Greek civil wars involved rebels who fought outside of the city-state and launched attacks from less controllable areas of operation; examples include Boeotia, Samos, and Corcyra. These resemble contemporary insurgencies that utilize difficult terrain.

A second counterargument could seek to discount the applicability of current scholarship because it is often based upon ethnic civil wars, which may not be relevant in a time that preceded the modern concept of nation-states. In the civil wars within each city-state, the combatants often viewed themselves as members of the same state; without ethnicity, along what lines did they oppose each other? This counterargument can be refuted by the fact that they were divided along faction lines. No matter the era, some form of identity will be made politically salient by entrepreneurs; in ancient Greece, this identity was primarily oligarchic versus democratic. Besides, there was an element of ethnicity in some Hellenic civil wars, including the helot revolt at Ithome and the Potidaean and Mytilenian civil wars against Athens. Thus, the model stands.

State-Societal to Systemic: How Civil Wars and Proxy Conflicts Contribute to Hegemonic War

Reinterpreting Thucydides’s Power Transition Thesis

The second half of the model seeks to demonstrate how civil war (produced by the first half of the model) reverberates back to the systemic level to exacerbate great-power competition and shape great powers’ behavior before and early in interstate war—in this case, between Athens and Sparta.

There is evidence that, at the systemic level of analysis, both Athens and Sparta feared...
civil war within their own blocs, especially if the other helped to cause it. To maintain the strength of their respective blocs and spheres of influence, the two powers were expected to “have arranged the affairs of the various states so as to suit” themselves.49 This included maintaining friendly regimes and preventing civil war amongst their allies. For instance, Corinth and Corcyra (and eventually Athens and Sparta) were willing to involve themselves in the civil war in Epidamnus for the sake of maintaining their broader alliance structures.50 Athens was willing to risk war with Sparta and Corinth to preserve Potidaea due to the latter’s importance in the Athenian Empire.51 Perhaps the most telling example of the need to prevent civil war occurred during the Pentecontaetia, when Spartans feared that the Athenians may listen to the people of Ithome and become sponsors of “some revolutionary policy.”52 Their fears were realized when Athens exported democracy and directly intervened on behalf of the democratic faction in Samos and launched a civil war there.53

A nuanced interpretation of Thucydides’s central thesis reveals that Sparta’s policy vis-à-vis its allies, and the Spartan desire to prevent civil war and Athenian intervention within its allies, was central in causing the Peloponnesian War. While Thucydides’s pithy quote about the rise of Athens making war inevitable receives the most attention, he modifies this thesis later in Book One of History of the Peloponnesian War.54 He claims that the Spartans voted to break the treaty and declare war because they feared that Athens would only increase their now-dominant power over Hellas.55 Furthermore, Thucydides argues that Athens began to “encroach upon Sparta’s allies” when it reached the zenith of its strength. This, he argues, spurred Sparta to “employ all her energies in attacking and, if possible, destroying the power of Athens.”56 Thus, it was not the power transition itself that caused the Peloponnesian War. Rather, it was the point at which the power transition crossed a key threshold: when Sparta perceived Athenian power as a threat to its vital national interests.57 But why did Sparta care so much about its allies and their political stability, enough to risk a destructive interstate war with Athens?

The answer is that ancient Greek alliances likely involved more commitment from the member states than modern-day alliances. Whereas policymakers today discuss the issues of combined operations in coalitional warfare, Greek allies were expected to yield vast quantities of troops and tribute to the leaders of their respective blocs, whether it be Athens or Sparta.58 Alliances were a significant asset for the two great powers in the Hellenic world. As a result, these ancient alliances resembled confederations more than current alliances, especially as the two sides crystallized their blocs during the Pentecontaetia.59 This explains how the Athenians, through naval might, easily converted the alliance structure of the Delian League into the more centralized architecture of the Athenian Empire; the institutions for this broader Athenian state were already mostly in place due to the heavy commitment of Greek alliances. While Sparta did not do the same for its Peloponnesian League, it still used its alliance framework to consolidate an impressive sphere of influence over the Peloponnese.60 Thus, it can be argued that the Peloponnesian League more closely resembled a confederation than an alliance in modern terms. Contemporary scholars must avoid the mirror-imaging bias when evaluating the alliances of the Peloponnesian War—these alliances were far more robust than current alliance structures, which made the state-societal, organizational, and individual phenomena that informed these networks (e.g., those
that contributed to civil wars) key to belligerents.

Within these confederation-like alliances, not only was there potential for civil wars within city-states, but also between city-states and their broader bloc. This illustrates Mark Kauppi’s concept of “bi-multipolarity,”\(^6\) which describes multipolar alliance structures nested within (and informing) the dominant overall bipolar system. Bipolarity shaped the multipolarity beneath it, as each great power vied for influence over the smaller city-states. Each feared that their respective city-states would change allegiance to the other bloc, due to either a civil war within the city-state (e.g. six of the eight case studies in this paper) or due to a broader decision within the city-state to change allegiance (e.g. Potidaea and Mytilene). This prospect of multilayered civil war was an important phenomenon, as both great powers sought to avoid losing their alliance structures and the vast gains that they provided. This incentivized Athens and Sparta to obsessively watch over their spheres of influence and prevent convulsions across levels of analysis. For this reason, Athens and Sparta tried to stabilize the international environment during the Thirty Years’ Peace by crystallizing their alliance structures and preventing or disincentivizing states from switching allegiances.\(^6\) This strategy failed, and the Peloponnesian War began in 431 BCE.

Ultimately, the cause of the Peloponnesian War is closely tied to civil war and alliance networks. It was a combination of underlying systemic factors and factors at lower levels of analysis. The power transition from Sparta to Athens was necessary, but it was not sufficient in causing the Peloponnesian War. Power transition is dynamic; it is always occurring, as states become powerful for exogenous or endogenous reasons. No relative power position is static. Thus, if this power transition is a constant, it cannot be used to explain interstate war.

**Building the Second Half of the Model: The Importance of Civil Wars and Proxy Conflicts to State Interests**

According to the analysis thus far in this paper, it is not the power transition itself that causes hegemonic war, but rather the moment when this power transition affects certain policy objectives and national security interests set by actors at lower levels of analysis, whether individual, organizational, or state-societal. As demonstrated above, Sparta’s objective was maintaining the political stability of its allies and preventing civil wars within its bloc, and thus preserving its alliance structure. When Athenians intervened in civil wars like Epidamnus and Potidaea—and Spartan decision-makers felt that the Athenian Empire would further encroach on its alliances—Sparta went to war.

Here, we see that war is not triggered when the power transition approaches or crosses the level of parity; Sparta allowed Athens to grow in power for a long time, potentially to the point where Athens surpassed parity and became the region’s new hegemon.\(^6\) Rather, we observe that hegemonic interstate war is triggered when the power transition affects involved states’ objectives and interests, such as defending allies and protecting certain economic or political conditions. This indicates the importance of both underlying factors at the systemic level and factors at the lower levels of analysis. The interest of maintaining allied stability and avoiding civil war, as well as the decision to defend this interest, were just as important in causing the Peloponnesian War as the overall power transition, which threatened the interest in the first place.
As two of the eight civil war case studies considered in this analysis occurred during the Peloponnesian War, it is imperative to discuss the importance of alliance structures early in the war. Due to their advantages, these alliance structures were considered centers of gravity. During the Allied Congress at Sparta in 432 BCE, the process of fostering revolt among Athenian allies was seen as a viable Peloponnesian strategy to cut off Athenian revenue. Moreover, Sparta sent an ultimatum demanding the end of the Athenian Empire, demonstrating the Spartan objectives of overturning the status quo (which had yielded significant advantages to rising Athens) and fracturing the empire. This would set the stage for Spartan attempts to incite revolts during the war. While Athens drew power from its allies to exhaust the Spartans, Sparta’s strategy of annihilation involved “the attempt to bring about the dissolution of the Athenian Empire,” which “would be done through either encouraging apostasy or aiding revolts of Athens’ allies.” As the war escalated after the first few years, the two sides sought to aid friendly factions and political entrepreneurs within city-states of the other side’s bloc in an attempt to fracture it. Thus, the centrality of alliances and the fear of civil war shaped the early conduct of the Peloponnesian War, just as much as they shaped the causes.

As a result of this analysis, the second half of the model is as follows: civil war at the state-societal level of analysis reflects changes in allegiance and the overall power transition at the systemic level. When this power transition threatens a great power’s objectives and interests, set by decision-making entities at lower levels of analysis, the result is a hegemonic interstate war. The centrality of alliance shapes the subsequent conduct of that war, as both great powers seek to disrupt the other’s bloc by inciting civil war as an element of a broader grand strategy.

Conclusion

Convulsions of instability pervade all levels of analysis. As demonstrated by the model, systemic great-power competition changes the relative balance of power within smaller states. It shapes the incentive structures of individuals and organizations to launch a state-societal civil war, which accelerates and escalates great-power competition to potential levels of hegemonic interstate war. Of course, hegemonic war is not inevitable; on the contrary, the importance of lower levels of analysis within the model reveals that decisions could have been taken to avoid tension, such as shifting priorities away from alliance coherence and balance of power politics. Regardless, the confluence of factors across levels of analysis demonstrates that instability is like an avalanche: the more it ripples outwards from the source, the more momentum it acquires. The History of the Peloponnesian War demonstrates this, detailing how as immediate crises continued and civil war within the alliance structure became more of a threat, great-power war became increasingly probable.

The model presented in this paper can continue to cycle repeatedly once a hegemonic interstate war begins. Specifically, systemic great-power competition between Athens and Sparta in peacetime resulted in civil wars in Boeotia, Samos, Epidamnus, Macedon, Potidaea, and Plataea. Such instability affected the priorities of Athens and Sparta, resulting in a systemic hegemonic war which led to more state-societal civil wars at Mytilene and Corcyra early in the Peloponnesian War. The process would presumably continue from here, acting along the already-established channels of the model from systemic to state-societal back to systemic;
this is repeated until the instability is resolved. While beyond the scope of this paper, an avenue for further research would be an exploration of how this model continued to cycle throughout the Peloponnesian War, and perhaps how it was finally resolved at the end of the conflict.

This model easily complements what appears to be Thucydides’s true argument. As opposed to many accounts, Thucydides does not argue that the systemic power transition was the sole cause, nor that the war was inevitable; these assumptions in IR literature are based upon inaccurate translations. Rather, the power transition is the “truest cause,” influencing the more immediate causes. Indeed, as argued by Arthur Eckstein “Thucydides postulates a power-transition crisis… which finds expression in the ‘complaints and disputes’ between states.” This argument corresponds to the model presented in this paper. It was not the power transition in and of itself that caused the war, but the effect of the power transition on the objectives and interests decided at lower levels of analysis; it was the interaction and confluence of factors across levels, not an inevitable or deterministic process. In this case, the immediate factor was civil war, and it constantly interacted with the systemic instability wrought by great-power competition; the systemic instability incentivizes state-societal civil war and intrastate violence, which further accelerates the systemic instability and exacerbates the original competition. Thus, the systemic process indeed “finds expression in” the phenomena at lower levels of analysis, in this case, civil war. The systemic power transition is arguably the truest cause because it sets the conditions of instability for the entire system and the other levels of analysis, but alone the power transition cannot cause anything without interaction among variables at lower levels of analysis.

This model bears implications for present and future conflicts. The importance of state-societal factors demonstrates that civil wars and proxy conflicts may be significant hybrid spaces where the conditions of great-power competition and interstate war are set; this certainly seemed to be true in the Cold War, and it appears to be true again today. Indeed, while the Russo-Ukrainian War is not a civil war, it is a proxy conflict that was shaped by, and will subsequently shape, great-power competition. The Kremlin likely perceived the invasion as a viable strategic option due to Washington’s focus on great-power competition with China—thus, the potential power transition from the United States to China provided an apparently permissive environment for decision-makers in Moscow to launch the invasion. The same may be said about Hamas’s attack on Israel in October 2023, and Beijing may be similarly eyeing Taiwan in the near future. It is impossible to know the extent to which these current and prospective conflicts were shaped by systemic factors, especially so until decision-making records and other primary-source documents are eventually released. However, as the United States’s global hegemony continues to decline, further destabilization of the international system is likely. In turn, regional and local actors at lower levels of analysis may seize the opportunity to pursue their interests in ways that have been considered inconceivable since the end of the Cold War.

Per the second half of the model, these conflicts at lower levels of analysis have already begun to reverberate back to the systemic level. In response to ongoing wars involving (or threats to) U.S. allies and partners like Ukraine, Israel, and Taiwan, the Biden Administration has more forcefully expressed the United States’s commitment to defending its friends. If the PRC violates this apparent red line (e.g.,
invading Taiwan), this may be the violation of national interests necessary for the United States to launch a hegemonic war. But it is also possible that U.S. policymakers will not prioritize allies as much as Athens and Sparta did 2,500 years ago; after all, civil wars and proxy conflicts may not be as threatening as they were in the Hellenic world, where alliance architectures were more centralized and integrated. After all, the United States has not yet directly intervened on behalf of Ukraine or Israel, despite providing robust support. As such, some amount of pressure on its allies and partners is to be tolerated before triggering a military response. The mechanisms by which these smaller conflicts exacerbate great-power competition and contribute to a potential hegemonic war will depend upon how the United States and China (among other actors) define their vital interests.

Ultimately, the model presented in this paper is not meant to be a template applied rigidly to every situation; rather, it is a framework for conceptualizing and determining the pervasion of instability and war across levels of analysis, and how the processes at one level can accelerate the deterioration of another. While the specific dynamics may be different today, or in any conflict, the confluence of these factors and the channels across which instability travels are essential for students of war to understand.

About the Author: Christian Trotti is a Master’s candidate with the Security Studies Program (SSP) at Georgetown University. His research and published analysis focus primarily on defense strategy, military operations, and technological and organizational innovation in national security. At SSP, Trotti is also a Graduate Teaching Assistant for the core course “Strategy, Policy, and Military Operations,” a Research Assistant specializing in historical case studies of military operations, and the Associate Editor for Defense at the Georgetown Security Studies Review. He was formerly an Assistant Director of Forward Defense at the Atlantic Council, where he led its defense innovation and modernization portfolio.
1 Thucydides. *History of the Peloponnesian War*, trans. Rex Warner (London, UK: Penguin Books, 1954), 1:23. For all citations of Thucydides, this paper will follow the convention established by other secondary sources, in that it will cite the Book number and Passage number (rather than the page number), as follows: “Book #:Passage #.”
4 Thucydides, 1:23.
6 The levels-of-analysis framework in international relations scholarship was first advanced by Kenneth Waltz as a means to categorize different independent and dependent variables. Waltz describes three levels, or ‘images’: factors relating to the balance of power and interests among states are assigned to the systemic level of analysis; factors relating to the character, policies, behavior, and decisions of regimes and broader societies belong in the state-societal level of analysis; and factors relating to the idiosyncratic decisions and beliefs of individual leaders and decision-makers are captured by the individual level of analysis. He later added organizational levels between the state-societal and individual levels to describe factors relating to the institutional interests, organizational behavior, bureaucratic politics, and lobbying powers of different factions within or beyond government. See: Kenneth N. Waltz, *Man, the State and War* (New York, NY: Columbia University Press, 2001).
7 Thucydides, 3:82.
10 Levy and Thompson, 194.
11 Fearon and Laitin, 75.
13 Kalyvas and Balcells, 421.
15 Thucydides, 3:82.
16 Waltz, 99.
17 Thucydides, 1:75-1:77.
18 Levy and Thompson, 193.
19 Waltz, 99.
20 Thucydides, 1:113.
21 Thucydides, 1:112.
22 Thucydides, 1:115.
23 Thucydides, 1:24.
24 Thucydides, 1:24-1:25.
25 Thucydides, 1:57.
26 Thucydides, 1:56.
27 Thucydides, 1:56-1:58.
28 Thucydides, 1:58.
29 Thucydides, 2:2.
30 Thucydides, 3:2.
31 Thucydides, 3:3.
32 Thucydides, 3:2.
33 Thucydides, 3:70-3:81.
34 Lake and Rothchild, 44.
35 Kauppi, 110.
36 Thucydides, 1:24.
37 Thucydides, 1:101-1:104.
38 Thucydides, 3:82.
40 Thucydides, 3:82.
41 Thucydides, 3:82.
42 Kagan, 43.
43 Thucydides, 2:2.
44 Kalyvas and Balcells, 420.
45 Kalyvas and Balcells, 420.
46 Thucydides, 3:82.
47 Fearon and Laitin, 76.
48 Thucydides, 3:82.
49 Thucydides, 1:76.
50 Thucydides, 1:24-1:55.
51 Thucydides, 1:56-1:65.
52 Thucydides, 1:102.
53 Thucydides, 1:115.
55 Thucydides, 1:88.
56 Thucydides, 1:118.
57 Kauppi, 105.
58 Thucydides, 2:9.
59 Thucydides, 1:96-1:108.
61 Kauppi, 110.
63 Kauppi, 110.
64 Thucydides, 1:122.
65 Platias and Koliopoulos, 383.
66 Platias and Koliopoulos, 388.
67 Thucydides, 3:82.
68 Eckstein, 773.
69 Eckstein, 763.
70 Eckstein, 763.
Arms, Ideology, and Alignment: Analyzing U.S.-Soviet Realignment During the Ogaden War

Matthew Fiorelli

How does ideology interact with arms shipments to influence alliance formation? Using archival research, this paper analyzes the realignment of Cold War superpowers and warring states during the 1977 Ogaden War between Ethiopia and Somalia. During the conflict, Soviet advisors and military forces flew directly from Mogadishu to Addis Ababa in a stunning reversal of strategic and tactical support. Similarly, American leaders withdrew support from Ethiopia and realigned their strategic interests with Somalia. The effects of this rapid and complete realignment resulted in Ethiopian victory and contributed to the downfall of the American-Soviet détente.

This paper argues that ideology and arms shipments proved determinant in the realignment of patrons and protégés during the Ogaden War. This finding contradicts Stephen Walt’s theory of alliance formation, which suggests that states will only follow ideological preferences for alliances when they are already secure. Instead, by applying Mark Hass’s causal mechanism of conflict probability, I argue that growing gaps between the ideologies of protégés and their respective patrons led to realignment. International bribery, or the shipments of arms and munitions, also proved to be a causal factor—rather than a consequence—of realignment.

In a new era of great power competition, the Ogaden War shows that ideology is an important dimension of geostrategic competition. Competing with the People’s Republic of China (PRC) for alliance partners will require American leaders to highlight the ideological contrast between the United States and the Chinese Communist Party. The Ogaden War also reveals that, while arms shipments may help create alignments, they are ultimately a weak tool of superpower influence. However, when the interests and ideologies of a protégé nation overlap with those of a patron, providing military and economic aid is a valuable tool for strengthening an informal alliance.

Introduction

On July 13, 1977, at approximately 3 AM, Somali mechanized divisions launched an invasion into Ethiopia, targeting the Ogaden region, a sprawling area of approximately 125,000 square miles. Termed “the Ogaden War,” this conflict marked the culmination of Somali efforts to seize a contested region of Ethiopia populated by Somali people. Driven by irredentist claims, Somali leader Mohammed Siad Barre sought to unify “greater Somalia” for cultural, ethnic, and economic reasons.

Despite heavy fighting involving fighter jets, tanks, and artillery, Somalia’s forces rapidly advanced across the Ethiopian lowlands. However, after seizing 85% of the region within weeks, Somali forces faltered upon encountering hostile highlands. Overstretched supply lines and the superiority of the Ethiopian air force began to take a toll on Somalia’s army, which was not equipped for a protracted conflict. This stalemate dragged on for six months until a massive intervention by Soviet and Cuban forces gave Ethiopia the upper hand. By March 1978, Somalia’s defeated forces withdrew, having suffered 10,000 casualties compared to more than 20,000 Ethiopian casualties.
The dynamics and implications of the Ogaden War reached far beyond the Horn of Africa. The conflict played a critical role in shaping the conduct of the Cold War, just as the dynamics of the Cold War proved decisive in shaping the outcome of the Ogaden conflict. Yet, perhaps the most surprising aspect of the war is the realignment that occurred between competing great powers. A longstanding ally of the United States, Ethiopia relied heavily on American arms, equipment, and funding before the Somali invasion. Conversely, Somalia’s fighting force relied on Soviet weapons, funding, and training. At the start of the war, more than 4,000 Soviet military and technical advisors were stationed in Somalia. But, by the end of the war, Soviet advisors had been transported to Ethiopia where, along with 15,000 newly arrived Cuban soldiers, they proved decisive in defeating Somali forces.

The Soviet realignment was rapid and total, with top Soviet General Vasili Petrov—a longtime advisor to Somali forces—flying directly from Mogadishu to Addis Ababa, where he promised to bring the Somalis “to their knees.” American advisors and permanently stationed military personnel were likewise evacuated from Ethiopia during the conflict, and American foreign policy leaders realigned American interests with Somalia. The ultimate effects of this rapid and complete realignment resulted in Ethiopian victory and contributed to the downfall of the American-Soviet détente.

Why did this complete realignment occur? This paper argues that ideological shifts and “international bribery” proved determinant in the realignment of patrons and protégés during the Ogaden War. While ideology and arms shipments are viewed as two separate causal variables, their interactions are important in understanding realignment during the Ogaden conflict. Before the war, a military coup in Ethiopia overthrew the emperor and gave rise to a Marxist military junta. As ideological differences between the United States and Ethiopia grew, both sides became increasingly hostile toward one another and believed that their interests were bound to conflict. This led the United States to cease arms shipments to Ethiopia. Perceiving an opportunity for increased influence with an ideologically aligned regime, the Soviet Union began shipping arms to Ethiopia, replacing the United States as an arms supplier. Meanwhile, Somalia’s decision to invade Ethiopia, despite Soviet disapproval, revealed the strength of Somali nationalism and the weakness of Marxist socialism as a binding force between Somalia and the Soviet Union. This led the Soviet Union to cease arms shipments to Somalia, which then turned to the United States for increased aid. As the war progressed, the United States found its interests more closely aligned with Somalia and increased its economic aid to the country. Thus, the combined effects of ideological realignment and international bribery led to a rapid and complete strategic realignment.

This paper will proceed in three parts. First, it will provide a historical description of the Ogaden War within the context of the Cold War. Second, it will argue that, far from playing a subordinate role, ideology was a primary determinant in the patron-protégé realignment. This finding contradicts Stephen Walt’s theory of alliance formation, which suggests that states will only follow ideological preferences for alliances when they are already secure. By applying Mark Hass’s causal mechanism of “conflict probability,” I argue that growing gaps between the ideologies of Ethiopia and Somalia and their respective patrons led to a realignment. Having addressed ideology, I next turn to arms shipments as the second causal variable. I find evidence for Walt’s
“international bribery” hypothesis: that increased aid will lead to a tighter alliance. I also argue that, in the case of the Ogaden War, arms sales were more likely to be the cause of patron-protégé alignment, rather than merely a result of it. While this specific finding contradicts Walt’s theory, I find support for his prediction that “bribery” is a weak tool of superpower influence. I conclude by applying this case study to contemporary U.S.-Sino competition.

History of the Ogaden Conflict

The post-colonial Horn of Africa was home to five territories inhabited by Somali people. In 1960, two of these territories—former British Somaliland and former Italian Somaliland—gained independence as the Somali Democratic Republic. The other three territories include modern-day Djibouti, Kenya's northeast region, and Ethiopia's Ogaden region. In 1969, the elected president of Somalia was assassinated as part of a military coup. The resulting military junta, headed by Mohammed Siad Barre, embraced socialism and aligned itself with the Soviet Union. By 1977, the Soviet Union had access to naval facilities at Kismay, Mogadishu, and Berbera. However, from its inception, the central aim of Somalia's foreign policy was to reunite the five Somali territories in the Horn of Africa. Somali nationalism ultimately proved a much stronger political force than Somali socialism. Ethiopian nationalism, on the other hand, was not ethnic but territorial by nature. As the oldest independent country in Africa with a population ten times larger than Somalia, Ethiopians are ethnically heterogeneous. Yet, like Somalia, Ethiopians in the 1970s were fiercely patriotic and refused to consider ceding Ogaden to Somalia. Under the leadership of Emperor Haile Selassie, Ethiopia became a stalwart regional partner to the United States in the Horn of Africa. Ethiopia housed a major U.S. military and communications facility at Kagnew Station in the port city of Asmara. By 1973, Somalia’s intention to invade the Ogaden had become obvious to American diplomats in Ethiopia, who urged the government to increase military assistance to Ethiopia.

A significant setback to U.S.-Ethiopian ties occurred on September 12, 1974, when a group of senior Ethiopian military officers known as “the Derg” arrested 82-year-old Emperor Selassie and established a military junta. Two months later, the initial leaders of the reformist regime were violently replaced by the leaders of a radical socialist faction. Domestic pressure to cut spending abroad coupled with concerns over Ethiopia’s socialist ties to the Soviet Union led to a decrease in American support for Ethiopia in 1975, though the United States continued as Ethiopia’s primary supplier of arms and equipment. By December 1976, however, repeated efforts by Ethiopian leaders to attract Soviet support succeeded with the signing of a Soviet-Ethiopian military aid pact. Simultaneously, Soviet influence in Somalia continued to grow with Moscow building a naval and communications base for its fleet near Berbera. As the Ford Administration prepared to transition political power to Jimmy Carter, a U.S. Department of State review of U.S. foreign policy towards Ethiopia recommended the “middle course” of three options. Rather than continue the current level of support or disengage from Ethiopia entirely, the middle option proposed to “retain the potential for influence… to capitalize on future changes in Ethiopian leadership.”

By the time President Jimmy Carter took office in 1977, the pace of events in the Horn of Africa had rapidly accelerated
toward a major conflict between Ethiopia and Somalia. Somali President Barre, concerned by increased Soviet military aid to Ethiopia, requested U.S. military aid as a means of decreasing Somalia’s dependence on the Soviets. Yet despite decreased influence in Ethiopia, National Security Advisor Zbigniew Brzezinski felt that the United States had not “reached the point where we feel we should give up Ethiopia in exchange for Somalia.”

Faced with rapidly growing Soviet influence, the Carter Administration was initially unwilling to double down on Ethiopian support or abandon Ethiopia for Somalia. It was therefore decided that the United States would continue to provide economic assistance to Ethiopia while suspending all future military support for the government. This decision triggered an immediate backlash from the Ethiopian government, which announced that all U.S. military advisors, including the few remaining personnel at the naval base in Kagnew, would be expelled from Ethiopia. On May 6, 1977, Ethiopian leader Lieutenant Colonel Mengistu Haile Mariam publicly signed two friendship pacts with the Soviet Union and privately signed a military aid pact estimated at $400 million, which was more than the total amount of aid that the United States had provided to Ethiopia in the last three decades.

Meanwhile, Soviet efforts to “prod Ethiopia and Somalia into joining South Yemen in a federation of Socialist Red Sea states” failed due to strong pan-Somali feelings in Mogadishu. Seizing the opportunity to increase Soviet influence in Addis Ababa, the Soviet Union opted to support both Somalia and Ethiopia, hoping that it could leverage its influence to prevent a large-scale conflict. Somalia interpreted Soviet opportunism as a threat to Soviet-Somali ties and responded by again engaging the United States in hopes of receiving arms and equipment. The Somali Ambassador to the United States highlighted the growing Soviet threat as a compelling reason for American support to Somalia and emphasized the symbolic importance of direct arms shipments from the United States to Mogadishu. In response, President Carter promised to increase economic aid but expressed hesitancy in providing military aid to the Somalis. Carter explained his logic by stating: “we do not want to compete with the U.S.S.R. because we want Somalia to be non-aligned and not dominated by anyone.” Carter nevertheless promised to work with U.S. allies to equip Somalia with adequate defensive capabilities. This apparently satisfied Barre, who was finalizing preparations for the invasion of Ogaden.

On July 13, 1977, tens of thousands of Somali soldiers, supported by tanks and aircraft, invaded Ogaden against a severely outnumbered Ethiopian force. By August 3rd, American intelligence reported that Somalia controlled two-thirds of Ogaden. In a memorandum to Brzezinski titled “Where Do We Go With Somalia?” the national security staffer responsible for Africa, Paul Henze, summarized the impressive Somali feat:

The Somalis have played a wily game with everybody and, so far, have come out way ahead. They have built up a highly effective army with Soviet equipment and advice. They have put it to use against Soviet advice (at least so it seems) and are, in effect, blackmailing the Soviets into continuing military assistance to them… Claiming to feel threatened by Ethiopia, the Somalis extracted promises of aid, and some actual aid, from a wide range of countries who would like to see them draw away from their
friendship with the Soviets during the very period when they were putting the finishing touches on their plans for invading Ethiopia and seizing nearly a third of its territory. The Somali operation into Ethiopia is one of the most skillful the world has witnessed in many years.\textsuperscript{29}

Hamstrung by Carter’s commitment to avoid implicating the United States in any human rights violations, the administration had limited options with which to respond. Instead, administration officials hoped that Soviet opportunism would be “seen everywhere as self-serving and narrowly materialistic” in contrast with the American response.\textsuperscript{30} Unwilling to withdraw from the region completely, the Carter Administration found itself equally unwilling to leverage the all-important currency of arms shipments and thus was unable to promote its stated interests.

Throughout July and August, the Soviet Union continued to send arms and equipment to both Ethiopia and Somalia. However, it became increasingly apparent to the Soviets that nationalism, rather than a shared commitment to Soviet Marxism, was driving Somali decisions. Soviet leaders thus condemned the Somali incursion in September and ceased all arms shipments to Somalia in October.\textsuperscript{31} In response, President Barre renounced the Treaty of Friendship with the Soviets and expelled all Soviet and Cuban military and civilian advisors from Somalia.\textsuperscript{32} The Soviet Union also initiated a massive airlift to Ethiopia, transferring dozens of fighter jets and helicopters, 500 tanks, more than a thousand surface-to-air-missiles, as well as small arms, mortars, rockets, and artillery guns.\textsuperscript{33} Moreover, within the next four months, approximately 15,000 Cuban fighters and 1,500 Soviet advisors arrived to support the Ethiopians.\textsuperscript{34} From September 1977 to January 1978, Somali progress stalled as they failed to capture and hold the strategic Ethiopian cities of Harar and Jijiga. Equipped with newly arrived Soviet equipment and bolstered by Cuban fighters and Soviet advisors, Ethiopia launched a major counterattack in late January.\textsuperscript{35} Despite this massive Soviet intervention, the United States continued to rebuff Somali requests for arms shipments due to Carter’s resolve to embargo both countries until hostilities ceased and territorial sovereignty was restored. Ethiopia’s successful counterattack continued until March 23, when Addis Ababa declared that the entire Ogaden had been liberated.\textsuperscript{36}

As the conflict waged on, American policymakers increasingly feared that Ethiopia would invade Somalia in retribution. While this concern proved unwarranted, predictions that President Barres regime would not survive its defeat in the Ogaden eventually proved true. Struggling to attract Western assistance, Barre’s already damaged image suffered even more as he levied new taxes to offset decreased assistance from the West. Widespread discontentment sparked a civil war, eventually ousting Barre and leaving Somalia as a failed state for decades.\textsuperscript{37} Ethiopia, meanwhile, leveraged its new relationship with the Soviet Union to embark upon a socialist program of collectivization. Tragically, these reforms, combined with widespread drought, contributed to a severe food shortage that led to the death of one million Ethiopians by 1985. While Soviet influence proved decisive in determining who won the war, such influence provided relatively little strategic payout for the Soviet Union. The conflict had major implications for the Cold War, however. The political fallout undermined arms-control discussions between the United States and the U.S.S.R.,
leading Jimmy Carter’s national security advisor to later claim that “SALT lies buried in the sands of Ogaden.”

**Ideology and Realignment**

Political ideology played a decisive role in the realignment of patrons and protégés during the Ogaden conflict. Following the 1974 coup d’état, Ethiopian military leaders adopted a Marxist political ideology. This ideological shift led to greater ideological differences between the United States and Ethiopia. As a result, both sides ascribed increasingly hostile intent to each other, leading to a rift in the alignment. Simultaneously, greater ideological similarities between Ethiopia and the U.S.S.R created a perceived overlap in the political identities of both regimes, leading to increased trust between leaders and similar strategic interests. The Ogaden conflict also exposed the strength of Somali nationalism as the primary motivator of Somali foreign policy. Unable to leverage their influence to prevent Somalia’s invasion of Ethiopia, Soviet leaders realized the weakness of Marxist-Leninist ideology as a binding force in Soviet-Somali relations. These growing ideological differences between Somalia and the Soviet Union did not go unnoticed by American foreign policy leaders. Instead, American leaders capitalized on the Soviet-Somali rift and partially realigned America’s interests in the Horn with Somalia.

The impact of shifting ideologies on realignment during the Ogaden conflict squares with Mark Hass’s theory on ideology. Hass argues that “ideologies, or actors’ foundational principles of domestic political legitimacy, are likely to impact leaders’ foreign policies by affecting their perceptions of the threats that others pose to their central domestic and international interests.” He finds that leaders judge the threats that other states pose to their own state by examining how the other state organizes their society. Rooted in social identity theory, Haas’s causal mechanism, described as “conflict probability,” suggests that greater ideological differences between leaders of different states will lead them to “attribute hostile intent to one another and believe that their interests are bound to conflict.” This causal mechanism is particularly relevant to the realignment that occurred during the Ogaden conflict.

Ethiopian and Somali realignment with competing great powers began concurrently, though full realignment happened quicker with Ethiopia than with Somalia. Leading up to the 1974 Ethiopian coup, American leaders sought to bolster Emperor Selassie’s regime and prevent a Marxist military junta from seizing power. Discussing the growing instability of the Ethiopian regime with other senior American leaders, Henry Kissinger argued that the United States should “not remain passive and just accept the new government that emerges.” Thus, throughout 1974, the United States continued to supply Ethiopia with arms and equipment to strengthen the faction of Ethiopian military leaders who wanted to remain aligned with the United States. Ultimately, these efforts failed, and the newly installed Marxist military junta turned to the Soviet Union for both economic and military aid.

By the time Jimmy Carter took office in 1977, the ideological divide between Ethiopia and the United States had eroded trust and diminished shared interests. A 1977 report outlining the administration’s understanding of America’s interests in the Horn of Africa highlighted Ethiopia’s increasingly hostile perception of American ideology. In reference to Ethiopia, the report states:
The controlling members of the Provisional Military Administrative Council (PMAC), the supreme body in the Ethiopian Provisional Military Government (EPMG) which took power in 1974, view the U.S. with deep suspicion. Our quarter-century of close friendship and of generous support for Haile Selassie make it difficult for the PMAC to believe that the U.S. can sincerely desire a cooperative relationship with those who overthrew the Emperor….the PMAC has adopted what it calls a Marxist-Leninist path for the achievement of its social-reform goals, making it ideologically difficult for the PMAC to trust the motives towards itself of the world’s major capitalist power.  

The same report also says the following of Somalia:

Somalia’s leadership has a legitimacy problem, having overthrown an elected government, and has sought a similar method of solving it by espousing “scientific socialism” as the best path for realizing the people’s will for a better life… All of this would make it difficult for Somalia to affect a dramatic reduction of the Soviet presence there should it wish to do so. However, the prospect of a major Soviet military supply relationship with their Ethiopian foe has caused the Somalis to consider whether it might be in their national interest to readjust their relations with the USSR, and with the U.S.  

Thus, by early 1977, Somalia’s increasingly weak commitment to socialism was evident to American policymakers. It was also becoming obvious to the Soviet Union.

Efforts by Cuban leader Fidel Castro to bring Ethiopia and Somalia together failed as he found the Somalis “more irredentist than socialist.” Regarding Ethiopia, American policymakers recognized that Ethiopian leaders viewed American interests as hostile to their own interests. They also viewed American ideology as threatening to undermine Ethiopian legitimacy. As a result of this report, the Carter Administration ceased military aid to Ethiopia and began sending economic aid to Somalia. American diplomats also encouraged Saudi Arabia to offer military aid to Somalia to draw Somalia away from the Soviet Union. Both real and perceived changes in ideological differences between patrons and protégés catalyzed this realignment.

Mark Haas contrasts his emphasis on ideology with international relations realists, who deny “the importance of ideological variables to states' international choices.” Realists, such as Stephen Walt, argue that “although ideology does play a role in alliance choices,” its role is subordinate to considerations of power. Walt also finds that states will only follow ideological preferences for alliances when they are already secure, arguing that “one takes whatever allies one can get.”

Evidence from the Ogaden War suggests that, under certain circumstances, ideological considerations predominate security considerations. Ethiopian realignment from the United States to the Soviet Union was driven by ideological differences during a period in which Ethiopia’s position in the Horn of Africa was very insecure. Before the invasion, the Somalis outnumbered the Ethiopians militarily in virtually every category (battalions, tanks, aircraft, artillery, armor, etc.). In addition to the external threat, the Ethiopian government faced a domestic situation in which a revolving coup had
devolved into a “murderous power struggle.” Despite this dire situation, realignment was driven by shifting ideologies—both real and perceived—rather than by purely realpolitik considerations.

While Walt is generally dismissive of ideology in determining alliances, his nuanced analysis holds true on several fronts when studying the Ogaden War. First, Walt predicts that in a bipolar world—which he considers to be the most stable system—the importance of ideology may increase because all states are more secure. His logic is that “the caution that bipolarity imposes on superpower conduct may permit most other states to follow ideological preferences rather than security requirements.” In the context of the Ogaden conflict, this claim appears accurate when considering the realignment of Ethiopia. Both American and Soviet leaders sought to dissuade Somalia from invading Ethiopia. Moreover, Ethiopian leaders were offered arms and equipment from both superpowers and, therefore, had the luxury of choosing their alignment based on ideology. Had the Soviet Union stuck with Somalia and refused to align with Ethiopia, leaders in Addis Ababa would have been forced to remain aligned with the United States.

Walt also warns that the ideology hypothesis may become self-fulfilling. In reference to the Cold War, he points out that:

American beliefs in a Communist monolith led the U.S. to behave in ways that may have made the alliance of the Soviet Union and other leftist forces fare more cohesive than it would have otherwise been… the apparent importance of ideology in determining Cold War alignments may be less the result of “natural” Marxist solidarity than the naive American assumption that this was the case.

Evidence supporting this claim appears in America’s changing posture towards Ethiopia from the Ford administration to the Carter administration. During the Ford administration, Kissinger’s realpolitik prioritization of U.S. interests meant that America’s alignment with Ethiopia, though diminished, continued even after the Marxist coup. However, Carter’s concern for human rights and his emphasis on ideological considerations caused the United States to distance itself from Ethiopia even as Soviet influence in the Horn was growing. In summary, the Ogaden conflict provides evidence for some realist predictions regarding ideology. On balance, however, this case study suggests that ideology may occasionally feature prominently, if not predominantly, in determining alignments.

**International Bribery and Realignment**

Having accounted for shifting ideological differentials, this paper will now turn to the role of arms sales in the realignment of patrons and protégés. If ideology catalyzed realignment in the Ogaden, then arms shipments solidified it. According to a theory termed “international bribery,” providing economic aid or military assistance creates effective allies, “either by demonstrating one’s own favorable intentions, by evoking a sense of gratitude, or because the recipient will become dependent on the donor.” Walt rebuts this theory, arguing that it is erroneous to conclude that economic or military assistance is “the principal cause of alignment or a powerful tool of influence.” Evidence from the Ogaden conflict both refutes and affirms Walt’s argument. First, coupled with changing ideological differentials, a reorientation of military aid and assistance resulted in the realignment of patrons and protégés. This first piece of evidence appears to refute Walt’s hypothesis. However, the Ogaden conflict
also showed that, while arms may cause alignments, arms shipments are weak tools of superpower influence and, as a result, donors retain minimal leverage over recipients. This second piece of historical evidence affirms Walt’s hypothesis.

Walt argues that a large aid relationship is more often the result of an alignment rather than the cause of it. This may have been true in the Horn of Africa prior to realignment, when Ethiopia, under the reign of Haile Selassie, was a stalwart regional partner to the United States. Shared strategic interests, coupled with Ethiopia’s willingness to house American military forces in strategic locations on the Red Sea, led the United States to support Ethiopia with economic and military aid. Likewise, shared ideological interests, coupled with the housing of Soviet forces at multiple port cities in Somalia, led the Soviet Union to support Somalia with both arms and equipment during this period. However, as regional tensions between Ethiopia and Somalia escalated, both African countries placed increasing importance on the shipment of military arms and equipment. At the same time, a growing ideological divide between both sets of patrons and protégés put all alignment options back on the table. At the beginning of 1977, Ethiopia and Somalia were soliciting arms shipments from both superpowers. In this environment, military aid became a key cause of the alignment, rather than merely a result of it.

Following the 1974 coup, Ethiopia’s entreaties to the Soviet Union for military aid were initially rebuffed. But in May 1977, the Soviets offered Ethiopia $400 million worth of military equipment and publicly signed two friendship pacts. It was at this moment that Ethiopia completed its realignment from the United States to the Soviet Union. That same month, American military personnel and diplomats were expelled from Ethiopia following the revelation that Carter would cease all arms shipments to Ethiopia. The Soviet Union’s decision to send Ethiopia such a massive amount of military aid made it easy for Ethiopia to sever ties with the United States and fully align itself with the Soviet Union.

In contrast, Carter’s unwillingness to meet Somali President Barre’s request for increased military aid slowed Somalia’s realignment from the Soviet Union to the United States. For the first two months of the war, Somalia continued to rely on Soviet arms and equipment and maintained Soviet advisors in prominent positions of leadership within the Somali military. Despite growing Soviet support for Ethiopia, Somalia was forced to maintain its alignment with the Soviets. Ultimately, Brezhnev’s decision to cease arms shipments to Somalia decisively ended the longstanding Somali-Soviet alignment. With no other options, Somalia attempted to align itself fully with the United States and repeatedly pleaded for arms and materiel. Carter’s steadfast resolve to maintain an arms embargo on both countries prevented the United States from achieving most of its strategic objectives in the region and ceded influence to the Soviet Union. Nevertheless, increased economic assistance solidified the U.S.-Somali alignment, leading American policymakers to reprioritize diplomatic relations in the Horn. In contrast to Walt’s theory on alliance formation, arms shipments thus played a decisive role in the realignment.

However, the Ogaden conflict is not a complete rebuttal of Walt’s theory. Instead, the war provides evidence for Walt’s prediction that international bribery is, in fact, a weak tool of superpower influence. Ultimately, both American and Soviet leaders failed to leverage arms shipments as an effective means of attaining geostrategic
influence. Under Kissinger, American efforts to leverage arms shipments to bolster the pro-West faction of the Ethiopian military junta failed. Similarly, despite providing Somalia with military arms and assistance for years, the Soviet Union failed to deter Somalia from attacking Ethiopia. Walt argues that “unless the supplier is the only available source of economic or military aid, leverage will be limited because the recipient can always obtain it elsewhere.”\(^{54}\) This was clearly the case in both examples cited above. As Walt predicts, client states “can usually threaten to shift suppliers if their interests are not being served.” For the Soviets, failure to deter the Somali invasion can be partially attributed to the superior size and strength of the Soviet-built Somali army. This lends support to Walt’s claim that “the provision of aid can be self-defeating because it strengthens the recipient’s position and reduces the need to follow its patron’s wishes.”\(^{55}\)

** Applying Lessons from the Ogaden to Sino-U.S. Competition **

In summary, this paper finds that ideological differences between patrons and protégés presented an opportunity for total realignment during the Ogaden conflict. Having deposed the pro-American leader, Haile Selassie, the Marxist military junta in Ethiopia viewed American capitalism as hostile to its domestic and international aims and naturally aligned with the Soviet Union. By contrast, the Somali invasion of the Ogaden exposed the strength of Somali nationalism and the weakness of Somali socialism. This led the Soviet Union to cut ties with Somalia and forced Somalia to realign its interests with the United States. In addition to ideology, arms shipments proved to be a causal factor—rather than a consequence—of the realignment. Despite tremendous sunk costs in Somalia’s military, the Soviet Union decided to arm, train, equip, and even lead the Ethiopian counterattack against Somalia. Soviet and Cuban support enabled Ethiopian leaders to align with the Soviets rather than the Americans. For Somalia, the Soviet decision to cease all arms shipments forced Somalia to realign its foreign policy interests as it sought to receive support from the United States. For the patron regimes, arms shipments proved to be a weak tool of superpower influence and ultimately bolstered Ethiopian and Somali agency.

What does this case study teach us about ongoing Sino-U.S. competition? First, the Ogaden War shows that ideology is an important dimension of great power competition. In a book chapter titled “Contesting the Periphery,” Hal Brands highlights the importance of ideology in geopolitical competition in peripheral regions. Brands suggests that American leaders should deliberately align foreign policy with American ideals. Doing so promotes a long-term strategy that undermines the People’s Republic of China (PRC)’s political ideology in the global south\(^{56}\) and helps American leaders mobilize the domestic and international support needed to sustain a protracted strategic competition.\(^{57}\)

The PRC’s current strategy in the global south relies on the popularity of the “China Model,” which combines “political authoritarianism and economic capitalism, to prove to some African countries that economic development and political stability can triumph over the need for a sound democratic system.”\(^{58}\) To win, the United States must expose the China Model—and the neo-authoritarian foundation upon which it is built—as undesirable. The United States should offer nations an alternative method of economic and political growth rooted in both economic and political openness. This
will require American policymakers to highlight the ideological contrast between the United States and the PRC. The Ogaden War shows that ideology matters, and foreign policy leaders must account for this dimension of strategic competition.

The Ogaden conflict also reveals that, while arms shipments may occasionally help create alignments, they are ultimately a weak tool of superpower influence. Instead of using arms shipments to create strategic alignments in which the United States holds little leverage, American leaders should only commit to arms shipments when the recipient state’s interests are clearly aligned with American interests. Hal Brands argues that the United States should limit its strategic liabilities when contesting the PRC’s influence in the periphery. This means that the U.S. should generally avoid overcommitting resources, such as military aid, to decrease costs and preserve sustainable involvement. Yet, when the interests and ideologies of a recipient nation overlap with American interests, providing military aid is a valuable tool for strengthening an informal alliance.

Contesting and winning the periphery will require U.S. leaders to outmaneuver the PRC’s growing bid for global influence.

American leaders should account for the varied ways in which arms and ideology impact alignment choices. These lessons from the Cold War, some of which lie buried in the sands of the Ogaden, are worth uncovering and reexamining as American leaders craft a comprehensive strategy to defeat the PRC’s hegemonic aspirations.

About the Author: Matthew Fiorelli is an active-duty Army officer currently teaching in the Department of Social Sciences at the United States Military Academy. He graduated and commissioned from West Point in 2014 and received a Master’s in Public Affairs from Princeton’s School of Public and International Affairs in 2023. The views reflected here are his own and do not represent the official position of the United States Military Academy, the United States Army, or the Department of Defense.
Invasion began on July 13, dated the start of the war as July 23, 1977. However, more recent Ethiopian archival records indicate that the invasion began on July 13th.

In this paper, a patron is defined as a powerful nation that provides some form of assistance to a weaker state (protégé) in a manner that serves the patron state’s perceived interests.

Given the dynamics of the Cold War, neither the patrons- the United States and the Soviet Union, nor the protégés- Ethiopia and Somalia, were “secure” during this the period leading up to the Ogaden conflict.


In this paper, a patron is defined as a powerful nation that provides some form of assistance to a weaker state (protégé) in a manner that serves the patron state’s perceived interests.

Given the dynamics of the Cold War, neither the patrons- the United States and the Soviet Union, nor the protégés- Ethiopia and Somalia, were “secure” during this the period leading up to the Ogaden conflict.


In this paper, a patron is defined as a powerful nation that provides some form of assistance to a weaker state (protégé) in a manner that serves the patron state’s perceived interests.

Given the dynamics of the Cold War, neither the patrons- the United States and the Soviet Union, nor the protégés- Ethiopia and Somalia, were “secure” during this the period leading up to the Ogaden conflict.


31 Jackson, 29.


33 Darton, “Somalia Expelling All Soviet Advisers,” 200. The estimated value of Soviet weapons transferred to Ethiopia in less than a year exceeds $1 billion.

34 Darton, 204.

35 Tareke, 657.

36 Tareke, 660.

37 Sam Wilkins, “Buried in the Sands of the Ogaden: Lessons from an Obscure Cold War Flashpoint in Africa.”

38 Woodroofe, 2.


40 Haas, 11.


44 Document 7. Memorandum from the President’s Assistant National Security Affairs (Brzezinski) to Vice President Mondale; undated, https://history.state.gov/historicaldocuments/frus1977-80v17p1/d7.


48 Walt, 24.

49 Tareke, 638.

50 Walt, 24.

51 Walt, 26.

52 Walt, 27.

53 Walt, 27.

54 Walt, 28.

55 Walt, 29.


Reining in the Iranian Nuclear Threat: The Unviability of a JCPOA Revival and the Need for a New Game Plan

Logan Duarte

This paper explores the evolving dynamics of U.S. policy towards the Iranian nuclear threat by tracing the historical trajectory of U.S.-Iran relations from collaborative efforts during the Qajar and Pahlavi dynasties to the adversarial post-1979 Islamic Revolution era. With a particular focus on the 2015 Joint Comprehensive Plan of Action (JCPOA), this study examines the United States’ 2018 unilateral withdrawal from the JCPOA in favor of a “maximum pressure” campaign. In doing so, this study addresses the effectiveness of the maximum pressure campaign, the conditions that facilitated the original JCPOA negotiations, the possibility of restoring the largely abandoned deal, and potential future strategies to curb Iran's nuclear ambitions. Drawing on recent scholarly research and open-source data, the paper offers the following findings: (1) The U.S. decision to unilaterally withdraw from the JCPOA and apply a maximum pressure approach to Iran’s malign activities backfired, diminished U.S. credibility among the international community, and ultimately jeopardized U.S. national security interests; (2) The United States lacks the leverage and multilateral consensus it once did, making a JCPOA revival unlikely in the short-term. Additionally, even if a JCPOA revival were possible, there is reason to believe it would be largely ineffective; (3) Future U.S. policy should be regionally engaged, prioritizing diplomacy while leveraging the Intelligence Community (IC) and weighing covert action authorities as short-term shaping mechanisms to create the conditions for a long-term deal.

Introduction

During the years of the Qajar and Pahlavi dynasties, the United States and Iran maintained friendly relations, unfathomable in today’s hostile relationship. The birth of the Islamic Republic following the 1979 Islamic Revolution marked this transition point. Since then, Iran has made spreading anti-American sentiment among the Iranian population a core element of its rule. Echoed through the revolutionary mantra of “Death to America” and vows of revenge for the targeted killing of General Qassem Soleimani, this sentiment remains alive today. The Iranian Supreme Leader Ayatollah Ali Khamenei’s November 1, 2023 address is but one example of this, stating, “The situation between America and Iran is this: When you chant 'Death to America!' it is not just a slogan—it is a policy.”¹ The Iranian regime’s malign behavior—in particular, its rapidly advancing nuclear weapons program—threatens the national security interests of the United States, its partners, and allies by undermining global non-proliferation norms. Moreover, it fuels the fire of an already tense regional security dynamic in the Middle East. A sign of hope came from the 2015 Joint Comprehensive Plan of Action (JCPOA), commonly referred to as the “Iran Nuclear Deal.” Iran placed limits on its nuclear weapons program and expanded access to the International Atomic Energy Association’s (IAEA) monitoring and inspection practices in exchange for sanctions relief. However, President Donald Trump unilaterally withdrew the United States from the JCPOA in 2018, claiming it failed to protect U.S. national security interests and stating that it was “one of the worst and most one-sided transactions the United States has ever entered into.”² When
the United States withdrew from the JCPOA, Washington not only reinstated all previous sanctions lifted under the deal but unleashed a series of new ones as part of a “maximum pressure” campaign towards Iran.³

This paper is organized into three sections. First, the paper argues that the U.S. decision to unilaterally withdraw from the JCPOA and apply a maximum pressure approach to Iran’s malign activities backfired, diminished U.S. credibility among the international community, and ultimately jeopardized U.S. national security interests. Second, the paper attests that during the time of the JCPOA negotiations, the United States held leverage through a strong sanctions’ regime coupled with robust multilateral support. This leverage has since shifted in favor of the Iranian regime, making a JCPOA revival unlikely in the short term. Additionally, even if a JCPOA revival were possible in the near future, there is reason to believe it would be largely ineffective. The paper concludes with policy recommendations. Future U.S. policy adjustments should prioritize diplomacy by multilaterally engaging the region as an ecosystem rather than a geopolitical vacuum. This policy should be complemented by a robust counter-proliferation program utilizing the U.S. intelligence community (IC) and weighing covert action authorities as a short-term shaping mechanism to create the conditions for a long-term deal.

**Iran’s Nuclear Program from Inception to the Revolution**

The Iranian nuclear energy program began in 1957 after the United States and Iran agreed to a civilian nuclear cooperation arrangement, known as the Cooperation Concerning Civil Uses of Atoms, through the Atoms for Peace program.⁴ The Atoms for Peace program under the Eisenhower Administration aimed to assist developing countries in utilizing nuclear power for peaceful purposes. President Eisenhower, speaking at the United Nations General Assembly (UNGA) less than a decade after the United States dropped two atomic bombs on Hiroshima and Nagasaki during World War II, emphasized the need to use atomic and thermo-nuclear devices for “peaceful pursuits” and to “serve the needs rather than the fears of mankind.”⁵ The United States maintained a friendly relationship with Iranian Shah Mohammad Reza Pahlavi, referred to as “the Shah,” during his rule. As a result, the Shah became a benefactor of Atoms for Peace, which aligned with his vision for Iran. Akbar Etemad, the first chief of the Atomic Energy Organization of Iran (AEOI) and frequently referred to as the father of Iran’s nuclear program, articulated the Shah’s vision as aiming to achieve “an industrialized economy and high social standards.”⁶ This meant that Iran needed to resolve two main issues: “energy needs and acquiring high-tech industries.”⁷ Given that other countries that pursued a nuclear industry also achieved substantial advancements in other technologies (e.g., the United States and the Soviet Union), nuclear capabilities were a solution to both.⁸

Friendly relations came screeching to a halt in 1979 with the Islamic Revolution led by the exiled opposition leader Ayatollah Ruhollah Khomeini. The revolution deposed the Shah, replacing the monarchy with a hybrid Islamic Republic headed by the Supreme Leader and President. The President holds considerable authority as the second most powerful individual in Iran. Article 113 of the Iranian Constitution specifies that the President is responsible for “implementing the Constitution and acting as the head of the executive.”⁹ Of greater pertinence to the formulation of the country’s nuclear policy, Article 176 names
the President as chairman of the Supreme Council of National Security. However, under the theory that “political authority derives from religious authority,” the Supreme Leader is the “guardian jurist” whose ultimate religious and political authority guides that of the President. Delineated under Article 110 of the Iranian Constitution, the Supreme Leader’s extensive list of authorities includes supreme command of the armed forces, the declaration of war, and dismissal of the President. Despite the President’s chairmanship of the Supreme Council of National Security, the Supreme Leader has the final say on all national security and foreign policy matters, including nuclear agreements with foreign powers.

All nuclear cooperation with the United States ceased when Ayatollah Ruhollah Khomeini came to power. Khomeini canceled pending nuclear contracts initiated under Shah’s rule. He argued the nuclear effort was “Western” and ran counter to the Islamic principles behind the revolution. This skepticism of nuclear cooperation, however, was not unilateral. Reluctant to trade with the new Iranian regime, the United States and other foreign nuclear suppliers also barred civilian nuclear cooperation with Iran. The United States’ imposition of sanctions on Iran in November 1979 amidst the Iran Hostage Crisis and Washington’s support for Baghdad in the 1980 Iran-Iraq war further soured U.S.-Iran relations. Within a few short decades, the same nuclear research reactor that the United States supplied Iran in 1967 for “peaceful pursuits” was used instead for plutonium reprocessing experiments and Polonium-210 production. This seemingly negligible 5-kilowatt research reactor laid the groundwork for present-day capabilities such as the 1000-megawatt Bushehr Nuclear Power Plant and Natanz Enrichment Complex.

Though these watershed moments in U.S.-Iran relations contributed to Iran’s growing ambition for a nuclear deterrent, they were not the primary driver. Saddam Hussein’s 1980 invasion of Iran and use of chemical weapons during the war catalyzed, to a greater extent, Iran’s desire for nuclear weapons. U.S. sanctions and lack of broader international support for Iran during the war intensified the Iranian sentiment of being alone in defending itself against the existential threat its neighbors posed. From Tehran’s perspective, this international isolation demanded a solution that would ensure another existential crisis like the Iran-Iraq War would not happen again. Nuclear weapons pursuits became one such solution.

U.S.-Iran Policy: The JCPOA and the Maximum Pressure Mistake

The sanctions employed during the 1979-81 Iran Hostage Crisis also marked a pivotal point for U.S. policy towards Iran. Subsequent administrations have since utilized sanctions in their Iran policy. The United States did not issue its first significant easing of these sanctions against Iran until 2016, when the IAEA certified that Iran complied with the standards agreed to in the JCPOA. Iran consistently demonstrated a willingness to circumvent global non-proliferation norms throughout the Obama Administration. For example, in September 2009, a U.S.-involved Western intelligence effort unveiled an Iranian uranium enrichment facility, which it covertly attempted to conceal from IAEA inspection requirements. In February 2010, Iranian President Mahmoud Ahmadinejad increased uranium enrichment from 3.5% to 20%. Despite insisting that the increased enrichment was for peaceful purposes, his rhetoric indicated otherwise. Less than a week later, he made the following statement
in an address commemorating the anniversary of the Islamic Revolution: “The Iranian nation is brave enough that if one day we wanted to build nuclear bombs, we would announce it publicly without being afraid of you.”

The 2015 nuclear deal brought relief to the growing Iranian nuclear threat that plagued the Obama Administration. Iran largely complied with the agreement and limited its nuclear program’s expansion. This progress began to erode after President Trump unilaterally withdrew the United States from the agreement on May 8, 2018. In substitute, he installed a “maximum pressure” campaign of hardline sanctions and escalatory rhetoric. Beginning in July 2019, the IAEA verified that some of Iran’s nuclear activities exceeded JCPOA-mandated limits. This period of rising tensions emboldened Iran’s willingness to ramp up the program and use its advancement toward a nuclear weapon as a bargaining chip, continuing this strategy to this day. IAEA data on Iran’s stockpile of low-enriched uranium confirms this through a significant decline in Iran’s stockpile during the years of JCPOA compliance, followed by renewed expansion efforts shortly after the United States’ withdrawal.

Former President Trump’s decision to withdraw from the JCPOA and apply a “maximum pressure” approach to Iran’s malign activities backfired. Not only did it diminish U.S. credibility among the international community, it ultimately jeopardized U.S. national security. Trump’s hardline approach held two primary objectives: (1) prevent Iran from advancing its nuclear weapons capability; and (2) force the Iranian regime back to the negotiating table to set new terms—one in which the United States would hold greater leverage due to its overwhelming sanctions on Iran. These goals backfired in unison. Instead of striking a new deal that would result in the reduction of Iran’s nuclear capacity, Iran responded by increasing its nuclear capacity. By January 2020, Iran increased its high-enriched uranium (HEU) production to 20 percent, nearly seven times the 3.67% enrichment stipulated under the JCPOA. By the end of the Trump Administration, after continued escalatory rhetoric and saber-rattling by both the United States and Iran, this had increased to 44 percent or twelve times that which was permitted under JCPOA.

No JCPOA signatories endorsed Trump’s maximum pressure policies. The only global support came from Israel and some Arab Gulf states like the United Arab Emirates (UAE). However, toward the latter end of the Trump Administration, these Gulf partners began engaging in de-escalation attempts with Iran, given its fevered response. These de-escalation efforts demonstrated the severity of the situation and the regionally destabilizing effects of the maximum pressure campaign. Additionally, the lack of international support significantly contributed to the ineffectiveness of the maximum pressure campaign. Not only did Trump’s approach make sanctions evasions easier, but unilateral economic pressure had also not reined in Iranian malign activity.

A Game of Leverage: Feasibility of a JCPOA Revival

The counter-productivity of the maximum pressure campaign set back years of diplomatic progress and changed U.S.-Iran relations. This setback leaves one wondering if a JCPOA revival is possible, what conditions between negotiating parties existed at the time that permitted its success, and if they still exist today. The successful negotiation of the JCPOA hinged upon one major factor: leverage.
The United States, along with allies and partners, held the leverage during the JCPOA negotiations due to two principal factors. Firstly, as previously mentioned, since 1979, the United States had sanctions programs targeting Iran’s economy, which were quite effective. Beginning in 2010, the Obama Administration increased the scope of U.S. sanctions by targeting Iran’s oil exports and other economic sectors. Many experts attribute Iran’s decision to enter into the JCPOA negotiations as a way to manage the negative impact of these expanded sanctions on Iran’s economy. Secondly, the United States also had strong multilateral support for isolating Iran in the years leading up to the JCPOA. In addition to the United States, the deal included the five permanent members of the UN Security Council (China, France, Russia, the United Kingdom, and the United States) plus Germany—collectively known as the P5+1—and the European Union. This powerful coalition carried weight, which helped bring Iran to the negotiating table. The same coalition is not likely to come back together again today. European partners were very angry with the United States’ 2018 withdrawal. Brian O’Toole, nonresident senior fellow with the Atlantic Council’s GeoEconomics Center, described the European dissatisfaction, stating: “While President Trump was overseeing a scorched earth sanctions campaign against Iran, he was doing roughly the same with respect to traditional U.S. alliances, especially those with Germany and the North Atlantic Treaty Organization (NATO).” Furthermore, even with European partner support, the current state of global affairs makes it unlikely that the United States can convince China and Russia to push Iran to the negotiation table like it did in 2015.

Today, Iran has furthered its enrichment of uranium to 60 percent, with some particles found as high as 83.7 percent purity, just shy of the 90 percent threshold for weapons-grade uranium. It is worth mentioning that a September 2023 quarterly IAEA Iran Verification and Monitoring Report indicated that “the amount of 60 percent highly enriched uranium (HEU) produced during this most recent reporting period was about half of the amount produced during the previous reporting period.” However, the fact that Iran has made little to no progress on outstanding issues raised by IAEA’s monitoring of its nuclear facilities taints these positive findings. Namely, Iran has yet to address traces of nuclear materials found at two unauthorized sites. Despite the reported September 2023 slowing, Iran’s nuclear program is as advanced as it has ever been and does not show a desire to turn back now. The latest December 26, 2023, IAEA report evidences this reversal. The report highlighted that Iran increased its rate of production of uranium enriched to 60% at Natanz and Fordow nuclear facilities between January and June 2023. As a result, Iran will be able to triple its monthly production rate of uranium enriched up to 60%. The U.S. Department of State condemned the increase in a joint statement on December 28, 2023, alongside the governments of France, Germany, and the United Kingdom. The statement called Tehran’s decision a “backward step” which “carries significant proliferation-related risks.” This production increase is the latest action in the series of steps Iran has taken to roll back its nuclear commitments under the JCPOA and further increases Tehran’s leverage at the negotiating table. The more advanced Iran’s nuclear program becomes, the higher Iran’s demands will be in exchange for dialing it back. Additionally, the closer Tehran gets to its goal of achieving a nuclear deterrent, the less willing it will be to give it up. Every step Iran takes to enhance its nuclear program will further increase already heightened
U.S.-Iran tensions and make any attempt at a JCPOA revival unlikely to materialize—at least in the short term. The United States must first address this leverage gap to meaningfully negotiate a long-term nuclear deal with Tehran.

Additionally, even if a JCPOA revival were possible, there is reason to believe it would be largely ineffective. According to State Department officials, even if Tehran were to resume implementing its 2015 JCPOA obligations, its fissile material production timeline would be less than one year. As explained by Jon Wolfsthal, a National Security Council official during the Obama Administration, this one-year breakout goal is meant to provide enough time “to generate an international response to any Iranian move to build weapons.” Given that Iran’s nuclear program has now reached the point at which it could enrich enough uranium for five fission weapons within as little as three weeks, a JCPOA revival is not currently a viable solution.

**Considerations for Future Policymaking**

If reviving the JCPOA is dead on arrival, the question then becomes: What future U.S. policy adjustments would help to rein in Iran’s nuclear weapons program? A new deal is necessary because unilateral sanctions from the Trump Administration’s maximum pressure campaign were not only ineffective but counterproductive, allowing Iran to overtly ramp up its nuclear program efforts. The United States must put diplomacy at the heart of a new strategy towards Iran. It must engage the region as an ecosystem rather than a geopolitical vacuum, a key missing piece of the Trump Administration’s maximum pressure campaign. The maximum pressure campaign and Iran’s escalatory response increased instability in the Middle East. Recent developments such as the Israel-Hamas War and Iran-sponsored proxies firing upon U.S. troops in Syria and Iraq also add to this instability. However, the present situation also creates incentives for some countries in the region to pursue avenues for de-escalating tensions. The United States must recognize this critical diplomatic opportunity and pursue engagements with regional stakeholders as a pathway to reopen talks. A December 2021 report from the Center for American Progress reiterated the importance of this point, stating, “America’s strategic reengagement in the Middle East requires a balanced approach that seeks to address the threat perceptions of key partners and de-escalates tensions across the region...Without an improved regional security environment, any future deal with Iran will not last long.”

U.S. policy should also be complemented by a robust counter-proliferation program utilizing intelligence collection and covert action when appropriate. Covert sabotage operations, including cyber operations, have been effective to varying degrees in disrupting and delaying Iran’s nuclear program. The Stuxnet cyber-sabotage operation of the Iranian Natanz nuclear facility—first uncovered in 2010 and unofficially attributed to a U.S.-Israeli partnership—is one such example. Estimates report that the Stuxnet malware disabled 1,000 of the total 9,000 centrifuges at Natanz. Additionally, as David Sanger argues in his *Confront and Conceal*, an additional goal of Stuxnet was to embarrass the Iranians. Some of Iran’s most renowned scientists and engineers were dismissed as incompetent, unable to explain what was happening to the malfunctioning industrial control systems at Natanz. The malware was only later discovered by a non-Iranian security researcher, another blow demonstrating Iran’s inability to protect its own networks from penetration. However, decision-makers—in this case namely the
U.S. President—must understand the risks of utilizing covert action authorities.

Covert operations always run the risk of surfacing or causing more damage through unintended second and third-order consequences. For example, Stuxnet was a catalyst for Iran’s decision to heavily focus on developing a national cyber capability. In April 2015, Andretta Towner, a Senior Intelligence Analyst at CrowdStrike, stated that Iran’s budget for cyber security increased 1,200 percent from 2012 to 2015.48 Fast forward to the present day, Iran’s cumulative experience in the cyber domain represents a relatively high level of operational maturity, and the regime firmly embraces cyber operations as an instrument of national power.49 The U.S. Intelligence Community 2023 Annual Threat Assessment called Iran’s growing expertise and willingness to conduct aggressive cyber operations “a major threat to the security of U.S. and allied networks and data.”50 While Stuxnet was a tactical success by disrupting Iran’s nuclear program, Iran’s response to develop cyber capability was an unexpected consequence, taking away from the operation’s success at the strategic level. Unintended strategic level consequences make covert action, including offensive cyber operations, ideal as a short-term shaping mechanism. However, if the United States aims to stop Iran from developing nuclear weapons, a long-term deal brokered through diplomatic engagement is necessary.

Robert M. Gates, former Director of Central Intelligence and Secretary of Defense, in a declassified 1981 memorandum to then Director of Central Intelligence William J. Casey, provided important criteria for the use of covert action. He stated that “covert operations are best suited to tactical situations where success will bring an immediate short-term gain” and that it is “a tactic to be repeated or expanded upon in the absence of the ingenuity, will, or money to come up with a viable long-term overt option.”51 The aforementioned lack of negotiating leverage on the part of the United States to secure a viable deal with Iran in 2023 reflects this absence. The current situation meets Gates’ 1981 criteria, given Iran’s current leverage and the unlikely circumstances in which Congress or the White House will be willing to make concessions to Iran in the lead-up to the 2024 elections. However, Gates also cautioned that “covert operations can rarely achieve an important objective alone.” Considering the Stuxnet example, decision-makers should not rely solely on intelligence collection and covert action as a long-term solution to rein in the Iranian nuclear threat.

Covert operations offer high-risk-high-reward outcomes and present an opportunity for taking direct action to delay, disrupt, damage, or deny Iranian nuclear capabilities. However, they cannot directly change Iranian compliance with global non-proliferation norms. Securing long-term Iranian compliance with these norms requires a new game plan to overcome the existing leverage gap and lack of international consensus. Firstly, the United States must continue to engage regional stakeholders in efforts to de-escalate tensions. Additionally, the United States must continue intelligence collection efforts while carefully weighing the use of covert action authorities to overcome the current leverage gap in any potential U.S.-Iran negotiations. Leveraging the IC will ensure diplomats and policymakers have a decision advantage in their work, but addressing these factors precludes any hope of a favorable diplomatic outcome in the near future.
About the Author: Logan Duarte is a graduate student in Georgetown University’s Security Studies Program, concentrating in Technology & Security. He is currently a Legislative Research Assistant in the U.S. House of Representatives, supporting the work of a senior member of the House Foreign Affairs Committee (HFAC). A former high school Portuguese teacher, Logan received his B.A. in Portuguese and an M.A. in Teaching (MAT) from the University of Massachusetts Dartmouth. In his free time, he is a published poet, most recently featured in the new book “Into the Azorean Sea: A Bilingual Anthology of Azorean Poetry” and in Gâvea-Brown (Brown University).
29 Thomas, “U.S. Sanctions on Iran.”
33 O'Toole, “Rejoining the Iran Nuclear Deal: Not So Easy.”
37 U.S. Department of State. “Joint Statement on the Latest Iranian Nuclear Steps Reported by the IAEA.”
38 U.S. Department of State. “Joint Statement on the Latest Iranian Nuclear Steps Reported by the IAEA.”
39 O'Toole, “Rejoining the Iran Nuclear Deal: Not So Easy.”
40 Kerr, Paul K. “Iran and Nuclear Weapons Production.”
41 Kerr, Paul K. “Iran and Nuclear Weapons Production.”
The Double-Edged Sword of Diplomatic Immunity: The ICJ and the Case Studies of Germany v. Italy and Mohammed bin Salman

Apurva Ramakrishnan

International law is an important pillar of the international system, and international organizations frequently serve as an enforcement mechanism for these rules. However, there is rarely a concurrence between them. This paper examines the two components of jurisdictional immunity—state immunity and sovereign immunity—with the help of the case studies of Germany v. Italy and the murder of Jamal Khashoggi by then Saudi Crown Prince Mohammed Bin Salman (MBS) to assess the dichotomy that exists between law and practice. The case studies highlight two distinct applications of diplomatic immunity and prove that loopholes exist within the legislation that powerful states may exploit for their own gain. The paper further argues against the effectiveness of international organizations such as the International Court of Justice (ICJ) in enforcing their laws. Instead, the paper establishes that the enforceability of international laws is highly subjective, and that the moral ambiguity of the states leads to uncertainty in the international system. Ultimately, this paper finds that there is no better alternative for the existence of international laws; however, better ways of enforcing them through international legislative reform and policy amendments need to be devised.

Introduction

In the aftermath of one of the most disastrous global conflicts, the Second World War, states across the globe came together with a common consensus to establish an international community governed by international law. According to Article 39 of the United Nations (UN) Charter, the UN’s main function was to promote interstate peace and harmony. To do so, it derives its body of rules and norms from international treaties, customary law, and general principles of International Humanitarian Law, along with judicial decisions laid down by both national and international courts. Though, in theory, UN member states consent to abiding by and enforcing international laws, state behavior is largely motivated by their respective national interests, making law enforcement by an external legislative body challenging. Therefore, in addition to international law emphasizing the consent of states, the UN also recognizes that national sovereignty is of unparalleled importance to each state.

This paper asserts that the effectiveness of international law is subjective and contingent on state morality and interests. Examining the concept of diplomatic immunity, the paper delves into the contrasting outcomes of two case studies and the role played by the International Court of Justice (ICJ) to argue for the necessity of periodically updating the frameworks of international law to ensure its relevance amidst new geopolitical and social contexts.

One of the most prudent concepts in international law is that of ‘diplomatic immunity.’ It is the principle that legally upholds the sovereign jurisdiction of a state over itself and its subjects, meaning that certain officials are exempt from the jurisdiction of courts outside their country of citizenship. The Vienna Convention on Diplomatic Relations (VCDR) of 1961 is an international agreement between sovereign states that outlines the rights and privileges of diplomatic missions and gives them the freedom to carry out their duties without...
hindrance or interference from the host state. As per the United Nations Convention on Jurisdictional Immunities of States and Their Property, “a state enjoys immunity, in respect of itself and its property, from the jurisdiction of the courts of another state subject to the provisions of the present Convention.” However, the Convention also clearly delineates the cases in which states may not be subject to such immunities. Article 12 of the Convention talks about “personal injuries and damage to property,” wherein a state cannot invoke immunity against proceedings in the host country in the case of injury or death.

Diplomatic immunity is a broad concept covering several types of immunity that a state can invoke based on its diplomatic relationship with other states. This paper will examine two facets of diplomatic and jurisdictional immunity: state immunity and sovereign immunity. While often confused as tangential, “state immunity” derives itself from the ‘Act of State’ doctrine, which in essence, declares that no state can come under legal proceedings or receive judgment from a foreign government for acts committed in its own territory. On the other hand, ‘sovereign immunity’ is the tenet that states that no sovereign state or head of such a state can be subject to a lawsuit in any court without their consent. This indicates that the head of the state is immune not only to any legal proceedings but should a legal proceeding take place, is also immune from the enforcement of any related verdicts.

**Case Studies Background**

*“The Act of State”: Germany v Italy*

Between 2004 and 2008, several Italian courts rendered a series of rulings awarding compensation to plaintiffs who were the victims of war and human rights violations perpetrated by the German Reich throughout WWII. Finally, in 2008, Germany filed a suit against Italy in the ICJ, claiming that Italy had violated international law because it permitted Italian courts to file civil claims against Germany, in clear violation of Germany’s jurisdictional immunity that it derived from international law. Germany also requested the court to rule that Italy had violated its immunity by imposing restrictions on Villa Vigoni, a German state property located in Italy. In making this argument, Germany did not contest the facts of Italy’s case itself but argued that the country’s case was invalid under international law. Italy had breached the rule that “one state cannot and should not exercise jurisdiction over (the acts of) another state. This principle is based on the notion of sovereignty and, thus, the (legal) equality of all states.”

In 2012, the ICJ ruled that the actions of Italy were a clear violation of international law. Given that both Germany and Italy ratified the UN Treaty on the Settlement of Disputes and the European Convention on the Peaceful Settlement of Disputes, the restriction imposed on Villa Vigoni was a violation of an international treaty. Finally, since Villa Vigoni was property used for government and not commercial purposes, the restrictions imposed were also a violation of the United Nations Convention on Jurisdictional Immunities of States and Their Property.

*Sovereign Immunity: Mohammad bin Salman and the Jamal Khashoggi Dilemma*

Jamal Khashoggi was a Saudi Arabian journalist who was known to be overtly critical of the Saudi regime, especially of the then Crown Prince Mohammed bin Salman’s (MBS) allegedly liberal policies, which Khashoggi believed masked his actions as a despotic human rights abuser. When MBS began to crack down on those who critiqued him, Khashoggi went on a self-imposed exile to the United States,
where he wrote for the Washington Post. In 2018, Khashoggi visited the Saudi consulate located in Istanbul to secure official divorce papers so he could marry his current fiancée. Surveillance footage showed him entering the consulate, but he never left.\footnote{15}

Initially, MBS denied any knowledge of his whereabouts, and the Embassy claimed that Khashoggi left through the back door. However, shortly after, Saudi Arabia confirmed his death and blamed 15 or so rogue agents. MBS denied all involvement in the murder, although he took personal responsibility for the role of Saudi officials in the killings.\footnote{16} A Saudi court held a trial for 11 accused members, of whom eight were found guilty, and five were sentenced to death.\footnote{17} An investigation conducted by the Central Intelligence Agency revealed audio recordings that took place at the time of the murder, and information was uncovered that the assassination of Jamal Khashoggi was authorized by a high-ranking individual within the Saudi government—all fingers pointed toward the Crown Prince.\footnote{18} This information was further corroborated by a now-declassified report by the Office of the Director of National Intelligence, which stated that MBS had a definitive role in authorizing the capture and execution of Khashoggi.\footnote{19}

A series of lawsuits were initiated against MBS in Turkey, Germany, and the United States, among others. A petition was also filed against him in the International Criminal Court. However, no judgment has yet been passed because MBS refuses to participate in the legal proceedings, and he enjoys legal sovereign immunity as the titular and diplomatic head of Saudi Arabia.\footnote{20} Fearing the potential souring of diplomatic relations as a key ally and trade partner, he was also granted immunity by countries like the United States.\footnote{21} While Saudi Arabia has the right to invoke the UN Statute on Sovereign Immunity, there is no denying that there have been several breaches of international and humanitarian laws in this case.

**Analysis**

The laws surrounding diplomatic immunity are one of the major tenets of diplomacy and are arguably one of the oldest practiced foreign policy measures among states. Some of the clauses, such as ones against the persecution or harm of diplomatic agents, derive from customary law and are an integral part of humanity.\footnote{22} With the advent of global liberalization, the notions of sovereignty and consent started gaining more traction and eventually became inseparable components of international law. This paper delves into the legal principle of immunity because it is perhaps one of the most studied and debated ones, with no clear outcome on how it should be exercised. In the case of Germany v. Italy, diplomatic immunity seemed justified, but for Khashoggi, the law shielded a suspected murderer from the consequences of his actions.

The biggest advantage to studying the contrasting outcomes of Germany v. Italy and Khashoggi’s murder is that it enables a deeper understanding of the workings of international laws in the real world. Both cases have vastly different circumstances. In the Khashoggi case, one person—the ruler of a sovereign state—was accused and summoned before the ICJ. However, in the case of Germany v. Italy, we look at the state itself in its entirety. Ultimately, however, this paper intends to prove the dichotomy that exists between the effectiveness of international law and the international organizations that exist to protect and promote it. In the case of Germany v. Italy, the ICJ ruled in favor of Germany, highlighting an outcome where the application of international law was
upheld, emphasizing respect for state sovereignty and treaty obligations. Conversely, in the Jamal Khashoggi case, despite compelling evidence of human rights violations and murder allegedly authorized by a high-ranking official, the international response was characterized by a lack of legal accountability. This divergence in outcomes illustrates a pattern where the effectiveness of international law becomes contingent on factors beyond legal considerations, reflecting the complex interplay between state interests, diplomatic relations, and the limitations of international organizations.

Jurisprudence of International Organizations: The International Court of Justice

As a body established based on the consensus among states, the ICJ, in practice, serves more as an advisory body than it does as a formal court. There is no denying the fact that in the aforementioned cases, there were several violations of state sovereignty. The commonly propagated defense of considering the embassy as a territory of Saudi Arabia still negated the sovereignty of the receiving state, i.e. Turkey, as per the Vienna Convention on Diplomatic Relations. Additionally, the American-authorized investigations into the death of Khashoggi conducted both in Turkey and elsewhere was a clear breach of the sovereignty of the Kingdom of Saudi Arabia, despite the then-U.S. President Donald Trump’s assertion on the importance of Saudi-U.S. relations. Amidst all this, as an international organization, the ICJ faced a dilemma, as the United Nations and all its constituent organizations had an obligation to uphold and promote the Responsibility to Protect (R2P) doctrine. The R2P doctrine held that if a state was unable to protect its population, the international community, through the UN and its organizations, had the responsibility to intervene using diplomatic, humanitarian, and other means. By failing to take appropriate action against the matters brought before it, the ICJ risked setting a precedent for inaction against human rights violations and failed in its duties too.

Herein lies the dilemma—since the foundation of the UN is based on mutual respect for sovereignty, the ICJ’s ability to hear cases is also limited by the principle of state sovereignty, which gives states the right to choose whether or not to participate in ICJ proceedings. This has led to situations such as the one with Khashoggi’s murder, where states can simply refuse to accept the jurisdiction of the ICJ or abide by its decisions, and nothing will come of it. This could be why the clause on consent and immunity gained the upper hand over all other concerns presented before the court. Ultimately, despite being an international apex court (i.e., the highest international legislative body), it is still just a body of states with no inherent source of power. Therefore, the ICJ has no independent power nor any independent influence on the outcome of events.

Justifying the Use and Need for International Law

The idea of an international legal system came about as a measure of checks and balances and as a way to promote a peaceful coexistence based on mutual respect. The law protecting the jurisdictional immunity of sovereign states was also based on such principles. Given the aforementioned facts, the concept of an international organization forcefully enforcing its decisions is infeasible, although individual states have taken such measures.

If legality truly derives itself from codified rules that are binding and supreme, then in the context of international organizations,
who will enforce these laws? Some argue that the subjectivity in enforcing international law is driven by the dominance of Western powers who seek to enforce their own interests in the world order. In practice, the states that have more power and influence over the development and enforcement of international law control its effectiveness. For example, it is often Western states that have the resources and capability to enforce international law more effectively through the use of military force or economic pressure. This can lead to unequal outcomes in the enforcement of international law, where Western interests are protected and advanced while the interests and rights of non-Western states are marginalized or ignored. In the case of Mohammad Bin Salman, no major state was willing to risk losing its diplomatic ties with Saudi Arabia due to the subsequent economic and political consequences, thereby proving how the actions of the West (or lack thereof) shape the outcome of global events. And MBS, too, used diplomatic immunity as leverage to escape the consequences of his involvement.26

Then, in that case, is the best way to ensure the effective working of the international system by appealing to the humanity of states and state actors? This has worked in several cases such as humanitarian interventions in conflict-prone regions like Kosovo or Sierra Leone, or arguably even through treaties on nuclear non-proliferation, which appeals to collective humanitarian interests. The pathos of public pressure on their respective governments has mostly delivered swift and satisfying responses from the concerned actors. However, are states morally obligated to follow international law? Eric Posner, famed international legal scholar and American lawyer, argues that violating a law is a morally neutral action.27 This essentially means that the legality of a particular action is distinct from its moral imperatives and that the mere act of breaking a law does not automatically carry a moral judgment in the eyes of a court of law. In the case of Germany v. Italy, the ICJ’s ruling favoring Germany was based on the laws of jurisdictional immunity, not retroactive justice based on moral considerations against the Nazi regime. Similarly, in the case of Khashoggi, while some may argue that moral and ethical considerations should motivate the ICJ to punish MBS for Khashoggi’s murder, the fact that MBS is the head of a sovereign state meant that he could legally exercise the doctrine of sovereign immunity, exempting him from trial or punishment by foreign courts. In both cases, there are different moral considerations for what “justice” looks like, but the ICJ is obligated to act in accordance with objective international laws—not subjective moral obligations.

Additionally, since an overarching commitment to international law by a state’s authorities does not guarantee that it is endorsed by every person within that state, there is also no effective way for the international community to enforce an obligation upon a state and its citizens—even one of following international law. Ultimately, it boils down yet again to the necessity of international laws in the first place. International laws are crucial even if they are not always fully enforceable. While enforcement mechanisms for international laws can be limited, their existence can still influence state behavior and serve as a moral and political framework for resolving conflicts and promoting cooperation between nations28—as in the case of Germany v. Italy.

International laws provide a shared understanding of acceptable behavior among nations, promote stability in international relations by providing a framework to
address global challenges, and act as a moral compass by holding states accountable for their actions. However, the existence of international laws enables us to question the sequence of events that took place, and using these laws as our basis, we can criticize the non-action of states.

International law often acts as a mechanism for influencing state behavior, fostering accountability, and shaping global norms even when legal avenues are challenging to pursue. It also establishes a normative framework that outlines acceptable conduct among states and thereby serves as an important tool for soft power projection among states, where compliance with international law is normalized as a more favorable action. Therefore, despite the challenges in enforcing international laws, they remain an important tool for promoting peace, stability, and justice in the international system.

Policy Recommendations and a Way Forward

The case studies above highlight the fact that the enforceability of international laws is highly subjective and is contingent on the power, morality, and vested interests of states. Therein lies, yet again, the conundrum of the international system—Who will enforce international laws? Who will balance the line between state sovereignty and the United Nations’ Responsibility to Protect? It is very clear that despite primary intentions of neutrality, it is impossible to separate decisions at the UN from power politics. Additionally, an interesting factor is that, unlike common law in most countries, international law does not follow the doctrine of binding precedent. The doctrine, in essence, establishes a historical precedent or example that lower courts must follow when deciding cases with comparable facts or similar legal issues. Since there is no court above the ICJ, the doctrine of binding precedent does not apply. If the ICJ takes on legal violations on a case-by-case basis, there is no guarantee that subjectivity can be totally eliminated from the hearings.

Diplomatic immunity came about as a measure of goodwill. However, diplomatic immunity—as the statute itself supports—should not become immunity against human rights violations. In the case of Germany v. Italy, the ruling of the ICJ was well justified because Italy’s actions were a breach of the sovereignty of the German state. However, in the case of Jamal Khashoggi, this law protected an alleged murderer. By deeming the ruler and the state as the same entity, it permitted atrocious acts to be carried out in the name of the ruler. Neither the perpetrators nor the ones who reportedly ordered and authorized the killing were ever punished. Contemporary events prove that even when legal standards are clear, the available mechanisms for enforcing compliance with international law are fragmentary and unreliable.

Does this negate the necessity of international law? Not necessarily. International law exists in place as a mechanism of protection, and the law in itself is inherently neutral. It is neither good nor bad; however, how these laws are exercised and enforced by both international organizations and states is what deems the effectiveness of the law. Ultimately, it is better to have these laws because they provide at least a framework for justice to keep anarchy in check. Of course, no state lives under any illusion that international law alone will facilitate peaceful coexistence. International law is ultimately simply a means for states to either further their own interests or form a basis for their protection—which, of course, is in no way a guarantee.

Additionally, while the ICJ plays a crucial role in upholding and enforcing international
law, it is severely limited by its lack of enforcement mechanisms, jurisdiction, and the principle of state sovereignty. These limitations make it difficult for it to effectively enforce international law and provide remedies for those who have suffered harm. However, neither states nor state agents must be infallible before the law. Ultimately, it falls upon the responsibility of the state to separate itself from the defendant(s)—especially in cases such as Mohammad Bin Salman’s. Effective and unbiased compliance measures must be enforced and encouraged. If the primary enforcer of international law should be the international community, some steps can be taken by individual actors and international organizations to monitor and ensure enforcement. The starting point for this would be the codification of diplomatic immunity as a conditional law.

*Diplomatic Immunity as a Conditional Law*

The international community could promote the conditional application of diplomatic immunity instead of accepting it as an absolute law, wherein the statute can be revoked in cases of human rights violations or criminal activities. The idea would be to introduce flexibility into the application of diplomatic immunity, allowing for its revocation in specific circumstances, such as instances of human rights violations or criminal activities. This approach recognizes the importance of balancing the privileges afforded to diplomats with the imperative of holding individuals accountable for severe breaches of international law. There have been several instances of international law or legal principles that incorporate flexibility based on certain conditions or circumstances, including laws on violations of state sovereignty in the name of humanitarian interventions. The conditional nature of diplomatic immunity would thus demonstrate a commitment to upholding human rights and international legal standards, highlighting the responsiveness of the international legal system to the complexities of evolving contemporary challenges.

To enforce the conditionality of the law and assess the relevant contexts where diplomatic immunity can be invoked, there is a need for the establishment of an independent body not affiliated with the diplomatic mission in question to investigate potential allegations of misuse. This body would amend the United Nations Convention on Jurisdictional Immunities to include mechanisms to review and revoke diplomatic immunity in cases where there is credible evidence to support the allegations of international law violations.

*Establishing Measures of Checks and Balances*

There is always the danger of hegemonic power politics interfering with potential investigations and judgments. In Khashoggi’s case, several powerful countries—including the United States, whose intelligence reports were among the first to link the authorization of the killing to Crown Prince Mohammed Bin Salman—were unwilling to take the case forward or push for deeper investigative efforts. A way around this would be for international organizations like the United Nations to assume an autonomous role in investigations and establish a framework for multilateral consensus where all member countries get a say in pursuing further investigations of alleged crimes and may choose to participate in them. To ensure unbiased and fair efforts, the established body must be comprised of officials from diverse countries—both with and without vested interests in the state/sovereign under investigation. The investigations must also be conducted with utmost transparency and by those with legal expertise in the field of international laws.
and crimes. Transparency would not only ensure public accountability and a more careful assessment of evidence but could also serve as an educational endeavor to highlight the utility of conditional immunity to balance the need for diplomatic privileges with the necessity of accountability. The promotion of legal pluralism would also be an effective way to circumvent Western dominance in international laws, by acknowledging diverse legal traditions and approaches across cultures, which would contribute to a more inclusive and culturally sensitive international legal framework.

Finally, an effective way forward would be to have a measure to periodically review and assess the utility and effectiveness not only of the provisions of conditional diplomatic immunity but also of the entire body of international law and the courts that seek to enforce them. To this effect, states must constitute an independent committee that would conduct in-depth data collection and analysis, consult the public and stakeholders (including legal experts, academics, diplomats, and human rights organizations), and review emerging legal frameworks at regular and defined intervals. International organizations must have legislative frameworks that allow for amendments based on the findings and maintain a continuous dialogue with the public—providing regular updates on new developments, seeking their input, giving notice of any modifications to the laws, and delineating the rationale behind them.

Regular evaluations and communication with the public will allow for evolving international norms based on adjustments from experience as well as emerging geopolitical and security challenges.

**Conclusion**

Minimizing the scope for the misuse of international law begins with a review of existing frameworks and whether they serve the purpose they were intended for. The case of Germany v. Italy indicated the importance of ensuring that international law is not selectively applied or manipulated for political purposes. The ICJ’s ruling, in this case, highlighted the necessity for unambiguous legal frameworks that protect states from unwarranted legal actions while fostering a balance that respects the rights of victims of war and human rights violations. On the other hand, Khashoggi’s case showcased the flaws in the current framework, where the protection of sovereign interests can potentially be exploited to evade accountability for egregious violations of international law.

Ultimately, the pursuit of a more just and effective international legal system requires a collective commitment to upholding humanitarian principles and ethical codes of conduct, transcending the boundaries of realpolitik for the betterment of global peace, stability, and justice. Regular legislative amendments are necessary to ensure the international community is not bound together by rules that no longer reflect the realities of the contemporary socio-political environment. Finally, more creative confidence-building measures and diplomatic efforts need to be devised, where collective humanitarian well-being and moral and ethical codes of conduct take precedence.
About the Author: Apurva Ramakrishnan is a Master’s student in the Georgetown Security Studies Program from India. Her interest in diplomacy and security can be attributed to her experiences living in the Middle East for over 15 years and her tenure in the Indian Ministry of Defense’s National Cadet Corps. She is keenly interested in counterterrorism and counter-radicalization in Asia and is currently pursuing the Terrorism and Substate Violence concentration at SSP. Apurva is also a columnist at the Georgetown Security Studies Review, works as a graduate T.A., and is a Campus Ambassador for the Security Studies Program.


8 “Equipo Nizkor - The Court Finds That Italy Has Violated Germany’s Immunity by Allowing Civil Claims against It for Violations of International Humanitarian Law Committed by the German Reich in Italy.,” n.d. [https://www.derechos.org/nizkor/lm/picj1.html](https://www.derechos.org/nizkor/lm/picj1.html).

9 “Equipo Nizkor - The Court Finds That Italy Has Violated Germany’s Immunity by Allowing Civil Claims against It for Violations of International Humanitarian Law Committed by the German Reich in Italy.,” n.d. [https://www.derechos.org/nizkor/lm/picj1.html](https://www.derechos.org/nizkor/lm/picj1.html).


11 International Crimes Database, “Jurisdictional Immunities of the State (Germany v. Italy: Greece Intervening),” *International Crimes Database*, 2012. [https://internationalcrimesdatabase.org/Case/1231/Germany-v-Italy/](https://internationalcrimesdatabase.org/Case/1231/Germany-v-Italy/).


Russian Influence and Disinformation Operations in the Balkans

Vibor Cipan and David Kirichenko

The Balkans, a region renowned for its intricate history and diverse ethnic tapestry, has been a continual theater of political strife for centuries. Its relative geographical proximity to Russia, especially to countries like Serbia, has been a critical factor in shaping the region's political landscape. This complexity makes it an ideal target for Russian disinformation campaigns, which seek to destabilize the political landscape and further Russian interests. These operations aim to sow discord, exacerbate regional tensions, and undermine European unity, particularly in broader European support for Ukraine. This strategic manipulation of information is a critical tool in Russia's broader objective to challenge European stability and assert Russian influence. This paper outlines key enablers of Russian propaganda and influence manifested through soft power narratives, political affiliations, and disinformation campaigns. Additionally, the paper offers four key dimensions through which Russian influence and impact are evident: the shifts in public perception or opinion; the strategic shaping of election campaigns by media influence; the direct policy implications that follow; and the impact of regional stability and Euro-Atlantic integration. Lastly, we offer recommendations for regulatory frameworks and media literacy programs tightly integrated with the broader societal and policy implications.

Understanding the Historical and Political Landscape and Context

A region of stunning natural landscapes and rich historical narratives, the Balkans is often considered Europe's crossroads. Its intricate religious and ethnic composition mirrors a complex past, with Bosnia and Herzegovina, Serbia, Montenegro, North Macedonia, and Kosovo at its core. Although the Balkans lack definitive political or geographical boundaries, these countries are integral to its identity. Croatia and Slovenia, bordering the Balkans on the west, reflect a Western orientation dating back to the 7th and 8th centuries, demonstrated by their memberships in the European Union (EU), North Atlantic Treaty Organization (NATO), Eurozone, and Schengen Area. Conversely, Serbia and Montenegro are influenced by a history of Eastern traditions, as seen in their Orthodox Christianity and Cyrillic script.

The region's history was further shaped by the Ottoman Empire's expansion in the Middle Ages, nearly reaching Zagreb and Vienna at its peak. Areas of Bosnia and Herzegovina and Serbia experienced Ottoman rule for centuries, notably altering the religious landscape of Bosnia and Herzegovina from predominantly Christian to religious diversity, including a significant Muslim population. Post-Ottoman decline, Austria-Hungary's expansion led to conflicts with independent Serbia (supported by Russia), culminating in the Balkan Wars and World War I, making the region a focal point for clashes between Western and Eastern interests, especially involving Russia.

After World War II, Yugoslavia formed as a federation of six republics (Slovenia, Croatia, Bosnia and Herzegovina, Serbia, Montenegro, and Macedonia) and two autonomous regions within Serbia (Vojvodina and Kosovo). Led by Josip Broz Tito, Yugoslavia established a unique political identity balancing between the Western and Eastern Cold War blocs, notably as a founding member of the Non-Aligned Movement. Despite being a
socialist state, Tito's regime was considered somewhat more liberal compared to its contemporaries, and under it, Yugoslavia did manage to develop its industry and improve the general well-being of its people.

Nevertheless, Yugoslavia could not suppress long-standing ethnic divisions, which were often met with force, killings, and persecution. The dominant role of the Serbs in this context generated resistance from other republics, particularly Croatia and Slovenia. This set the stage for the devastating Yugoslav wars of the 1990s. The Yugoslav People's Army (JNA), primarily commanded by Serbians, initially aimed to quash Slovenia's independence efforts, leading to a brief conflict that presaged longer, more destructive wars in Croatia and Bosnia and Herzegovina. In Croatia, the Homeland War ravaged cities like Vukovar and Dubrovnik, causing extensive casualties and leading to the devastating Bosnian War, characterized by ethnic cleansing and international intervention. The Bosnian War, with its substantial death and injury toll, was the most lethal post-WWII conflict in Europe until the recent Russian aggression against Ukraine.

Post-war, Bosnia and Herzegovina were divided into two main regions: Republika Srpska, based on Greater Serbian territorial aims, notorious for the Srebrenica genocide against Bosniaks, and the Federation of Bosnia and Herzegovina, primarily Bosniak and Croat. Additionally, the District of Brčko was established, representing Bosniaks, Croats, and Serbs more equitably. In the late 1990s, the focus shifted to Kosovo, where independence efforts were met with harsh Serbian retaliation, leading to international intervention against Serbian forces. Despite Kosovo's declaration of independence, Serbia and several other nations do not recognize Kosovo’s claims.

Currently, the Balkan region remains politically fragmented. While most countries are NATO members, and some belong to (or are candidates for) the EU, the diverse regional alignment fosters various international influences. Serbia, for instance, though officially neutral, heavily leans towards Russia and has become central to Russian disinformation campaigns in the area.

**Geopolitical Significance**

The geopolitical stage of the Balkans has long been a focal point of contention between the power centers of the West and East - most notably, Russia. The allure of this region for Russia is multi-faceted, extending beyond cultural affinity to strategic imperatives. As the EU and NATO welcomed more members, the Balkans became a frontier of influence, where East and West vie for hearts, minds, and political allegiance. Russia, acutely aware of the geopolitical significance of the Balkans, has endeavored to cultivate allies and assert influence in this region. The stakes are high, and the implications are profound. The geopolitical tussle in the Balkans is not merely a regional affair but a microcosm of larger global dynamics. The unfolding story reflects broader tensions between Russia and the West, with the Balkans as both a mirror and a stage.

**Current Political Climate**

**EU Integration Efforts and Alignment**

Montenegro and Serbia

In the context of EU integration, both Serbia and Montenegro are notably more advanced than their Western Balkan peers. This progress is not merely a regional matter but also carries significant geopolitical
implications, particularly concerning Russia's regional influence. EU integration offers a pathway toward stronger ties with Western Europe, promising economic, political, and security benefits. For Serbia and Montenegro, it means greater political stability, economic growth, and strengthening democratic institutions. However, this shift towards the West is seen as a move away from Russia's traditional influence in the Balkans. Historically and culturally tied to the region, Russia perceives EU integration as a threat to its strategic interests—especially with Serbia.

Domestically, the political narratives in Serbia and Montenegro fluctuate between pursuing EU membership and maintaining historical connections with Russia. This reflects the geopolitical tension between Western and Russian influences in the region. Moreover, aligning with the EU necessitates comprehensive reforms and adherence to its standards. This can clash with the entrenched political interests in Serbia and Montenegro, complicating their path to EU membership.

Finally, the success of EU integration depends on several factors: Serbia and Montenegro's political determination to meet EU criteria, the changing global political dynamics, and the general EU sentiment towards further enlargement.

Bosnia and Herzegovina

Bosnia and Herzegovina, especially its smaller entity, Republika Srpska, is highly susceptible to Russian disinformation campaigns. Milorad Dodik, the leader of Republika Srpska, openly aligns with Russia, exacerbating this vulnerability. The close ties among Serbia, Republika Srpska, and Montenegro enable the seamless spread of media narratives, including Russian disinformation, impacting the political climate in the region. These dynamics pose significant hurdles for Bosnia and Herzegovina's aspirations to join the EU and NATO, which faces opposition from Serbia and Republika Srpska, both closely linked to Russia. Considering the expansion of EU and NATO influence as a threat to its regional dominance, Russia actively opposes these integration efforts through disinformation and political alliances. In contrast, Croatia, a member of both the EU and NATO, champions Bosnia and Herzegovina's integration into these Euro-Atlantic structures. Croatia's advocacy stems from its own experience with EU accession and the belief in the stabilizing effect of European integration. As a neighbor with a significant Croat population in Bosnia, Croatia is invested in the region's stability. It leverages its EU position to support Bosnia's aspirations, underscoring the benefits of EU membership, such as improved political and economic stability and adherence to the rule of law. However, this pursuit is complicated by evolving EU standards, geopolitical shifts, and a general weariness towards EU enlargement.

The Impact of The Russo-Ukrainian War

Russia’s invasion of Ukraine in February 2022 had a profound impact on the Western Balkans, affecting Serbia, Montenegro, and Bosnia in significant ways. The invasion triggered political and social reactions that reflect the complex interplay of influences and loyalties. Russia was quick to play that card.

In Serbia and Republika Srpska, the invasion strengthened already strong pro-Russian sentiments. These areas, having deep-rooted Orthodox Christian ties with Russia, have witnessed public demonstrations supporting Russia's actions. The pro-Russian rallies signify not just a shared religious identity but also a political stance against Western influence in regional affairs. This alignment has historical
precedence, with the Serbs of both Serbia and Republika Srpska being Moscow's longstanding partners in the Balkans.12

Contrastingly, in other Balkan states, the war's brutality has revived unsettling memories of the Balkans' tumultuous past, particularly the conflicts of the 1990s. There is an underlying fear that the aggression seen in Ukraine could spark similar unrest closer to home, given their relatively recent war experiences with Serbia.13 As mentioned earlier—those nations fret that if Russia is victorious, Serbian nationalists may be encouraged to avenge their losses in the wars over Bosnia and Kosovo and challenge their outcomes.

The Russo-Ukrainian war further prompted a demographic shift. Russians, seeking to escape the repercussions of their country's actions, moved into Serbia and some to Montenegro. Over 200,000 Russians have emigrated to Serbia since Russia’s full-scale invasion of Ukraine.14 These nations, offering visa-free entry to Russians, have become havens for those fleeing the crisis. This shift could reshape these Balkan states' political and social fabric, potentially tipping the scales in favor of Russian geopolitical interests.15 In essence, the more Russians come to the region, the greater the political ties will become with Russia. It is now a common sight in places like Belgrade to see people wearing pro-war and pro-Russian symbols.16

In reaction to these developments, some Balkan governments cautiously started distancing themselves from Moscow, wary of being drawn into the orbit of a nation now widely regarded as a pariah by the West. Furthermore, the European Union, cognizant of the delicate balance of power in the region, has reinforced its peacekeeping forces in Bosnia as a precaution against the destabilizing effects of the invasion.17

Serbia’s Political Maneuvering and the EU’s Appeasement

In the context of the Ukrainian crisis, Serbia's strategic positioning reflects a precarious balance between Eastern and Western alliances. President Vučić, while committing to the European Union path, notably avoids sanctioning Russia, a move indicative of Serbia's effort to maintain favorable ties with both sides.18 Vučić's administration is marked by diplomatic ambiguity. Despite the Russo-Ukrainian war, Serbia remains on its equivocal path, juggling Western expectations and its historical ties with Russia.19

This is exemplified by Serbia's acquisition of an "extremely favorable" gas deal with Russia during the global condemnation of Moscow's actions in Ukraine.20 The gas deal is one of several instances of ongoing Serbian-Russian engagements. For example, in August 2022, Serbian minister Aleksandar Vulin met Russian Foreign Minister Sergei Lavrov in Moscow, receiving honors from the Russian Ministry of Defense, underscoring the nations' close relationship. The European Union, recognizing Serbia's strategic role in regional dynamics, adopts a somewhat conciliatory, if not appeasing, approach.21

Considering Serbia's interactions with Russia, this could be viewed as an attempt to maintain influence in the Balkans.22 Serbia's geopolitical clout allows Vučić to extract benefits from both Russian and EU partners, exploiting the EU's desire to sustain influence in the region against Russian assertiveness. For instance, as mentioned earlier, after Russia invaded Ukraine, the EU significantly increased its peacekeeping force in Bosnia. This move, aimed at reinforcing security, also signals to Serbia and Republika Srpska the EU's protective, rather than punitive, approach -
potentially precarious message in this volatile region.

**How Pro-Russian Disinformation Works in the Region**

**Reasons for Disinformation Susceptibility**

Understanding the previous (geo)political and historical dynamics is essential for understanding the Western Balkans' geopolitical landscape, particularly considering the overarching influence of Russian interests in the region and the strategic importance of these countries to the EU and NATO's efforts to counteract Russian influence.

These developments reflect the ongoing tug-of-war for influence in the Western Balkans, where Western institutions seek to promote democratic values and integration. At the same time, Russia asserts its influence through local allies like Serbia’s Vučić, Bosnia’s Dodik, and Montenegro’s Mandić, who embody a different, scarier vision for the region's future.

The close social, historical, and cultural ties between Serbia, Republika Srpska, and Montenegro facilitate the daily flow of media narratives, including disinformation, across borders. This interconnectedness provides a conducive environment for the propagation of Russian disinformation, significantly impacting the political landscape.

**Key Enablers of Russian Propaganda and Influence**

In March 2019, the Russian Academy of Military Sciences held its annual defense conference with Valery Gerasimov, Chief of the General Staff. During that occasion, he presented what is coined the Gerasimov doctrine (a somewhat ill-named term). His approach emphasized non-military means to achieve political and strategic goals.

Blending traditional military tactics with information warfare, it aims to destabilize societies from within, leveraging political divisions and societal vulnerabilities. His speech made several key references to the increasing role of the information dimension. Even earlier, Gerasimov claimed that the new reality of wars would show the significant importance of informational dimensions - even going so far as to say that future wars will include the transfer of hostilities precisely into this area, pointing out that information technology is essentially becoming one of the most promising types of weapons.

With this in mind, it is possible to observe Putin’s chaos efforts as an extension, or even an implementation, of parts of this “doctrine” in the Balkans. Under Putin, Russia leverages this chaos strategy of creating and exploiting disorder as a form of geopolitical maneuvering in the Balkans, characterized by unpredictability and opportunism.

The Kremlin’s strategy in the region involves inflaming ethnic tensions and encouraging protests through disinformation and political affiliations, aligning with General Gerasimov’s doctrine of hybrid warfare. This approach aims to destabilize the region, distract the West, and maintain Russian influence. The Western Balkans’ susceptibility to Russian influence, stemming from historical, cultural, and political ties, makes it a fertile ground for implementing Putin's chaos theory.

However, the response from international actors like NATO and the EU, as seen in their reaction to the events in Kosovo, indicates a recognition of the need to counter this destabilizing influence.

Russia’s propaganda and disinformation network in this region generally relies on soft power narratives, political affiliations, and disinformation campaigns—perpetrated...
against Serbia, Montenegro, and Republika Srpska (in Bosnia and Herzegovina). The broader geopolitical impact of Putin's chaos theory in the Balkans lies in its ability to generate uncertainty and instability, not just within the region but also in its relations with the West. This strategy serves multiple purposes for the Kremlin: it keeps the Western Balkans in a state of flux, preventing full integration with Western institutions like the EU and NATO, and it diverts attention and resources from other areas of strategic importance to Russia, such as Ukraine. 

To counteract the disruptive influence of Putin's chaos theory, a multifaceted approach is required. But before that is even possible, it is necessary to understand each of the elements—soft power narratives, political affiliations, and disinformation campaigns—and how they are implemented in this region.

Soft Power Narratives

Russia's primary soft power narratives in the region pivot around anti-Western opposition and the espousal of civilizational Orthodox brotherhood. Moscow employs the discourse of the Russian Orthodox Church against perceived Western "moral decadence" to vilify the process of European integration. This narrative resonates with local societies in Bosnia’s Republika Srpska, Serbia, and Montenegro, which find an ideological affinity with Russia's positions due to their Orthodox Christian heritage. Those positions leverage historical, cultural, and religious ties to foster pro-Russian sentiment in the region. Key narratives include the shared Orthodox Christian faith, pan-Slavism, and opposition to Western ideologies.

Russia also plays the Kosovo card (using the Kosovo issue to influence and strengthen its ties with Serbia as a tactic of leveraging a sensitive and significant regional issue) when appealing to the national pride of Serbia and Serbs in the region. It does so by positioning itself as the “protector of the Serbs,” and Serbia’s Vučić often turns to Russia for support in these matters - a fact that Russia knows how to use. Russia leverages the Kosovo issue to strengthen its ties with Serbia, often framing itself as the defender of Serbian interests. This strategy resonates strongly in Serbia, where Vučić frequently seeks Russian backing on matters related to Kosovo. Russia's involvement is strategic, capitalizing on Serbia's position and the historic tensions in the region to reinforce its influence.

Political Affiliations

Another area where Russia really can count on strong support and fertile ground is the current state of political affiliations in the region. In Republika Srpska, the political leadership, notably Milorad Dodik, who has close ties with Russian President Vladimir Putin and Serbia's pro-Moscow leader Vučić, openly aims to align Bosnian Serb territories closer to Russia and Serbia. Dodik's actions, which have earned him sanctions from the U.S. and the U.K., reflect a broader inclination towards Russian influence in the region. Furthermore, Dodik awarded the Russian president a medal for his 'patriotic concern' for Republika Srpska, underscoring the deep-seated political connections with Moscow.

Serbia's political landscape is notably marked by the enduring allegiance towards Russia, spearheaded by the nation's President Vučić. Amidst the broader geopolitical turbulence, particularly the ongoing conflict in Ukraine, Vučić has meticulously consolidated power, with his pro-Russian stance seemingly solidified by the unfolding events in Eastern Europe. This alignment is not merely rhetorical but manifests in substantial political events. A
significant milestone was the foreign policy agreement signed between Serbia and Russia, a gesture underscoring the synchronized stance of the two nations on certain international matters.\textsuperscript{33}

The depth of this political camaraderie between Serbia and Russia is further elucidated by the personnel helming critical establishments within Serbia. A case in point is the pro-Russian head of Serbia's state intelligence service, Aleksandar Vulin, who resigned recently. His allegiance to Moscow didn’t come without repercussions; it led to him facing sanctions from the United States. This episode doesn’t just underline the individual's pro-Russian tilt but also reveals the extent to which Russian influence has permeated the echelons of Serbia’s political and security apparatus.\textsuperscript{34}

Lastly, in Montenegro, the political arena is a microcosm of the larger East-West dichotomy, with Russia's influence casting a long shadow despite the country's strides towards EU integration. The formation of coalitions between pro-Russian and pro-EU factions is a telling illustration of this dynamic, reflecting the tug-of-war of influences that characterizes Montenegro's political landscape.\textsuperscript{35} These unlikely coalitions are not mere political expedience but underscore the extent to which Russian influence persists, even as Montenegro navigates the path towards European integration.\textsuperscript{36}

The resonance of pro-Russian and anti-Western sentiments isn't confined to the corridors of power but extends to the broader political spectrum, influencing pivotal decisions. Recent government formation efforts in Montenegro offer a glimpse into this influence. Pro-Russian and anti-Western groups wield significant sway, steering political decisions in a manner conducive to Russian interests.

Case Studies: Disinformation Campaigns in Serbia, Montenegro, and Bosnia and Herzegovina.

The success of Russian disinformation campaigns in Serbia, Republika Srpska, and Montenegro can be attributed to pre-existing soft power narratives and regional political affiliations. These narratives and affiliations, as explained earlier, provide fertile ground for Russian influence operations, which attempt to undermine entities like the European Union and NATO. The media landscape in these regions is also a significant factor, and its understanding is critical.\textsuperscript{37}

Serbia

In Serbia, Russia established a regional base for the Sputnik news agency in Belgrade. Sputnik Serbia disseminates narratives suggesting that the EU and NATO are weak, fragmented, and aggressive, thereby challenging further integrations of the Western Balkans into these Western structures.\textsuperscript{38} The disinformation techniques adapt to local sensitivities; in Serbia, Sputnik is presented as a Serbian media outlet rather than a conduit for Russian state propaganda, making its messages more palatable to the local audience. This approach leverages the historical, cultural, and geopolitical ties between Russia and Serbia, amplifying narratives that resonate with Serbian society, such as shared Slavic brotherhood and Orthodox Christian values.\textsuperscript{39}

One of the most significant Russian media presences in the region is the state-owned news agency Russia Today (RT), which, despite EU sanctions, launched an online platform named RT Balkan in Serbian.\textsuperscript{40} The EU banned RT and Sputnik due to their roles in spreading disinformation and justifying Russia's war in Ukraine.\textsuperscript{41} However, RT Balkan has managed to skirt...
these sanctions, setting up operations that provide content across various social media platforms and planning to initiate television broadcasting by 2024.

The narratives disseminated through these outlets often reflect the official Russian stance on international affairs. For example, Serbian media, heavily influenced by Russian propaganda, initially presented the invasion of Ukraine as a defensive action against a Western threat to Russia. This narrative taps into historical grievances, such as the NATO bombing of Serbia in 1999, to find resonance among the Serbian population. In this sense, the disinformation campaigns are both retrospective, invoking past conflicts, and prospective, shaping perceptions of current events. Other common topics used in these efforts are claims that EU and Western sanctions are “shot into the foot” of the EU and the West and claims that Serbia can’t impose sanctions since “China and Russia are our only friends.”

The implications of these narratives are significant. They engender a skewed perception where the EU and the West are depicted as adversaries, whereas Russia and the People’s Republic of China (PRC) are portrayed as allies. Such distortions are not merely prevalent in fringe media but are evident in public opinion polls where Russia and the PRC are perceived as larger contributors to Serbia than actual major donors like the EU, Germany, and the USA.

Montenegro

In Montenegro, Russian influence tactics have included disinformation campaigns designed to stoke political and ethnic divisions, to destabilize the country and hinder its integration with the West. Notably, Russian narratives have described Montenegro's bid for NATO membership as a provocation, and Russian agents have been implicated in an attempted coup d’état to instigate political unrest. AidData's profile on media ownership and potential foreign influence channels in Serbia and Montenegro highlights the complexities of the media market, identifying ownership as a potential entry point for Russian influence through elites and media owners. The strategic use of local media to republish content favorable to Russian interests is a tactic observed across the Balkans.

Bosnia and Herzegovina (Republika Srpska)

In Republika Srpska, Russian influence operates mostly under Serbian directives, relying on the same or very similar tactics and narratives. Rather than direct media ownership, Russia supports the long-serving president of RS, Milorad Dodik. As the long-serving president of Republika Srpska, Milorad Dodik plays a central role in Russia's influence strategy. Dodik's political alignment with Russia is well-documented. He often advocates for closer ties with Moscow and has opposed Western influence in the region. His control over local media is a critical aspect of maintaining this influence. Politically motivated legal attacks against critical outlets and selective funding for loyal ones are methods used to maintain this influence - something ongoing since at least 2012.

The most influential media outlet in RS, RTRS, has close ties with the RS government and Russia. Meanwhile, Glas Srpske, owned by Nezavisne Novine and allied with Dodik, does not show direct ties with Russia but is perceived as being in the pro-Russian camp.

Another interesting aspect of Russian involvement in Republika Srpska is through the Night Wolves, also known as "Nochnye Volki," a Russian motorcycle club with a significant presence in the political and
cultural landscape of Russia and Eastern Europe. Founded in the late 1980s, the club evolved from a group of motorcycle enthusiasts into an organization with staunch nationalist views and close ties to Russian President Vladimir Putin. Night Wolves have been blacklisted by the US government in 2014 due to their involvement in activities endangering the peace, security, stability, sovereignty, and territorial integrity of Ukraine during the Russian occupation of Crimea. Dodik's connections with the Russian motorcycle club "Night Wolves" and their activities in Republika Srpska (RS) in Bosnia and Herzegovina showcase a symbolic and practical alignment with Russian nationalist and separatist ideologies.

Dodik has used the occasion of RS Day, declared unconstitutional by the Constitutional Court of Bosnia and Herzegovina, to affirm his allegiance to Russia. The military-style parade near Sarajevo, part of the festivities, included thousands of ethnic Serbs. The parade featured around 2,000 Bosnian Serb law enforcement officers and the Night Wolves, who openly support Russian President Vladimir Putin. The group's presence and military display might indicate Dodik's strategy to align RS more closely with Russian nationalist and separatist sentiments.

During these events, an announcer praised the mission of the Night Wolves to "promote Orthodox Christianity and make the Republika Srpska as mighty and eternal as Mother Russia." The group, funded by the Kremlin, has been involved in pro-Russian activities in Ukraine's Crimea and Donbas regions since 2014. Dodik also awarded Putin with his administration's highest medal of honor, citing his "patriotic concern and love" for RS.

Lastly, the Bosnian Security Minister Selmo Cikotić confirmed that intelligence information indicated the presence of the Night Wolves, alongside other extremist groups such as the Russian mercenary group Wagner and the Ravnogorski Chetnik Movement (Ravnogorski četnički pokret), in RS celebrations. The celebration of RS Day has seen the attendance of members from radical and extremist groups, including Chetnik associations in Bosnia and Herzegovina, alongside the Night Wolves. Flags of pro-Russian separatist entities like the so-called "People's Republic of Donetsk" were displayed, indicating a broader nexus of nationalist and separatist ideologies converging in RS.

Russian Disinformation Targeting Support for Ukraine in the Balkans

Since Russia invaded Ukraine, the Russian Embassy in Bosnia has actively used social media to disseminate information on the conflict. These posts often refute claims against Russian soldiers, defend Russia's military involvement in Ukraine, and draw parallels between the situation in Eastern Ukraine and the 1995 Srebrenica genocidal massacre, where more than 8000 Bosniaks were tragically killed. These comparisons are not new; Russian authorities have previously likened the Donbas conflict to the Srebrenica massacre. In 2017, President Vladimir Putin warned that the actions of Ukrainian nationalist battalions in Eastern Ukraine could be more severe than the Srebrenica massacre. Russia has consistently denied the Srebrenica genocide, leveraging it to rationalize its 2014 incursions in Donbas. Karcic, an expert on the subject, notes that Russia uses this narrative to justify its actions in the region.

This strategy is part of Russia's broader propaganda efforts aimed at causing destabilization by stirring tensions and distracting the global community. Ukraine’s chargé d'affaires in Bosnia, Serhii Miniailo, points out Russia's concern over the support
Ukraine is garnering. He highlights the Balkans' strategic significance in European security, suggesting that Russia's actions in the region are part of a larger geopolitical game. James Rubin, the coordinator for the U.S. State Department's Global Engagement Center, emphasized the Kremlin's ongoing efforts to destabilize Western unity, particularly in its support for Ukraine. According to Rubin, the Kremlin employs information warfare as a strategic tool to sow discord among the United States, NATO allies, and European Union partners. He pointed out that these tactics are aimed at weakening the collective Western response to the Russian invasion of Ukraine, showing what their aims are to divide and break the West's unified stance. Rubin earlier commented, "The main source of the threat in this part of the world is Russian-generated disinformation, often repeated and acting as a hub through Serbian media platforms and then repeated and promulgated here in the Western Balkans."

In Serbia, the predominant source of information about the Russo-Ukrainian war for the average citizen is the influential pro-government mainstream media. This media often propagates disinformation to support Russian propaganda, which portrays Ukraine as a Nazi country that Russia is attempting to liberate. Serbian users are frequently exposed to viral posts, which spread across language barriers, containing fabricated claims about the presence of Nazi symbols among Ukrainian officials, soldiers, or within Ukraine itself. Additionally, Russian officials have promoted conspiracy narratives suggesting that Ukraine houses dozens of "U.S. laboratories for bioweapons aimed at the destruction of Russians." The Serbian media has quickly adopted and disseminated this narrative, furthering the spread of such misleading information.

Understanding the Landscape

It is crucial to understand that these disinformation campaigns are not isolated incidents but part of a broader strategy by Russia to influence the Balkans. The strategy includes establishing connections with local media outlets and ownership structures, which can serve as conduits for Russian state-owned media content and narratives that align with Russian geopolitical interests. The impact of such campaigns is multi-faceted, affecting public opinion, political decision-making, and the overall media environment in these countries.

This complex web of media influence, historical narratives, and geopolitical affiliations underscores the challenges faced by the Western Balkans in countering Russian disinformation, especially in the volatile context of the Ukraine invasion. The operations of RT and Sputnik in the Balkans, particularly Serbia, despite EU sanctions, exemplify the ongoing struggle for narrative control in the region and raise concerns about Serbia's alignment with EU policies.

In summary, the Russian disinformation campaign in the Western Balkans, propelled by media outlets like RT and Sputnik, capitalizes on historical alliances and cultural affinities. It leverages local media ownership and the absence of robust, independent journalism to disseminate narratives supporting Russian geopolitical aims, including depicting the Ukraine invasion in a light favorable to Russian interests. This situation represents a significant obstacle to the region's aspirations for integration with the West and underscores the strategic importance of independent media in safeguarding democratic values and processes.
Connecting the Dots

Russian disinformation campaigns materialize in a series of tangible outcomes in the Western Balkans. Its impact and effects can be analyzed and summarized through several critical dimensions: the shifts in public perception and opinion; the strategic shaping of election campaigns by media influence; the direct policy implications that follow; and the impact on regional stability and Euroatlantic integration.

Connecting these dimensions clarifies how Russian narratives sway popular sentiment, influence political landscapes, and steer regional policies. All these bear significant weight on the Western Balkans' trajectory toward European integration.

Public Perception and Opinion Shifts

The resonance of Russian disinformation in the Western Balkans is partly due to the skillful exploitation of historical narratives. The impact of Russian disinformation in Serbia is a stark example of how the exploitation of historical narratives can diverge public perception from reality. Despite the EU being the largest financial donor for Serbia, Russian media influence has led many in Serbia to believe that Russia and the PRC are the primary benefactors. This misconception is a direct result of the disinformation campaigns that exploit Serbia's historical and cultural ties with Russia, overshadowing the substantial economic assistance provided by the EU.

The effectiveness of Russian media like Sputnik and RT in Serbia is augmented by their use of the Serbian language, which facilitates the portrayal of these outlets as local rather than foreign, enabling easier dissemination of their narratives. They capitalize on Serbia's historical narratives and resentment from past conflicts to intensify pro-Russian and anti-NATO sentiment. This tactic is particularly potent since Serbian is not only the official language of Serbia and co-official language in Montenegro but is also widely understood in Bosnia, thus broadening the reach of Russian propaganda across the region. The local language usage embeds these outlets within the media landscape, masking their true origins and intentions.

Election Campaigns

In the Western Balkans, the landscape of election campaigns has undergone a notable transformation, largely influenced by the media narratives shaped by Russian-backed outlets. This shift in the media landscape has profound implications for nations like Serbia and Montenegro's political orientation and foreign policy choices, leading to a polarization that aligns with Russian interests.

In both Serbia and Montenegro, media outlets backed by Russia have played a crucial role in shaping public opinion during elections. These outlets often cast pro-Russian politicians in a positive light, amplifying their agendas and ideologies. This media strategy has been particularly effective in Montenegro, where a notable pivot towards pro-Serbian and pro-Russian leadership occurred.

The case of Montenegro is especially significant, considering its historical bid for NATO membership and its traditionally Western-leaning political stance. The shift in Montenegro's political landscape towards pro-Russian sentiments can be traced back to the strategic use of media narratives that portrayed pro-Russian politicians as protectors of traditional Montenegrin and Serbian Orthodox values. This portrayal resonated with a significant portion of the population, who felt a cultural and religious kinship with Serbia and, by extension, Russia. This media strategy has proven
effective, as seen in Montenegro's pivot towards pro-Serbian and pro-Russian leadership.\textsuperscript{61}

Similarly, in Serbia, the media's endorsement of politicians with pro-Russian leanings has had a discernible impact on election results, suggesting a significant correlation between the media's portrayal of candidates and their electoral success.\textsuperscript{62} The Serbian media landscape, influenced by both domestic and Russian-backed outlets, often presents pro-Russian politicians as champions of Serbian national interests, juxtaposing them against Western influences. This portrayal taps into a deep-rooted sentiment among many Serbians who view Russia as a historical ally and protector of Orthodox Christian values.

The effectiveness of this media strategy is evident in the electoral success of politicians who align themselves with Russian policies and interests. These successes suggest a significant correlation between the media's portrayal of candidates and their ability to garner electoral support. The media narratives shape public opinion and play a crucial role in defining the political discourse in Serbia.

Furthermore, the success of Russian media strategies in these elections showed just how effective Russian disinformation can be in shaping political outcomes. It showed the growing importance of media in modern democratic processes and how democratic societies need to be more vigilant against external influences that seek to sway electoral processes and public opinion.

\textit{Policy Implications}

The repercussions of Russian media influence are evident in key policy stances. In Serbia, the alignment with Russian foreign policy, fueled by Russian disinformation efforts, has led to a divergence from the typical EU-aligned foreign policy trajectory. This is evident in Serbia's refusal to join EU sanctions against Russia, which starkly contrasts the position held by most EU aspirants and member states.\textsuperscript{63} The Serbian government's approach, influenced by the strong public sentiment favoring Russia as depicted by local Russian-controlled media, suggests a policy-making process that is sensitive to the narratives shaped by these outlets.

Republika Srpska's stance on NATO is illustrative of Russian influence. The entity's resistance to NATO membership echoes the Kremlin's position, reinforced by the media landscape where Russian-backed outlets are prevalent. This media environment amplifies the anti-NATO sentiment, which aligns with the region's historical skepticism of Western military involvement. The local media's portrayal of NATO, steeped in the memories of past conflicts and the perceived protection of the Slavic brotherhood by Russia, is a decisive factor in shaping Republika Srpska's policies.

These policy stances underscore the broader geopolitical implications of Russian disinformation campaigns. By swaying public opinion and influencing the policy-making process, the campaigns successfully steered Serbia and Republika Srpska towards a more Russia-centric foreign policy, potentially affecting their international relationships and standing with Western institutions.

\textit{Regional Stability and EU Integration}

The pervasive reach of Russian disinformation campaigns has significant implications for regional stability and the trajectory of EU integration for Western Balkan countries. These campaigns, adept at leveraging historical tensions and fostering nationalist sentiments, sow distrust toward Western institutions, particularly the EU and NATO. In Serbia, this has manifested in a
reluctance to align with EU foreign policy. At the same time, Republika Srpska has bolstered opposition to NATO integration, both stances serving to fortify Russian influence in the region. 64

The resultant skepticism not only decelerates the EU enlargement process but also deepens internal divisions, threatening the fabric of multi-ethnic societies within the region. Such divisions are exacerbated by the media’s portrayal of the EU and NATO as adversaries. This narrative finds resonance amidst the populace still grappling with the legacies of past conflicts. Moreover, these internal divisions have geopolitical repercussions, emboldening non-EU actors like Russia and the PRC to further assert their influence, thereby complicating the Western Balkans' political landscape. This shift has broader geopolitical implications, especially in the context of the region's integration with European and transatlantic institutions. The political realignment towards Russia presents challenges for the EU and NATO, as it could potentially undermine their influence and objectives in the Western Balkans.

The challenge for the EU is to navigate this complex milieu, where the push for democratic resilience and unity is continually undermined by strategic disinformation efforts that capitalize on existing vulnerabilities. Addressing this challenge requires a multifaceted strategy that not only counters disinformation but also strengthens local media, bolsters civil society, and reassures citizens of the tangible benefits of EU integration. Such efforts are essential to reorient the Western Balkans towards a stable, prosperous, and EU-aligned future.

**Recommendations**

This final chapter synthesizes the complex web of Russian disinformation's impact on the Western Balkans, drawing on multiple sources to offer a well-researched and analytical perspective. Based on the EUvsDisinfo study, disinformation is identified not as the root cause but as a symptom of deeper societal and governance issues within the Western Balkans. 65 In assessing the current efforts to combat disinformation in the Western Balkans, it makes sense to consider the initiatives spearheaded by the EU, particularly those under the aegis of EUvsDisinfo. These initiatives have been central to identifying, analyzing, and debunking disinformation campaigns proliferating across the region.

The EU’s Code of Practice on Disinformation emerged in 2018 as a pioneering attempt to bring together industry stakeholders to tackle the spread of disinformation online. 66 It has been effective in certain respects, such as during the COVID-19 pandemic, where it played a key role in monitoring and ensuring the accountability of signatories in limiting the spread of related disinformation. However, an assessment of the Code reveals several critical gaps. Among the noted shortcomings were the lack of clear definitions for key concepts and the absence of relevant key performance indicators to measure the effectiveness of platform policies against misinformation. These issues highlight the challenges of applying broad, EU-wide strategies to the Western Balkans' unique and complex media landscapes.

The updated 2022 Strengthened Code of Practice on Disinformation seeks to build upon its predecessor by setting more ambitious commitments and involving a more diverse range of stakeholders. 67 This aims to bring about comprehensive improvements in the fight against disinformation. The revised Code
acknowledges that counter-disinformation efforts must be dynamic and adaptable, particularly in regions like the Western Balkans, where the effectiveness of such measures is contingent upon their alignment within specific sociopolitical contexts.

A study commissioned by the European Parliament’s Committee on Foreign Affairs underscores the importance of understanding disinformation as a symptom of the Western Balkans’ vulnerabilities, such as societal divisions and governance weaknesses. It suggests that a purely reactionary approach to disinformation is insufficient. Instead, there is a need for initiatives that foster societal resilience and strengthen democratic governance as a bulwark against the influence of disinformation.

The study correctly claims that disinformation does not result from poorly regulated digital media ecosystems or exogenous security threats. Rather, it is the result of a combination of structural vulnerabilities. It lists poor governance, geopolitical exposure, deep-seated internal enmities, and disunity as key areas but also presents them as a dynamic field of opportunities to deploy false information to attain political goals.

Finally, it concludes that, as a result, the policy responses recommended here do not reside entirely in the media or security domains but place significant emphasis on the domains of governance, public engagement, and diplomacy. With this in mind, it’s necessary to observe regulatory frameworks and media literacy programs and their impacts on societal and policy dynamics when positioned to fight disinformation successfully.

Regulatory Frameworks

Regulatory frameworks play a pivotal role in the fight against disinformation. The EU’s Code of Practice on Disinformation, initiated in 2018 and updated in 2022, represents a collaborative effort among industry stakeholders to mitigate the spread of disinformation. This self-regulatory framework has had successes, particularly during electoral periods, but its application encounters challenges.

For the Western Balkans, adapting this Code necessitates an in-depth understanding of the local media ecosystems and political dynamics to address specific vulnerabilities. It would involve not only translating the framework into local languages but further customizing it to align with regional legal standards and societal norms.

Key actions might include enhancing transparency in media ownership, improving the detection and reporting of disinformation, and fostering partnerships between government bodies, media organizations, and civil society to enforce these regulations effectively.

To further refine the regulatory frameworks for combating disinformation in the Western Balkans, the EU’s *acquis communautaire* could serve as a foundation, providing a set of standards that countries can inherit as part of their EU integration process.

Adjusting these frameworks to the local context could involve:

- **Regional adaptation of EU legislation** - Tailoring EU disinformation legislation to fit the Western Balkans' specific legal and cultural nuances. This adaptation ensures that the regulations are effectively integrated into the local legal system, considering the specific socio-political dynamics of the Western Balkans. It helps make the laws more effective and enforceable in the local context.

- **Stakeholder engagement and co-creation workshops** - Bringing
together policymakers, media professionals, and civil society to discuss and shape the local adaptation of regulatory frameworks. These workshops facilitate a collaborative environment, ensuring that policies are informed by a wide range of perspectives, including those most impacted by disinformation. As such, they promote ownership and commitment from various societal segments.

- **Transparency initiatives** - Implementing policies that increase transparency in media ownership and funding sources, particularly for online platforms. Enhancing transparency in media ownership and funding combats hidden influences and biases is crucial for building public trust in media. This will, in return, help create a more informed and critical audience.

- **Local language resources** - Creating accessible resources in local languages that explain the implications and responsibilities under the new frameworks. Providing resources in local languages ensures wider accessibility and understanding of the policies among the general population, leading to better compliance and participation in the democratic process.

- **Monitoring and evaluation systems** - Establishing systems to monitor the effectiveness of the regulations and adapt them as disinformation tactics evolve. These systems are vital for assessing the effectiveness of the regulations and enable continuous improvement, ensuring that the policies remain relevant and effective against evolving disinformation tactics.

- **Public consultation processes** - Engaging the public in the legislative process to ensure that the regulations reflect societal values and priorities, thereby enhancing public buy-in and compliance. Engaging the public directly fosters a sense of ownership and responsibility towards the implemented measures. It also helps identify potential blind spots and challenges in the regulatory framework, ensuring that it is comprehensive and effective. Public involvement in the legislative process is key to building trust and legitimacy, which is vital for successfully implementing policies to combat disinformation.

To be truly effective, any adapted regulatory framework must be accompanied by robust enforcement mechanisms and the capacity to evolve in response to the changing tactics of disinformation campaigns. This could include establishing dedicated oversight bodies with the authority to monitor, investigate, and act on violations of disinformation policies. Additionally, developing a rapid response system within these mechanisms would allow for swift action against disinformation campaigns as they emerge.

To maintain relevance over time, these frameworks should incorporate provisions for regular reviews, leveraging insights from media watchdogs, academia, and civil society. This approach ensures that regulations can adapt to evolving disinformation tactics, maintaining their effectiveness in a constantly changing information environment. This tailored approach should aim to bolster the region’s resilience to disinformation while respecting freedom of expression and media pluralism.

*Media Literacy*
Media literacy is essential in empowering individuals to navigate the complex landscape of modern information. It equips people with the skills to critically analyze the content, discern biases, and recognize misinformation and disinformation. The critical thinking fostered by media literacy is a cornerstone of a functioning democracy, as it enables citizens to make informed decisions.

Finland's success in media literacy education serves as a laudable model. Finnish strategies incorporate media education across all levels, from early childhood through adult education, and are supported by a wide range of stakeholders. These efforts contribute to Finland's robust defense against disinformation, as evidenced by its high ranking in the Media Literacy Index. The Finnish model also involves society-wide initiatives, engaging schools, libraries, government departments, universities, and NGOs in developing and enacting learning programs. Finnish students are equipped to differentiate between disinformation and misinformation, a skill that contributes to Finland's ranking as highly resistant to fake news.

Adapting these programs to the Balkans faces challenges and requires a nuanced strategy that acknowledges local educational structures, media landscapes, and cultural contexts. Educational infrastructures differ, and the political climate is more volatile. Implementing these programs requires accounting for the region's diverse cultural context and varying levels of media freedom. The programs must be localized, potentially through public campaigns, workshops, and the integration of media literacy into school curricula. Collaboration between government and civil organizations is crucial to address the region-specific challenges and enhance public resilience against disinformation.

Key elements of such an approach, however, can be enabled through the following initiatives:

- **Partnership-driven media literacy campaigns** - Collaborations between Balkan educational institutions, NGOs, and media organizations to develop localized media literacy resources.
- **Inclusive curriculum development** - Tailoring school curricula to include media literacy, focusing on the region's history, languages, and current media challenges.
- **Community media workshops and various stakeholder engagement methods** - Organizing workshops that engage communities directly, leveraging local languages and cultural touchstones to foster media literacy.
- **Public awareness campaigns** - Launching campaigns highlighting the importance of media literacy, drawing on local influencers and public figures to reach a broader audience.
- **Cross-sectoral media literacy platforms** - Creating online platforms that provide accessible media literacy tools and resources, catering to the diverse linguistic and cultural groups in the Balkans.
- **Ongoing professional development** - Ensuring educators and media professionals have access to continuous training on the latest media literacy techniques and disinformation trends.

In conclusion, the EU is intrinsically interested in promoting media literacy in the Balkans as a bulwark against Russian disinformation. With its commitment to the region's stability and democratic progress, the EU is well-positioned to provide
sufficient and continuous funding for these efforts. Such support would ensure the sustainability of media literacy programs, enhancing their effectiveness and long-term impact. By investing in the critical thinking capabilities of Balkan societies, the EU can contribute significantly to the region’s resilience against disinformation.

**Societal and Policy Impacts**

Disinformation’s corrosive effect on society extends beyond the spread of false information—it undermines the very trust that binds communities and informs policy decisions. With its intricate history and diverse societies, the Western Balkans face unique challenges in this regard. Disinformation can exploit and widen ethnic and political divides, influence election outcomes, and sway public opinion against critical reforms and international alliances—all of which have been observed and documented in detail in previous sections.

To counter these effects, a two-pronged approach that includes developing media literacy programs and strengthening regulatory frameworks is imperative—both of which have been identified and analyzed in previous sections.

Drawing from Finland’s example, where media education starts early and continues throughout one’s life, the Western Balkans can integrate these practices into their education systems. Students would be taught to distinguish between disinformation and misinformation and understand the motives behind media messages and the techniques used to influence audiences. This education would extend beyond schools, involving community outreach and adult education programs, ensuring that all segments of society are equipped to navigate the media landscape with discernment and skepticism of dubious sources. Moreover, civil society organizations play a crucial role in this educational endeavor, as they can provide localized training and resources tailored to specific community needs. They also act as watchdogs, holding media outlets accountable and championing truthful reporting.

In tandem with education, robust regulatory frameworks are needed to create accountability for disseminators of disinformation. These frameworks should be adaptable, with enforcement mechanisms that can respond swiftly to emerging threats and regulations that evolve alongside new media technologies and tactics used by purveyors of disinformation.

Together, these strategies form a comprehensive defense against the impact of disinformation on society and policy. They not only serve to educate and protect the public but also reinforce the democratic processes that are foundational to the Western Balkans' aspirations for EU integration and regional stability. The success of such initiatives, however, will critically depend on sustained commitment and collaboration among educators, policymakers, civil society, and international partners.

Furthermore, the effectiveness of the proposed strategies for combating disinformation in the Western Balkans heavily depends on two critical factors.

First, the EU’s ability to extend its influence into Balkan countries that are not EU members. The EU can extend its influence into non-member Balkan countries through strategic partnerships and aid programs. By offering expertise, funding, and support for media literacy and regulatory reforms, the EU can help build capacities that align with its standards. Moreover, the EU can leverage its position as an aspirational model for these countries (and most of them are already on the EU
integration path), particularly in the promise of future integration, to encourage adherence to democratic practices and reforms against disinformation.

Second, **incentivizing Balkan countries with pro-Russian political leadership to undertake these transformations** presents a unique challenge. Countries with pro-Russian leadership might be incentivized through a combination of diplomatic engagements and benefits linked to reform achievements. The EU can employ a nuanced approach, offering economic and political incentives for adopting anti-disinformation measures while respecting their geopolitical alignments. This could involve conditional aid, access to certain EU programs, and support in sectors critical to these countries. Moreover, highlighting the long-term benefits of stability, enhanced governance, and international integration can be key in persuading these countries to undertake necessary transformations.

While the strategies and proposed framework for combating disinformation in the Western Balkans hold significant promise, their implementation will not be without challenges. It's a sobering reality that these efforts will likely face resistance, not just internally but also from external actors such as Russia, who have vested interests in maintaining influence in the region.

The process will likely be lengthy, demanding both time and financial resources. However, despite these obstacles, the long-term benefits of fostering a resilient, informed society and safeguarding democratic processes far outweigh the immediate costs and complexities. The task ahead is formidable, but the risks of inaction — continued societal division, erosion of trust in democratic institutions, and vulnerability to external manipulation — make a compelling case for steadfast commitment and strategic action.

**About the Authors:** Vibor Cipan is an independent researcher and analyst focusing on topics related to geopolitics, OSINT, all things data, AI, and technology. He tweets @viborc and writes at viborc.com. Contact: vibor@viborc.com. David Kirichenko is a researcher and expert on Eastern Europe. His analysis is published widely in publications such as the Center for European Policy Analysis, Atlantic Council, The Hill, Irregular Warfare Center, and peer-reviewed journals. He can be found on the social media platform X @DVKirichenko.


https://warsawinstitute.org/western-balkans-war-ukraine.


50 Euronews, “Night Wolves.”

51 Euronews, “Night Wolves.”


59 CRTA, “Nothing is True.”


68 Greene, “Mapping Fake News and Disinformation in the Western Balkans.”

69 Greene, “Mapping Fake News and Disinformation in the Western Balkans.”


Georgetown Security Studies Review 85 Volume 11 | Issue 2
Malaya to Vietnam: The British Counterinsurgency Model and Its Replication Challenges

Steven Hu

In the aftermath of World War II, a surge of anticolonial struggles driven by ideological motives broke out in the territories formerly under Western European empires. Despite a growing global condemnation of militarism and imperialism, the Allied victors of the war endeavored to retain their colonial domains, largely in response to the Soviet Union’s expanding influence and the looming threat of global communism. In several instances, external pressures and the burdensome costs of maintaining these colonies led to significant guerrilla victories, exemplified by the Vietnamese revolution in Indochina and the Algerian War of Independence. These successful bids for decolonization showcased the effectiveness of armed resistance against technologically superior forces, sparking more independence movements globally.

Contrastingly, in Malaya, the British Empire quelled a widespread communist insurgency, marking a divergent narrative in the decolonization era. Military historians hail the British approach, known as the Briggs-Templer Plan, pioneered by General Harold Briggs and Field Marshal Gerald Templer, as one of the 20th century's most effective counterinsurgency operations. Nonetheless, when U.S. military forces applied this model in the Vietnam War less than five years later, it led to contentious and mixed outcomes. This paper explores the context of the Anti-British National Liberation War, the intricacies of the British counterinsurgency strategy, its problematic replication in Vietnam, and the broader implications for modern conflict resolution.

Introduction

The concepts of guerrilla war and insurgency have become synonymous with modern-day conflicts — low-intensity three-block wars have become increasingly prevalent in the contemporary world. However, a greater examination of history would point to countless examples throughout military history in which insurgent forces employed unconventional warfare tactics against a numerically and economically superior foe. In the era of decolonization and great-power competition that emerged in the aftermath of World War II, low-intensity conflicts and counterinsurgency campaigns became prominent in former European colonial possessions. Eastern and Western blocs further exacerbated these wars, as indirect support for ideologically aligned paramilitaries allowed nations to continue projecting their power without direct involvement.

For the former British Empire, the Malayan Emergency, also known as the Anti-British National Liberation War (1948–1960), was the perfect storm of both. With great powers rising in Asia and the Soviet Union and the animus of the global proletariat revolution, Malayan communist revolutionaries aimed to topple the colonial regime and establish a socialist system encompassing economic and political aspirations. With an empire on the verge of bankruptcy and combat experience during the Japanese occupation, the Malayan National Liberation Army (MNLA) had a unique opportunity to realize the goals of Marxism in Southeast Asia.

However, within 12 years, the Malayan Emergency ended with relatively low civilian casualties and damage to the Malayan economy. While communism spread like wildfire in Indochina and China
proper, the British military and the colonial administration disarmed the MNLA and later peacefully transitioned into an independent state.\(^2\)

In response to the Malayan Emergency, military historians view the Briggs-Templer Plan as one of the most successful counterinsurgency operations in the immediate postwar era — this success led the United States to attempt similar tactics during the Vietnam War. Furthermore, modern-day military commanders, policymakers, and world leaders have also referred to Malaya as their model for the three-block war in the Global War on Terrorism, hoping to find solutions to our contemporary security challenges.\(^3\)

**The British Empire in Malaya**

The origins of British Malaya date back to the late 18th century; in the era of archaic globalization and the early modern period, European Empires sought the luxuries of the Indo-Pacific.\(^4\) Over the centuries, powerful joint-stock companies and crown colonies established their presence in the islands of Southeast Asia. These islands offered abundant spices and natural resources and strategic trade routes for maritime commerce.\(^5\)

The British Empire subsequently established its first venture in Penang Province in 1786 and extended its influence to neighboring Borneo, Singapore, and the Malayan Peninsula in the following years.\(^6\) During the turn of the 19th century, the colonial administration brokered a deal with the four major British protectorates to form the Federated Malayan States (FMS) to improve civil administration, with a headquarters element in Singapore.\(^7\)

Rich in both natural resources and situated in an area of geopolitical importance, British Malaya would prove to be a profitable venture for the British Empire and an important outpost for their power in the Far East. Natural resources such as tin and rubber became flagships of the Malayan economy, as, in 1947, their rubber industry alone had a greater export than the British home isles, raking in $200 million in export revenue.\(^8\) Furthermore, the colonial administration financed significant infrastructure projects in Malaya, increasing the land value 42-fold between 1889 and 1909. Additionally, the crown colony of Singapore served as a strategically important station for the British Empire. To protect their overseas empires from overzealous imperial rivals, Britain established a series of “fortress colonies” to defend their overseas possessions.\(^9\)

Dubbed the “Gibraltar of the Far East,” Singapore’s strategic importance was so vital that during the interwar period, Britain had devised that in the event of a major crisis, the Royal Navy’s main fleet would regroup at the island as their new staging area.\(^10\)

When Japan fully realized its imperial ambitions in World War II and launched a massive campaign to conquer all of Asia, Britain was unwilling to give up its regional hegemony in Southeast Asia. With multiple economically and geopolitically vital colonies such as India, Hong Kong, and Burma in the Indo-Pacific, the British would play an instrumental role in the Pacific War along with her American and Chinese allies.

**The Ethnic and Ideological Origins of the Malayan Communist Movement**

With the fall of Singapore and the subsequent Japanese occupation of Southeast Asia, World War II would empower the underlying racial and ethnic tensions that eventually produced Malaya’s communist anti-colonial mass movement. For centuries, ethnic Chinese enclaves had formed in the Federated Malayan States and Singapore, consisting of 40% of the population.\(^11\) The majority of Chinese-
Malays worked in labor-intensive trades such as coal mining and rubber farming, while a small minority of the educated business class became part of the wealthy elite.\textsuperscript{12} Despite decades of migration, the Chinese diaspora in Malaya was also very connected to their ancestral homelands in China Proper. Dr. Sun Yat-Sen founded his successful anti-Qing movement, the Tomenghui, in the city of Georgetown in Penang and attracted significant political support and donations from Chinese Malays to fund the monumental Xinhai Revolution in 1912.\textsuperscript{13}

Thus, when the communist mass movement first gained traction in mainland China in the 1920s, the diaspora in Malaya found resonance with Mao’s proposed proletariat revolution, as they were politically ostracized and treated as second-class citizens by ethnic Malays.\textsuperscript{14} Since the British only provided schooling for the indigenous population and not the Chinese, communist literature and teachers made their way to Chinese communities in colonial Malaya and Singapore.\textsuperscript{15} This inequality led to the formulation of several communist entities, such as the South Seas Communist Party, the Malaysian Communist Party (MCP), and the Malaysian General Labour Union (MGLU).\textsuperscript{16} Along with the growing Vietnamese communist movement led by Ho Chi Minh in Indochina, Marxist-induced anti-colonial sentiments would begin to grow quietly under the nose of European empires. Although the British became concerned with the spread of communist ideologies in the 1920s and early 1930s, they did not pose a major threat to its rule. The reality changed after the Japanese invasion on December 8th, 1941, in which the Japanese Empire swiftly conquered Malaya, North Borneo, and Singapore from the British.

With the British Empire defanged in the Far East, communist guerilla armies were the only means of resistance against the now juggernaut Japanese Empire. By this point, the MCP had formed a well-organized resistance of mostly Singaporean Chinese and became known to American and British leaders as the most effective anti-Japanese guerilla force in the region. In 1943, the MCP formed a temporary alliance with the British to counter the Japanese occupation of Malaya. It remains uncertain whether the British were fully aware of the MCP’s ultimate goal: to unite with Soviet Russia and China in supporting the independence of smaller races in the Far East and to assist the Japanese people in their anti-Fascist struggle.\textsuperscript{17}

To support MCP activities, the United Kingdom deployed several intelligence-led special operations military units to assist the Malayan People’s Anti-Japanese Army (MPAJA), the military wing of the MCP, with their guerrilla campaign.\textsuperscript{18} Support included training in tactics including subversion, intelligence gathering, and unconventional warfare tactics, as well as direct material support by airdropping more than 1.5 million pounds of equipment in the last eight months of the war. By the time Japan surrendered, the MPAJA had formed eight independent regiments consisting of around 10,000 fighters.\textsuperscript{19} Women also became integral to the MCP’s anti-Japanese activities by becoming propagandists organizing to proselytize the masses to the anti-imperialist cause in both rural and urban areas, mainly in the provinces of Penang and Perak.\textsuperscript{20} MCP female organizers also successfully formed trade unions, taught night classes, and eradicated literacy, resulting in female representation in leadership positions and favorable portrayals of the communist movement in the public conscience.
This cooperative effort would deteriorate after the World War II when the MCP no longer served as a strategic benefit to Britain and became an obstacle in its foreign policy. The communist movement in Asia gained significant popularity as former European colonies sought independence from their deteriorating colonial rulers and governments supported by the West. On the Korean Peninsula, communist revolutionary Kim Il-Sung founded the Democratic People’s Republic of North Korea after its partition on the 38th parallel, and the Indochinese Communist Party Chairman Ho Chi Minh launched an armed uprising in French Indochina. On mainland China, the collapse of negotiations led to the resurgence of the Chinese Civil War, a conflict sharply defined by the opposing national and communist parties, each receiving explicit support from the world's emerging superpowers, the United States and the Soviet Union, respectively.

In this geopolitical context, Britain faced a dilemma in this new world order, as the empire in which the “sun never sets” faced an unprecedented identity crisis. Although it emerged victorious after World War II, it came at a heavy human and economic cost; nations no longer respected London as the formidable superpower that “ruled the waves,” including the Malays who watched the Far East Command in Singapore surrender to Imperial Japan. Furthermore, Britain's empire status faced significant ambiguity; the new world order led by the United States renewed its former Wilsonian principles, which prioritized self-determination and discouraged the prospects of empire. Theoretically, it would mean that world powers would not get tangled in the domestic affairs of the post-colonial world and allow them to work out a suitable form of independent governance. Japanese and German imperialism left scars across all continents, leading to the worldwide consensus that colonialism is no longer an acceptable practice in contemporary and globalized society. Britain's fall from grace as the world's leading superpower, coupled with the forthcoming Cold War and new norms and conditions that would strip away most of its prized overseas possessions would be a daunting challenge for London.

With the emerging threat of communism in the Cold War, the British halted decolonization efforts and justified their colonial administrations as America sought the Anglosphere as bulwarks against the Soviet Union and Communist China. In Malaya, the MCP renewed their anti-imperialist ambitions and finally had the resources to carry out their proletarian revolution. With leftover equipment from World War II, premiere military training by British commandos and intelligence operatives, and Chinese ethnic unity forged by the Sino-Japanese War and Mao’s mass movement on their ancestral homelands, the Malayan Emergency would finally begin.

When the British sought to reestablish Singapore’s fortress status in 1945 and reclaim its hegemony in Southeast Asia, the MPAJA became the de facto power in many regions as conventional Allied troops only attacked as far as Borneo in the Pacific Campaign. Armed to the teeth with developing left-wing organizations and sentiments in Malayan society, the MCP initially hoped for the empire's peaceful and tactical retreat and to establish independence without bloodshed. They pursued a policy of a "united front," working with ideological allies such as bourgeois parties and using mass labor strikes to discourage the colonial administration from continuing an unprofitable investment.

The British, however, decided to hang on to its colonial possessions and did not accept the initially peaceful attempts at decolonization. This led the MCP to turn to
more violent measures, foreshadowing the forthcoming guerilla war. Between 1946 and 1948, Malayan communist insurgents used more coercive measures such as abductions, torture, violent rioting, and targeted assassinations.24 The MPAJA had also rebranded itself as the Malayan People’s Anti-British Army (MPABA) and began to specifically target rubber plantations and palm oil farms to disrupt the source of the crown’s profits.25

With British business owners under threat and escalating political violence, London decided to finally intervene. An emergency was officially declared in June 1948 by High Commissioner of British Malaya Sir Edward Gent, officially commencing the guerilla war between the British Empire and the Malayan communists.26

Emergency, not War

The conflict in Malaya was called an emergency for particular reasons. The Rubber Growers Association (RGA) was worried that the insurance companies would not protect businesses in the event of a civil war or rebellion, which would be detrimental to the industry and undermine Malaya’s ongoing economic diversification efforts as it relied on the capital generated from the rubber industries.27 British investments had reached almost 10 million pounds at the beginning of the insurgency — if there were a declaration of war, compensation for damage would become the government’s responsibility.28 Furthermore, the conflict designation would also dictate rules of engagement and the measures the British colonial administration could enforce to disperse the communist movement.29 Declaring martial law rather than an "emergency" would require a military government enacting strict emergency powers. This move could prove unpopular and detrimental since the Federated Malayan States relied on an indirect governance model.30 Thus, the British colonial administration adopted terms such as “terrorists” and “bandits” to deter escalations of “war.”31 The conflict, therefore, became known as an "emergency" to protect vital colonial interests and enable military mobilization, but also mitigate the negative externalities of economic and political fallout.

The Counterinsurgency Campaign

Despite the jungle environment being ripe for a guerilla war and many considerable obstacles to success, the British colonial administration achieved victory in a decade with minimal casualties. In 1950, Sir Harold Briggs, a career military officer appointed the successor to Edward Gent as the High Commissioner of British Malaya, outlined a strategy known as the Briggs Plan. This playbook would focus on three baskets: separate the guerrillas from the people, intelligence as a key to subverting guerilla operations, and effective deployment of territorial security forces.32 In the coming years, the Briggs Plan would prove to be an effective multi-faceted response that saw the efficient use of political and military intelligence, conventional and irregular warfare tactics, civil-military cooperation, and political warfare.

The major challenge in any three-block asymmetrical war is separating the civilian population from combatants. Collateral damage has been a historic way to proselytize apolitical bystanders to take up arms, and Malaya’s terrain would prove to elevate that obstacle. The urban environment is extremely dangerous for uniformed combatants, as the enemy has the advantage of blending in with a dense population and attacking built-up structures from all directions. The jungle, which would eventually become the proving grounds of the conflict, would come with similar threats; thick foliage and unforgiving terrain...
make it easy for guerilla armies to plant booby traps, conceal their forces, escape, and regroup from engagements as well as launch coordinated ambushes.

The Min Yuen also aided the MCP, a communist political organization that provided logistical support for the MNLA. This organization was comprised of ethnically Chinese squatter communities who fled to the fringes of the jungle to avoid Japanese labor camps and were ideologically driven to liberate Malaya from British occupation. Min Yuen agents helped recruit new fighters, smuggle and transport food and medical supplies, as well as dealing drugs and contraband to fund the war effort.

To combat this logistical network, the British colonial administration resettled over 400,000 Chinese squatters from jungle communities into "new villages" where the colonial administration could effectively police and control their movement. Relocated squatters were granted an ownership title of their newly settled land and also compensated for their relocation and anything that could not be moved, such as crops. Additionally, the civil administration prioritized providing infrastructure and services, including schools, water, and electricity, to ensure that the new villages did not resemble concentration camps. Resettlement officers, predominantly of ethnic Chinese origin, addressed resident concerns about fostering municipal autonomy by establishing local police forces and militias. In some cases, to prevent settlers from aiding the communists, the military erected installations such as guard towers, checkpoints, gates, and pillboxes. Furthermore, civil-military efforts dismantled logistic networks and enemy movement through a series of emergency regulations. These regulations included restricting transport on roads, arresting individuals without trial, and registering the population with identity cards.

Due to the ethnic composition of the conflict, many legal measures, albeit controversial, became highly effective in differentiating the civilian population from sympathizers and active communist guerrillas. Although the MNLA and Min Yuen were almost entirely Chinese, they still had the advantage of being an ununiformed fighting force that could disappear into the populace. A mass detention campaign between 1949 and 1951 detained over 110,000 without trial, followed by a controversial deportation campaign that explicitly targeted ethnic Chinese — most of whom did not possess citizenship. Approximately 26,000 were deported to Nationalist China, with another additional 3,000 repatriated to CCP-controlled mainland China after 1949.

Additional measures to distinguish insurgents and political operatives from civilians, police, and military forces included issuing over one million identification cards and enabling checkpoints and cordons to operate more effectively. The MNLA attempted to disrupt the system with forgery, disobedience, and destruction but was ultimately unsuccessful due to British counterintelligence efforts. By denying the MCP and MNLA freedom of movement and their primary advantage of blending into the civilian population, the colonial administration was able to free up their arsenal and focus on defeating the insurgency with the threat of collateral damage minimalized.

Coordinated intelligence efforts between police and military entities were also essential factors to the success of the Briggs Plan. At the onset, British intelligence faced many challenges. Britain had never fought a
counterinsurgency and had just spent the last few years combating conventional Japanese and German forces in large-scale naval and land maneuver warfare in clearly defined battlefields. Intelligence agencies also lacked cooperation, as the Security Intelligence Far East (SIFE) in Singapore and the Malayan Security Service were both unprepared for an escalation in violence. A new intelligence branch was created by Briggs to combat this deficiency, bringing in MI5 talent from the UK and operatives that had served in other British colonies and specifically Palestine. The intelligence relief effort consisted of Colonel Arthur Young, the Commissioner of Police of the City of London, and Director of Military Intelligence General Gerald Templer to succeed Gurney as the High Commissioner of Malaya. Templer was an experienced soldier by all measures, especially in intelligence and counterinsurgency. He served in the Mesopotamia campaign in the immediate aftermath of World War I to prevent the spread of Bolshevism and saw guerilla war in British Palestine first-hand. He then oversaw civil affairs command and control in West Germany after World War II and held many other staff positions in the empire's military and security apparatus before arriving at his new post in Malaya.

The Briggs-Templer Plan

With new leadership at the helm in 1952, many noted this year as a turning point in the Emergency, with Templer implementing his strategy within the framework of the Briggs Plan. He fostered greater collaboration between police and military entities, delivered specialized training to field officers, and created local field branches to provide accurate intelligence on a tactical level. This apparatus was expanded across British colonial possessions to combat communist movements, as intelligence practice and doctrine would be elevated outside the scope of the Emergency and into the greater Cold War context. The development of a police intelligence capability was also critical to mission success. The core focus of the Special Branch, the police intelligence apparatus, was to provide a counterintelligence capacity to decipher the MCP order of battle, command structure, and general activities. The effort succeeded by recruiting ethnic Chinese linguists and turning defectors into interrogators for the police, thereby curating a pool of informants. The Special Branch was also effective on the tactical level as they participated in site exploitation, collecting operation intelligence, and liaising with military units to coordinate information on the battalion, brigade, state, and district levels.

Winning hearts and minds through psychological operations were at the core of Templer's intelligence activities. He understood that whisking away ethnic Chinese support for the MNLA and MCP would doom the movement unconditionally. It was not enough to make the dichotomy between combatant and civilian, but wearing down popular support for the communist movement was essential for victory. Speaking to a crowd of public servants, Templer explained, "You cannot win this battle ... without information. The people who can give you that information are ordinary, simple people. They refuse to give you that information in more cases than not because they do not see that you are winning ... It is this confidence that is needed." In a show of peace through strength, the General would travel through new villages and rural Chinese communities in armored convoys but step out in bustling markets and town squares in his dress uniform as a show of force. He would then interact with the locals as if he were a campaigning politician - listening to their concerns, striking up
casual conversations, and even "pick up some bouncing baby in his arms." During Templer’s leadership, alternatives to the communist movement were established to break the MCP’s political monopoly on the Chinese population. The congregation of the Sino-population in “new villages” led to the creation of the Malaysian Chinese Association (MCA). This political party sought to represent the interests of ethnic Chinese, cultivate positive relations with the Malay government and the British colonial administration, and specifically “cleanse the negative image of the Chinese produced by the terrorist acts of the Malayan Communist Party (MCP), whose members were mainly Chinese.”

Founded by Kuomintang affiliated leaders Leong Yew Koh, Colonel H. S. Lee, and Tan Cheng Lock, the MCA renewed Dr. Sun Yat-Sen’s Nationalist movement that ethnic Malays had once supported fervently in 1912 as a counterweight to the fervent anti-communism adopted by the mainstream diasporic community.

Additionally, Briggs oversaw the creation of the Emergency Information Service (EIS) to wage a coordinated large information warfare campaign against the MNLA through the effective use of propaganda techniques, dropping almost 200 million pamphlets between 1953 and 1956. Traditional media, such as newspapers, radio broadcasts, and loudspeaker messages urging MNLA fighters to surrender, were also effective in lowering enemy morale and causing defections.

The use of Indigenous troops also expanded under Briggs’ and Templer’s communist eradication efforts under the strategic goal of “winning hearts and minds.” The use of local forces had not been unique to warfare at this point, as the benefit of having permanent security forces who understood the local culture and language was invaluable to colonial and invading armies from Alexander the Great to Dwight D. Eisenhower. Creating a “home guard” was imminent starting in 1950, as colonial police services included ethnic Chinese recruits in their year conscription. Although a small contingent, these constables would free up military units to tackle larger engagements and operations with the MNLA. By 1953, home guard units of mostly Chinese amounted to 250,000 soldiers. The combat effectiveness of home guard units was questionable, but it showed that the British were willing to trust the locals with their security, which helped disarm anti-imperialist sentiments.

The third basket of Briggs Plan involved direct action from the military and fighting the MNLA directly. As aforementioned, the Malayan operating environment was incredibly complex, and the British lacked experience fighting in thick jungles without defined belligerents. With police and intelligence units operating in the background to deter radicalization and help define the battlespace, the military was able to focus on their primary objective: seek out and kill or capture the enemy.

The military’s peak strength included 30,000 troops consisting of Ghurka, Australian, Fijian, and New Zealand battalions and regiments. Special forces units such as the New Zealand Special Air Service and Royal Marine Commando deployed alongside indigenous tribe members from Borneo to track and patrol the thick Southeast Asian jungle. A homegrown Multiracial Malayan Army was also created, with cadres of officers sent to the United Kingdom for training.

A great deal of the military’s success in the Malayan Emergency was their devolved command structure, as Headquarters and higher echelon leaders emphasized operational and tactical command and
control. During the early stages of the Emergency, the Far East Command in Singapore played a significant role in both military and police operations at all levels. With the arrival of Briggs and Templer, more autonomy was given to elements. This meant that if tactical units aligned with the commander’s intent, headquarters and staff elements were ready to supply all necessary resources, including training and equipment.

Moreover, specialized units concentrated exclusively on their well-defined combat operations, remaining focused on their specific mission tasks without being sidetracked by other objectives, such as "hearts and minds" campaigns or civil-military cooperation efforts by the conventional armies. Units deployed at the platoon level with a greater focus on small-unit engagements with the enemy, proving to be effective in countering MNLA hit-and-run tactics designed to ambush overstretched convoys and larger formations.

The British military also demonstrated efficiency in combined arms tactics by incorporating both the Royal Air Force and Navy into their ground operations to create a coordinated symphony of death against the MNLA. The Royal Air Force did not engage in carpet bombing as it was easy to avoid aerial ordnance in a jungle environment, but once again focused their efforts on patrolling and intelligence. Aerial photography and air-to-ground communication helped land forces identify enemy troop movements with greater situational awareness and accuracy. Airpower was also effective in deploying paratroopers and supplies deep in enemy territory, as well as immediate medical evacuations for the wounded. Maintaining aerial superiority allowed units to operate outside the wire for prolonged periods without having to return to base for resupply or CASEVACs. The Navy was given the task of patrolling the coastlines to deny the transport of goods and movement by the communist insurgents and, at times, provided logistical support by landing amphibious troops and supplies on Malayan beaches.

By July 1954, two years after General Templer assumed command, monthly rates of engagement declined 37% from their peak in 1951 as the MNLA insurgency began to crumble. One year later, The Malayan Federation and the Government of Singapore began to offer amnesty for communist guerillas as the Emergency was winding down. In 1957, the Malayan Federation declared independence, meaning that the anti-colonial origins of the MCP were no longer valid. By that point, The Min Yuen and MNLA had a combined strength of less than 2,000, and their remaining true believers fled to the jungle or neighboring Thailand in hopes of securing a haven in Maoist China through other communist organizations. In 1960, the Newly formed Malayan Federation declared the Emergency over, concluding both the British colonial era and guerilla war in Malaya.

**Replicating Success - The United States in Vietnam and Beyond**

Since the end of the Emergency, there have been multiple attempts to replicate the British response’s success in other counterinsurgency campaigns. The Malayan “blueprint” was superimposed almost immediately by the United States during the Vietnam War, only four years after the conclusion of the Emergency. Furthermore, with the Western world consumed by a global campaign against terrorism, many scholars, politicians, and military leaders have looked to Malaya as a way to solve their military and geopolitical challenges.
into how to integrate law enforcement, military, and intelligence assets in counterinsurgency operations and three-block wars. To this day, we see many of these cooperative measures in operations against non-conventional actors, such as the emergence of fusion centers and overseas deployments of joint inter-agency task forces. However, a critical weakness in directly applying these tactics lies in failing to account for different operational environments and doctrinal warfare approaches. Military leaders and strategists often caution against preparing for the last war, as it may lead to failure in the next. Despite this, Western militaries frequently adopt previously successful strategies for new conflicts, ignoring the dynamics of evolving geopolitical contexts. While understandable, this bias towards proven methods often sidelines the need for operational innovation. This lack of adaptability, favoring tactics from past conflicts, has contributed to significant military setbacks in the 20th and 21st centuries. The adage 'no plan survives first contact' aptly describes the outcomes of contemporary interventions in the Middle East.

Typically, strategic adjustments have stemmed more from initial failures than from implementing a robust, environment-specific doctrinal framework. As the United States became gradually involved in the war in Indochina, American military leaders looked to the Briggs-Templer Plan as the strategic framework. Adopting this doctrine was logical, as the newly partitioned South Vietnam was another post-colonial state in Southeast Asia that struggled internally with a burgeoning anti-colonial communist movement. With limited experience in counterinsurgency in the Asia-Pacific, America looked to the experiences of their British allies for guidance.

U.S. advisors attempted to replicate the "new village" concept through the "Strategic Hamlet Program" in the early stages of the Vietnam conflict. However, this initiative by Washington became an abysmal failure. Not only were Templer's tactics morally questionable (and passionately debated to this day), but they also neglected the difference in population and culture. The new villages resettled ethnic squatter communities and had no hereditary nor legal ties to a specific region, as they were an ethnic minority that had migrated to Malaya. Contrast that with the South Vietnamese strategy of uprooting indigenous Vietnamese villagers into new communities where their security and freedom would be deprived. Furthermore, the plan was far too ambitious in contrast to Templer’s model; Saigon hoped to resettle over one million inhabitants in over 5,000 villages, which was not only financially and logistically unfeasible but also stretched security forces thin and unable to provide security for the locals.

Furthermore, the failure in the case of Vietnam did not account for the role of ethnic dynamics in Malaya. The Emergency, although driven by communism, was also fomented by the frustrations of subjugated Chinese Malays. Winning hearts and minds were geared towards the diaspora population, as most ethnic Malays had come to accept the status quo of British indirect rule for centuries. In contrast with Vietnam, the battlespace had defined battle lines with the 17th parallel. Everything north of that boundary was true mass believers who believed vehemently in anticolonialism, and a plea for ethnic unity would be futile. While Malaya had many ethnic local leaders under the supervision of the colonial administration, South Vietnam was effectively a puppet state that sought to promote the remnants of French religion, government, and culture that the majority
did not want. Plus, while the MNLA were almost completely isolated in their ideological struggle, the North Vietnamese received training, material aid, and diplomatic and political support from rival world powers such as China and the Soviet Union.71

In conjunction with superimposing the Briggs-Templer methodology without accounting for variables, the Vietnam experience also neglected some of the most critical counterinsurgency tactics essential to winning the war. While the British had used air power with restraint and defined mission sets, American forces engaged in various large-scale bombing campaigns. One of them is Operation Rolling Thunder, a three-year-long bombing campaign that dropped 643,000 tons of bombs in Vietnam, which led to minimal enemy casualties and no significant change in the fighting spirit of the enemy.72 On the contrary, extensive U.S. bombing campaigns unintentionally led to the radicalization of many civilian populations in joining communist guerilla forces, with the most notable being the rise of the Khmer Rouge in the Cambodian countryside as a result of Operation Menu. This extensive carpet bombing campaign saw 214 tons of bombs being dropped in a four-year period.73 This was based on Chief of Staff General Curtis LeMay’s controversial belief on enemy morale, that overwhelming airpower through excessive aerial firepower will not only eliminate high value and payoff targets but also reduce the enemy’s willingness to fight, a theory mainly grounded on his personal experiences in commanding major carpet bombing campaigns in World War II and Korea. This theory saw a short reemergence in the initial stages of the Iraq War as “Shock and Awe,” a strategy that asserted massive displays of military strength, including dominating airpower and the use of heavy ordnance, would deter the morale of the Iraqi Army.74 Already skeptical of its assertions, this practice quickly went out of favor in the Bush administration after lacking clear evidence of operational success.

Beyond the Vietnam experience, British veterans of the conflict also considered their successes in the Malayan Emergency as a proven model for conducting irregular warfare that could be applicable in other colonial uprisings. General Frank Kitson, who would go on to command the Parachute Regiment in Northern Ireland, advocated for replicating these tactics to quell other rebellions in the British Empire.75 The prioritization of a political solution concurrently supported by precise targeting operations, minimal use of firepower, and the use of civil affairs capabilities would be characterized as a “distinctly British approach to counterinsurgency” by war historian Thomas R. Mockaitis, asserting that Harold Briggs and Gerald Templer successfully adapted to the evolving standards for humanitarian law, the rules of war, and liberal democracy in the 20th century.76 However, the operational history of British military involvement in the post-colonial world illustrates a very different story, as the lessons learned from the Malaya experience also led to numerous failures. Under Kitson’s command in the Troubles, the Parachute Regiment fired upon unarmed Catholic protests in what became known as Bloody Sunday, significantly escalating the conflict against Irish separatists. Furthermore, they failed to suppress the Marxist uprising in Aden, paving the way for South Yemen’s independence. Operations in Cyprus, likewise, led to mixed results as the conflict ended with a political compromise between the various factions.

The Briggs-Templer Plan should not be interpreted as an outlier in the history of
irregular warfare and counterinsurgency. Rather, the key lesson from the British strategy in Malaya is not found in its tactical triumphs but in its mastery of operational art. The United States Army and the Marine Corps define operational art as a cognitive application and integration of war principles to achieve a campaign objective, supported by the “skill, knowledge, experience, creativity, and judgment” of the commander and their staff. Originating in the Soviet Union, operational art evolved into the overarching guidance for military planning in the West. This approach adapts to the increasing complexity of modern armed conflict, encompassing peripheral domains such as political objectives, information and psychological warfare, and economic resilience. Rather than directly applying military doctrine and strategy, operational art calls for the commander to determine “centers of gravity” to inform their tactics by examining all relevant factors. In the case of Malaya, Templer identified ethnic tensions and the popularity of communism as their “centers,” resulting in a strategy to isolate the adversary through military action while formulating political solutions concurrently. Therefore, the Briggs-Templer Plan or any other “textbook” case study of counterinsurgency or irregular warfare strategy cannot be directly applied to conflicts of similar anatomy, as the center of gravity differs with each environment. This is especially true for the coalition experience in the Global War on Terror; although there was an overarching ideological trend across the conflict zones in the Middle East, the “center of gravity” in each battlespace varied dramatically. While the political grievances were divided across ethnic and tribal lines in Afghanistan, Iraq’s source of sectarian violence was rooted in opposing religious interpretations. Thus, despite being under the same umbrella of a wider counterinsurgency campaign against transregional radical-Islamic militancy, each theatre had its own distinct challenges that demanded a nuanced understanding for effective operational design.

Conclusion

Although The Briggs-Templer Plan during the Malayan Emergency remains an example of one of the most effective counterinsurgency campaigns in military history, its replication challenges, especially in the case of the Vietnam War, underscore a fundamental flaw in strategic military thinking. As wars become increasingly complex with evolving political, economic, and social dynamics, doctrinal frameworks can quickly become incompatible with the operating environment. Therefore, the true lesson from Malaya may not lie solely in its tactical victories against the MNLA, but in the broader understanding that each counterinsurgency effort demands a unique, thoroughly considered approach, tailored to its specific geopolitical context.

About the Author: Steven Hu is currently entering his final semester for his Master’s Degree in Applied Intelligence. He received his Honors Bachelor’s Degree in International Relations and History from Victoria College, University of Toronto. His research interests include 20th-century history, civil-military cooperation, and Canadian foreign policy. In his free time, Steven enjoys reading, watching football, and confusing Americans with his Canadian accent.


5 Keat Gin, Ooi. Malaysia and the Cold War Era. Routledge Studies in the Modern History of Asia Series, 2020, 6

6 Keat Gin, 6.

7 Keat Gin, 26.


13 Komer, 4.

14 Komer, 4-5.


16 Komer, 5.

17 Mumford, 27.

18 Mumford, 28.


20 Mumford, 28.

21 White, 172.

22 Komer, 34.

23 White, 173.

24 White, 19.

25 The name MPABA was changed to MNLA shortly after the war began

26 Komer, 40.

27 Komer, 37.

28 Komer, 35.
44. Komer, 38-39.
47. Ramakrishna, 121.
48. Mumford, 39
49. Mumford, 39
50. Komer, 42.
51. Komer, 43.
52. Komer, 44.
53. Ramakrishna, 125.
54. Ramakrishna, 125.
55. Ramakrishna, 126.
56. Keat Gin, 62.
57. Mumford, 41.
58. Mumford, 36.
59. Mumford, 37.
60. Mumford, 37.
64. Komer, 50.
66. Komer, 52.
68. Hack, 228.
70. Tilman, 417.
71. Komer, 80.
Tightening the Screws: Examining the Efficacy of U.S. Sanctions Against Russia Amid the Russo-Ukrainian War

Nick Lekkas

The United States’ wide-scale sanctioning of the Russian economy in the aftermath of its brutal full-scale invasion of Ukraine is experiencing success. Despite intense debate on the efficacy of sanctions as a tool of American foreign policy, continuous efforts by Russia to circumnavigate sanctions and export controls, and the war in Ukraine showing no end in sight, sanctions are meeting their intended objectives. U.S. sanctions against the Russian Federation are a key tool in Washington’s economic arsenal as it seeks to punish Moscow for its role in a war that has decimated Ukrainian cities and slaughtered thousands upon thousands of innocents, all in a thinly veiled attempt at imperialistic expansion. In the nearly two years since the U.S. Department of Treasury and Department of State issued their first salvo of sanctions against Russian banks in 2022, Russia’s economy has lost substantial levels of revenue, been nearly completely isolated from the U.S. dollar, and is forcing itself to adjust to a wartime economy. Arguments continue that sanctions have not forced Russian President Vladimir Putin to turn his troops around and return home, and they will not; rather, sanctions will force Russia to have to make costly choices at the expense of either its war effort or its own people. While the true extent of the damage of Western sanctions may not be fully realized, we are only now beginning to see signs of severe economic harm that will only increase over time.

Introduction

Russia’s full-scale invasion of Ukraine on February 24, 2022, set forth a course of events that altered the world’s geopolitical landscape. Amid gut-wrenching scenes of World War I-style trench warfare in the 21st century, the United States has waged its own form of war against the Russian Federation since 2008 through the imposition of sanctions. Questions continue to be raised regarding the efficacy of the sanctions imposed upon the Russian Federation, often buoyed by common criticisms of U.S. sanctions against other adversarial countries. This analysis shows the unprecedented U.S. strategy is resulting in measurable impacts with goals and objectives misunderstood by both media sources and the wider public.

This paper examines the following questions: first, are sanctions against Russia key to U.S. national security interests? Second, what are the goals of these designations? Third, what are the common critiques of sanctions against Russia in response to its invasions of Ukraine, and do they lack substance? Finally, which areas of improvement exist within the United States’ sanctions program, and how can Washington act upon them?

Russia’s economic pillars are instrumental in continually fueling its war machine as the fighting in Ukraine approaches its third year. U.S. sanctions are increasingly cutting off vital revenue streams, import channels, and billions of dollars of funds previously available to Russia’s economy. As a direct result of sanctions, Russia is facing a devalued ruble, fewer clients, and increased logistics, transportation, and trading costs.

Policy Objectives and Tactics

As of February 2022, no program held the size, scope, and efficacy as that of “Russian Harmful Foreign Activities Sanctions,” one of the 38 sanctions programs listed on the website of the U.S. Department of
Treasury’s Office of Foreign Assets Control (OFAC).¹ Using Executive Orders 13660² and 13661³ in 2014, and then 14024⁴ in 2021, the United States (as of September 2023) had sanctioned a total of 1,694 individuals and 2,043 entities, per Ukraine’s “War and Sanctions” tracker.⁵ From an American policy perspective, these actions have aimed to benefit from a Russian economy that is not operating at its full capacity. A weakened Russian economy aims to limit its military options and ability to leverage its hard and soft power as it once could, allowing America to focus on other key foreign policy areas. Furthermore, it sends a message to both authoritarian powers and democracies around the world that foreign policy goals cannot be fulfilled by military force to seize territory. It is critical, therefore, from both economic and military perspectives that the United States continues to hold Russia accountable for its actions through sanctions.

Sanctions are a powerful mechanism through which the United States can inhibit Russian economic power and cut off key revenue streams from the West through the sheer power of the U.S. dollar. Most foreign businesses and financial institutions do not wish to run afoul of the Department of Treasury or find themselves on the receiving end of secondary sanctions, so they will often endeavor to comply with sanctions measures. They also act as a reminder of America’s steadfast support for Ukraine and highlight Russia’s role in the war. Press releases from both the Department of Treasury and Department of State serve as official notices of designations to individuals, companies, and property; these function as highly visible messaging platforms, effectively a “name and shame” strategy. Some sanctioned individuals, such as Olena Shapurova, a puppet official in Russian-occupied territory in Ukraine, do not hold significant assets.⁶ Their designations instead allow for the United States to emphasize actions that violate human rights or territorial integrity, especially of its allies and partners, are not tolerated.⁷

Even though U.S. sanctions can be effective, there is great debate around such a claim. Every sanctions program is created or adjusted based on the whims of policy, which rely on varying strategies and objectives. Some, such as the sanctions against South Africa’s government in the 1980s, were viewed as successful in coercing a change in Pretoria’s apartheid behaviors and aligning its new approach with American interests.⁸ Others, like the sanctions imposed on Venezuela beginning in 2006, are viewed as unsuccessful due to their role in triggering economic hardship for the Venezuelan population and the failure to oust Nicolás Maduro from power.⁹ These failed sanctions are often referenced in critiques against sanctions on Russia, citing similar consequences such as a failure to coerce a change in Russian behavior or the periodic evidence of evasion and circumvention schemes. When applied to a relevant scenario, these arguments are valid and important points to raise when discussing the value of sanctions and their necessity as a policy response. They are not, however, applicable in this current circumstance.

The primary reason the U.S. government’s current Russia sanctions program is unlike any other is due to its sheer scope and scale; never has a country with an economy like Russia been designated at such scale and intensity. Therefore, any comparisons to previous sanctions programs must be made with caution. Simply citing unsuccessful sanctions in another country does not mean that sanctions against Russia are doomed to fail. Too many variables separate different
country programs, which must be considered when comparing one with the others.

Current Russian sanctions are incomparable even to the previous major round in 2014, following the Russian seizure of Crimea. The U.S. government’s current authority contains vastly more legal powers and perhaps most critically it includes allies and partners in imposing parallel sanctions. Establishing a united sanctions coalition, including the United Kingdom, European Union, Japan, Australia, and New Zealand, has been a breakthrough in the imposition and enforcement of sanctions against Russia. This united front has demonstrated solidarity and like-mindedness between Western partners, granting further legitimacy to the U.S.’s sanctions program.

A key misconception surrounding U.S. sanctions against Russia is they are not achieving the desired results because Russia continues to fight in Ukraine. While theoretically, targeted sanctions could convince a government to cease military operations, sanctions alone have not and will not convince Vladimir Putin to order his troops to turn around and come home. The more feasible goal is one that systematically cuts off critical channels of revenue for the Russian economy, such as oil, gold, diamonds, or precious metals (while not causing unintended effects to global markets); denies access to Western markets, components, and investments; holds malign actors accountable for the atrocities committed in Ukraine; and reduces Russia’s capacity to wage war at an industrial scale over time. While powerful and effective, sanctions alone are not likely to deliver a Clausewitz-esque decisive blow in economic form; there is no singular designation remaining that will end the war. Rather, they have an accumulative effect that causes incremental death by a thousand cuts. With fighting steadily approaching its second full year, the effects of those cuts are starting to show.

Sanctions against Russia have caused serious damage to its economy, and the data show it. According to calculations by the European Commission based on data compiled from Russia’s Federal State Statistics Service (Rossat), the Russian economy contracted by 2.1% in 2022, with notable shrinkages in key manufacturing sectors like vehicle production. A large-scale pivot in Europe away from Russian oil and natural gas deprived Russia of its largest customer for its key export commodity, resulting in its energy revenues falling by nearly half in the first two months of 2023, compared with 2022. A decrease in European consumption of Russian oil has resulted in a sharp collapse in gas production levels, purportedly at their lowest since 1978: “Independent Russian-language news outlet Agentstvo reported that Gazprom ‘has never had such a low production rate in its entire history’ and that ‘the last time there was a similar figure was in the Soviet Union in 1978,’ a year when 372.1 bcm were produced.” As of the end of September 2022, the ruble sat at approximately 96.2 RUB to 1 USD down from 142 RUB to 1 USD on 1 January 2022, which triggered desperate interest rate hikes by the Russian Central Bank to stem the bleeding.

Sweeping designations against key manufacturers and suppliers of Russia’s military-industrial base have inflicted compounded misery onto a massive defense industry deeply constrained by staggering losses to Russian military equipment. While intricate and obfuscated evasion schemes have emerged to bypass Western export controls and sanctions on Russian defense companies, Russia “still struggles to produce weapons at the pace required by the war in Ukraine.”
Beyond targeting the defense sector, the U.S. government has recently begun tightening the noose around Russia’s future energy projects, a critical development in preemptively ensuring that vital future revenue streams for the Kremlin continue to be constrained. Current energy production has remained relatively unscathed (aside from the imposition of a G7 oil price cap with mixed effects) to ensure the flow of oil and natural gas to states that need it and to mitigate market shocks. Energy designations pose a more complex conundrum for the U.S. and its partners and allies; Washington is still searching for the right balance between curbing the world’s reliance on Russian energy, constraining Moscow’s export profits, and ensuring that oil and gas keep flowing to the countries which need it most. However, key energy projects in Russia’s arctic regions have been designated, with a particular emphasis on “entities and individuals involved in the development of key energy projects and associated infrastructure, including Russia’s Arctic LNG2 liquified natural gas project” as well as entities involved in the procurement of materials and advanced technology for future energy projects.

The emphasis on sanctions imposed on these two cornerstones of Russia’s economy is made clearer by the ever-growing reliance of Russian state-owned companies and other sanctioned entities on complex and obfuscated sanctions evasion schemes. Their role in circumnavigating export controls and sanctions often raises arguments that sanctions against Russia have proven futile. This is because companies can still get their hands on vital components, precursor goods, and dual-use items, which are critical for large-scale production of weapons systems, technological services, or other strategic deficiencies. Reports across the internet tell of very similar strategies, in which third- or even fourth-party suppliers and distributors of goods of interest (often Russian-owned) purchase essential products from Europe or the United States before shipping them along to their eventual final destination in Russia. These procurement networks pose a major risk to greater sanctions compliance, with some estimates claiming that Russian missile production, for instance, has expanded beyond prewar levels.

The Department of Treasury, State, and Justice have honed in on Russian sanctions evasion as a top priority, evidenced by arrests and continuous rounds of sanctions on procurement agents and their suppliers. This amounts to a large-scale game of a “whack-a-mole” policy against sanctions circumvention, given the ease with which Russian businesses can set up companies in third-country jurisdictions like Hong Kong, London, or the United Arab Emirates. However, circumvention networks do not necessarily present net gains for Moscow either. Russia has been forced to pivot to alternate jurisdictions, especially the People’s Republic of China (PRC), and has had to layer and repeat transactions with arms dealers, smugglers, money launderers, or third-party re-suppliers. There are valid policy concerns about potentially increasing tensions with major powers like the PRC for their facilitation in Russian sanctions evasion. The PRC is likewise playing a cautious game of its own. The goods that are predominantly sent to Russia fall under the category of components and goods needed to build finished products in Russian factories and have rarely qualified as lethal goods. This is due to Beijing’s perception that crossing a red line could trigger larger-scale sanctions on Chinese entities that rely heavily on the Western financial system to operate successfully. Both alternatives have increased the likelihood of theft, faulty products, and exorbitantly higher logistics,
contracting, and transportation costs. These elevated costs may be temporarily covered by the Kremlin’s recent announcement it would double its defense spending. Still, with Russia reeling from its sovereign debt default in 2022, soaring costs will inevitably come back to haunt it. According to an investigative journalist who focuses on national security topics, “Russia’s 2022 default is likely to poison its interactions with international financial markets for a long time to come, even if sanctions on its central bank are one day lifted. While the sanctions on Russia’s central bank already block it from tapping Western credit markets explicitly, the default will see it also have to deal with creditor claims.”

Yet another frequent claim against sanctions programs is the civilian harm caused by economic constraints and restricted access to the U.S. dollar—whether through limited access to medicine, agricultural goods, food, or other critical human needs. This is simply not the case in Russia, nor will that likely be the case anytime soon. OFAC’s website contains at least 36 active Russia-related General Licenses which permit the continuation of certain transactions with sanctioned Russian entities. The rationale behind this is mitigating undesirable consequences such as civil aviation safety, marine rescue services, agricultural commodities, medicine, or telecommunications. These General Licenses play a crucial role in ensuring that the designations imposed on businesses and their owners will have their intended effects on blocking transactions from accessing the U.S. financial system but continue to allow the free flow of key goods to Russian citizens or recipients of similar imports. According to trade data, Russian exports of wheat and fertilizer have not been negatively affected by Western sanctions, even as some oligarchs who own major fertilizer companies have been designated. If anything, trade figures show an increase in exports compared to prior years.

At a more personal level, reports among the Russian populace claim that “a strong majority of Russians (70%) are not concerned by the sanctions imposed on Russia,…eight-in-ten Russians report that the sanctions levied against Russia have not created serious problems for themselves or their families,…[and] a majority of those who say they have experienced serious problems as a result of sanctions imposed on Russia express support for the actions of the Russian military forces in Ukraine (61%), though at a lower level than the overall population (76%).” Though Russian polling and surveys must be taken with a healthy dose of skepticism, it speaks to a broader sentiment that sanctions are targeting the upper stratosphere of Russia’s economy rather than trickling down to negatively impact ordinary civilians not in Washington’s crosshairs.

Concluding Observations

In short, U.S. sanctions against Russia are concentrated on long-term economic pain points that aim to degrade Moscow’s ability to wage war on its neighbors. There are multiple areas for improvement to tighten compliance and hamper circumvention efforts. Still, its unparalleled size, scope, and scale, this unified front against the Russian economy is proving more fruitful than critics give credit. Sanctions are a critical component of the United States’ national security interests when applied correctly, and in this context, are an important method through which Washington can apply significant economic pressure on the Kremlin for its brutal invasions of Ukraine. All this can be accomplished without the deployment of troops or the concerns that come with a kinetic engagement against a country like Russia. Valid critiques of American sanctions policy are a key part of
ongoing conversations of how to best utilize sanctions without unintended consequences against the global economy, and though not applicable in this context, ensure that policymakers in Washington feel pressure to continually revisit and revise their goals and objectives. Using general licenses to mitigate undesired economic effects on key commodities markets, vital services, or agriculture means that the likelihood of sanctions causing unintended economic hardships or humanitarian disasters in Russia is significantly lower. Blocking access to the U.S. dollar means that Russian firms must resort to complex evasion practices that increase costs and a higher chance of poor-quality products. The ruble has decreased sharply, further diminishing Russia’s economic power as a major currency holder and forcing additional increases in important costs for its military and industry. While forecasts and estimates indicate that Russia’s economy will not collapse, its stopgap measures are finite, and the continued pressure of U.S. sanctions reminds Russia that the cost of waging war is exceptionally high and will remain so. There remains plenty of opportunity for improvement – the oil price cap is beginning to show signs of enforcement. However, levels remain too high to call it a definitive success, foreign goods continue to flow into Russia through illicit schemes, and Russia’s economy is now on a war footing. With continuous and rigorous pressure from the U.S. Government on loopholes and obfuscation methods, those areas for improvement can be addressed properly. Whether Russia turns to energy, financing, or military means to continue to wage its war in Ukraine, it will find obdurate obstacles in its way.

About the Author: Nicholas Lekkas is a fourth-semester student at Georgetown University’s Walsh School of Foreign Service and a member of the Security Studies program, concentrating on international security. He received a B.A. from Wheaton College in Massachusetts in Russian Studies and has worked in D.C. for various government departments and agencies as both a government employee and contractor on a number of issues in the Russia problem set. Email: nml70@georgetown.edu.


7 “Territorial integrity” is an important concept emphasized by the U.S. Government, particularly by both the State Department and the Treasury in their sanctions press releases, and the concept itself is a subsection of both E.O. 13661 and E.O. 14024.


It’s All Connected: The Impact of Russian Sanctions on Global Trade Relationships

Ashley Kelso

After the 2022 Russian invasion of Ukraine, the United States and European Union, along with other Western allies, imposed a variety of sanctions on Russian and non-Russian citizens and businesses suspected of supporting Russia’s military campaign, expecting that crippling Russia’s military would end the war. Sanctions have not worked as intended and, instead, have pushed Russia closer to the People’s Republic of China (PRC) and other eastern nations, particularly in the energy and manufacturing sectors. Western companies accepting substantial losses to shutter Russian operations have redistributed Russian market share and lined the pockets of Russian businesspersons. Russia has looked to the PRC and India to replace energy exports previously sent to Europe, and countries like Saudi Arabia have been pulled into the fray despite not being party to Western sanctions. In many cases, sanctions have benefited U.S. adversaries more than they have supported the end of the Ukraine War. As policymakers look to impose sanctions in future situations, they would be wise to consider whether the secondary and tertiary effects make sanctions the best tool of power to exercise.

Introduction

After Russia invaded Ukraine in February 2022, the United States and the European Union (EU) partnered to execute a variety of sanctions that would penalize Russia for its actions. Expanding previously imposed sanctions during Russia’s 2014 annexation of Crimea, the goal of new sanctions was to isolate Russia from the global financial system and reduce profits generated from exports to reduce its military funding, limit its military power in Ukraine, and ultimately bring an end to the war. For example, the United States and the EU established sanctions on individuals and companies (both Russian and non-Russian alike) that were suspected of financing or otherwise contributing to Russia’s war efforts. The Society for Worldwide Interbank Financial Telecommunications (SWIFT) promptly removed Russian banks from its systems, restricting Russia’s ability to make international money transfers. The Price Cap Coalition, which includes the Group of Seven (G7) countries plus the EU and Australia, instituted a price cap of no more than $60 per barrel on Russian crude oil sales. The United States also banned the import of Russian fossil fuels and restricted the ability of U.S.-owned companies to invest in Russian energy companies or projects.

Sanctions against Russia are ongoing and likely to continue for years to come as Russia sustains its Ukraine campaign. The sanctions imposed may not have ended the Ukraine War as intended, but they have caused a ripple effect across the globe by altering existing trade relationships, pushing Russia closer to the East. Sanctions led many Western companies to end their Russian operations, which affected private sectors across nations and resulted in new partnerships between Russia, the PRC, and others, particularly in the energy and manufacturing sectors. Sanctions, along with Russia’s repeated use of energy as a political tool, also led many EU countries to accelerate existing plans to divest from Russian energy, resulting in the redistribution of Russian energy exports and EU energy imports and contributing to a souring of relations between the United States and Saudi Arabia over an increase in oil production. This shuffling of
relationships, in many cases, has benefited U.S. adversaries more than the sanctions themselves achieved the intended goal. As policymakers carefully consider the tools of power available to them in any given scenario, they will need to decide whether the secondary and tertiary effects make sanctions worth using at all.

**New Private and Public Partnerships Replace Western Firms**

By nature, sanctions are not limited to the public sphere; they restrict the ability of the private sector and individual citizens to do business with any entity designated by the U.S. Department of Treasury’s Office of Foreign Assets Control (OFAC). Private sector companies are generally careful to avoid punishment by the U.S. government when complying with sanctions, often avoiding whole sectors instead of only what is exclusively designated. In addition to the governmental pressures surrounding sanctions, companies also face pressure from their boards, shareholders, and the public to respond to social and ethical issues.

In the face of Ukraine’s breach of sovereignty and reports of Russian humanitarian abuses, along with initial sanctions imposed in the weeks following the invasion, hundreds of multinational companies were quick to announce they would temporarily or permanently end their operations in Russia, despite many shouldering extensive losses to do so. For example, Shell transferred its 27.5% stake in the Sakhalin-2 liquified natural gas (LNG) plant, accepting a loss of approximately $1.6 billion. Netflix, Sony, and The Walt Disney Company all suspended their media and streaming services in Russia. Starbucks initially paused all store operations and later announced it would sell its 130 Russian locations. McDonald’s acted similarly, temporarily closing almost 850 stores before selling its Russian franchise at a loss of at least $1.2 billion. Heineken and Carlsberg both made plans to exit Russia, taking on estimated losses of $400 million and $1.4 billion, respectively, while they looked for buyers. Other popular companies, such as Adidas, H&M, Nike, IKEA, Pepsi, Apple, Google, Microsoft, Uber, and Nissan, also committed to limiting goods and services sold in Russia.

At the approval of Russian President Vladimir Putin, its supporters, and many Russian businesspersons profited from purchasing these closed Western companies, generally at below market value, as companies looked to exit the country hastily. For example, in July 2023, the Russian government seized Carlsberg and directed the sale of Heineken to aerosol company Arnest. Russian restaurateur Anton Pinsky and rapper Timati (who in 2015 released a song titled “My Best Friend is Vladimir Putin”) purchased the 130 closed Starbucks stores for 500 million rubles (4.7 million euros). The pair reopened most of those stores shortly after their purchase using the name “Stars Coffee” and an extremely similar logo. With a reported Starbucks revenue of 47 million euros in 2021, they will likely quickly recoup their minimal expenses. Alexander Govor, who previously owned 25 McDonald’s restaurants in Siberia, purchased the 850 McDonald’s stores on the market to expand his restaurant empire. Besides restaurants, Govor also owns businesses in the forestry, fishing and hunting, and oil industries.

Given the connections between Russian industry and the state, it is very likely that the profits made from these new business endeavors will eventually make their way to supporting (if not outright funding) Russia’s efforts in Ukraine, either through the form of government-issued defense contracts or indirectly through donations to government officials or affiliated organizations.
More importantly, however, the redistribution of market share that resulted from the loss of Western companies in Russia left gaps across industries and contributed to establishing new partnerships, with Russia increasingly looking to collaborate with the East and the latter more than happy to fill the void. For example, Chinese retailer Alibaba increased its sales in Russia, while Xiaomi (coined “the Apple of China”) expanded its share in the Russian smartphone market. Chinese retailer Alibaba increased its sales in Russia, while Xiaomi (coined “the Apple of China”) expanded its share in the Russian smartphone market.  

16 Emirates and Etihad Airways also both took advantage of Western airlines halting flights in Russia and over Russian airspace, increasing the number of routes flown despite the dangerous environment.  

17 In the energy sector, Western companies BP and Shell withdrew from their Russian joint ventures after U.S. sanctions restricted the ability of any U.S. company to invest in Russian projects, leading Russia to actively solicit additional investment from the PRC and others. For example, China and Russia are currently planning the joint development of oil and natural gas fields and a second pipeline from the Yamal Peninsula into the PRC, dubbed the “Power of Siberia 2.” This pipeline would transport oil and natural gas from Siberia through Mongolia into northern China, increasing Russia’s fossil fuel exports to China while diversifying China’s natural gas imports and expanding China’s foothold in the Arctic. Russia’s second-largest natural gas producer, Novatek, is also partnering with Emirati company Green Energy Solutions for LNG technology and with Turkish company Karpowership for a “floating power plant” to support its Arctic LNG 2 project. Previously, the Arctic LNG 2 project, which would allow Russia to increase its LNG production and exports, was almost entirely contracted by Western firms, including Technip, Total, Linde, and Siemens. U.S. and EU sanctions imposed throughout 2022 resulted in these companies leaving the project, forcing Novatek, the majority owner, to find new suppliers. Russia’s Zvezda shipyard, engineering and construction firm Nipigaz, and Sberbank are all expected to support the project moving forward. The China National Offshore Oil Corporation (CNOOC) and the China National Petroleum Corp (CNPC) also hold 10% stakes in the Arctic LNG 2 project and will assist Russia in continuing the project on time.  

18 This is not Russia’s first time dealing with Western companies leaving due to sanctions. In 2011, ExxonMobil partnered with Rosneft, the Russian state-owned oil company, and agreed to invest $3.2 billion to share technologies and develop offshore reserves in the Kara Sea. In 2014, Exxon was forced to leave the project due to U.S. sanctions after Crimea, and in 2020, Rosneft finally resumed the project with funding from the Kremlin. With the constant churn of companies jointly owning and supporting Russian energy projects, its increasing partnerships with the East may provide more stable business while supplementing Russia’s industry capability and capacity.  

19 In the manufacturing sector, the PRC is supplementing Russia’s production capabilities by heavily importing “dual use” goods that can be used for both military and civilian purposes. This includes items such as drones, lithium-ion batteries, heavy-duty trucks, shovels and construction equipment, helmets, vests, radios, and other communication technology, most of which go to support Russia’s military in Ukraine rather than Russian civilians. Semiconductors are another important “dual use” good in which trade between the PRC and Russia has more than doubled since the start of the Ukraine War. Historically, Russia has lacked the domestic capability to
manufacture advanced semiconductors and has outsourced production to Taiwan Semiconductor Manufacturing Corporation (TSMC). In the wake of Russia’s invasion of Ukraine, however, Taiwan followed the United States in instituting sanctions that banned the export of sensitive technology to Russia. Such sanctions restricted TSMC— and other semiconductor manufacturers like the Dutch-owned ASML—from working with Russia, leading Russia to leverage its relationship with the PRC and move its chip production to the partially state-owned Semiconductor Manufacturing International Corporation (SMIC).

In 2022, the PRC exported over $500 million in chips to Russia, a significant increase over the $200 million exported in 2021. The benefits of “dual use” goods are two-fold: by exporting “dual use” goods, the PRC can bypass U.S. and EU sanctions with claims they are exporting civilian goods while contributing to Russia’s war efforts. With contributions from the PRC and other Eastern firms in the Russian energy and manufacturing sectors, Russia can more easily withstand the sanctions imposed by the United States and the EU.

New Energy Partnerships After EU Divestment

When the United States and EU started to use sanctions and other tools of state power against Russia to slow its progress in Ukraine, Russia responded by once again leveraging energy as a tool of its own. Initially, Russia reduced the amount of natural gas flowing through pipelines such as Nord Stream 1, which transferred natural gas from Russia to Germany. By September 2022, Russia had entirely shut off the Nord Stream 1 pipeline. Given the dependent relationship between the EU’s fossil fuel imports and Russia’s fossil fuel exports, these measures effectively drove Germany further away from Russian energy.

Before the start of the war, the EU was Russia’s largest oil export market, accounting for about 50% of its total exports. Likewise, many EU countries were heavily dependent upon Russian energy. For example, Lithuania imported 97.5% of its fossil fuels in 2021 from Russia, while 58.9% of The Netherlands’ and 50.1% of Slovakia’s imports were Russian. Other EU countries were slightly less reliant: Hungary imported 35.5% of its fossil fuels from Russia, Finland imported 32.0%, Germany imported 31.1%, and Poland imported 28.7%. Germany was already preparing to end its dependence on Russian energy after experiencing Russia’s use of energy as a tool in 2021, when Russia leveraged natural gas flows to speed up the certification of the Nord Stream 2 pipeline under construction at the time. The Ukraine War accelerated these efforts: the EU banned coal imports from Russia in April 2022, banned Russian seaborne crude oil and petroleum products in June 2022 (nearly 90% of its Russian imports), and introduced measures to reduce its natural gas demand (by addressing internal consumption) in September 2022. By February 2023, the EU only accounted for about 8% of Russian oil exports.

As the EU divests from Russian energy, several countries have stepped up to provide energy to the EU, filling the supply gap that could not be addressed by the EU’s renewable energy sources and changing the existing relationships between energy partners in the process. The United States now heavily supplies LNG to Germany, Finland, and other European countries. Other countries, such as Norway, Poland, Kazakhstan, and Algeria, also contribute to the EU’s energy supply. On the other hand, with the loss of the European energy market, Russia turned to the PRC and the East to make up for those lost exports and
keep energy profits (which support its military actions in Ukraine through the Sovereign Wealth Fund) flowing. Russia redirected many of its European exports to Turkey, India, and the PRC, increasing its shipments by over 1.2 million barrels per day between February and April 2022. By February 2023, Russia sent more than 70% of its exports to India and the PRC, including what was previously exported to Europe. In at least one instance, Russia also redirected LNG earmarked for Japan to India and the PRC to discourage Japan from imposing sanctions like the United States.

Russia’s crude oil exports now account for approximately 40% of India’s and 20% of China’s imports. Additionally, Russia increased its exports to Africa, Turkey, and the Middle East by up to 300,000 barrels per day as of February 2023.

The shift in import/export relationships seen after the EU divestment of Russian energy also illustrates how non-direct players can become tangled in issues between only two or three original parties. With a pending European energy crisis in the second half of 2022 and after months of high oil and gas prices at home, the United States aspired to do more than just increase its energy exports to the EU: it saw an opportunity to both stabilize the global energy market and restore relations with Saudi Arabia. Relations between the United States and Saudi Arabia had been tense for months in the aftermath of the Jamal Khashoggi murder in 2018 and after President Biden called Saudi Arabia a “pariah” while on the presidential campaign trail. During negotiations, the Biden Administration requested that Saudi Arabia increase crude oil production to help offset the loss of Russian oil sources globally, which would help to stabilize rising oil prices. In exchange for this increased oil production and an extension to a two-month-old ceasefire with Iran-backed Houthi fighters in Yemen, Biden agreed to a future visit to Saudi Arabia, an important diplomatic step in renewing relations between the two nations.

In this case, Saudi Arabia, while not a direct party to U.S. and EU sanctions, was pulled into the fray as both the Organization of the Petroleum Exporting Countries Plus (OPEC+) Chair and one of the only oil-producing countries with the ability to increase its production capacity. However, the United States’ intent to partner with Saudi Arabia backfired, as the Saudis reneged on the informal agreement it made with the Biden Administration, ultimately cutting oil production by 100,000 barrels per day rather than increasing it by the intended 200,000–400,000 barrels per day. This move angered the Biden Administration, which accused the Saudis of supporting Russia’s war in Ukraine and threatened punishment. Although no direct punishment against Saudi Arabia was ever enforced, the Administration did threaten to support the No Oil Producing and Exporting Cartels Act (NOPEC) under consideration in Congress, which would have lifted sovereign immunity and allowed for antitrust prosecution of OPEC members if passed. Accordingly, relations between the United States and Saudi Arabia further soured rather than improved.

Sanctions Do Benefit the United States and Allies

In some cases, the shuffling of relationships has created positive outcomes for U.S. and EU allies despite various influences in addition to sanctions. For example, Finland was already reducing its Russian energy dependence before the Ukraine War by heavily investing in renewable energy technology. Finland has had a well-diversified energy supply for years and only consumes about 6% of its overall energy from natural gas, making the transition from...
gas to renewables less arduous than in Germany (where household energy generally comes directly from Russian LNG sources) and other European states with greater dependence.\textsuperscript{51} However, over the last year and a half, since Russia halted its LNG imports in May 2022 as a result of sanctions, Finland has surged its efforts into renewable energy sources such as wind and nuclear to combat Russian supply issues, leading to decreased over-reliance on fossil fuels, decreased contribution to oil and gas resource depletion, and increased research and development in renewables.\textsuperscript{52}

Separately, Taiwan was heavily incentivized to increase its investment by the Biden Administration’s Chips and Science Act, passed in early 2023, and by its continually deteriorating relationship with the PRC, where U.S. support will be necessary in the future event of forced reunification. With Taiwan supporting the United States in sanctions against Russia and the associated impact to TSMC with the loss of Russian production contracts, TSMC has looked to the West (rather than to China in the East) for increasing its overseas production capacity, creating U.S. jobs and contributing to the U.S. chip supply. TSMC is now investing approximately $40 billion to build two production plants in Arizona, a significant increase from the $12 million deal to build one plant that was previously announced in 2020.\textsuperscript{53}

\textbf{Using Sanctions as an Economic Tool of Power}

Moving forward, as nations look to exercise sanctions and other economic tools of power, leaders should carefully consider the secondary and tertiary effects these tools could have domestically and on both allies and adversaries to ensure that any policy tradeoffs are well worth the price. In today’s interconnected world, the use of sanctions abroad affects the private sector at home more than it did before globalization. It also ensures that both nations and companies not directly involved in the sanctioning efforts may suddenly find themselves involved, and not always to their benefit. When the United States and EU imposed sanctions on Russia, many Western companies doing business in Russia lost millions of dollars (collectively billions) as they complied with the restrictions. While the Western private sector took a hit, on the other hand, sanctions did contribute to reduced European dependence on Russian energy as countries like Germany and Finland created a diversity of supply. Saudi Arabia was pulled in due to U.S. attempts to stabilize global oil markets and found its already tense relationship with the United States somehow worse. At the same time, Taiwan and Japan, two U.S. allies, were also affected by Russia’s reactions to U.S. sanctions. Taiwan decided to stop producing Russian chips, and Russia preemptively pulled LNG away from Japan. The United States and Taiwan benefit from Taiwan’s pivot to the West. The United States also finds its relationship with Japan strengthened as Japan responds with Russian sanctions. Ultimately, in an attempt to punish Russia for its actions in Ukraine, the United States also consequently helped many of its adversaries, including the PRC, India, and Turkey, to profit as all three countries took advantage of the redistribution of Russian market shares and formed new trading relationships with Russia.

Sanctions are a popular economic tool of power used against U.S. adversaries, but they take time to work—to achieve their intended goal and for connected events to unfold. As the United States, EU, and other allies impose additional sanctions on Russia in 2024 and beyond, trade relationships between nations will continue to transform. The United States must continuously
scrutinize the effects of sanctions, intended and unintended, and remain conscientious that sanctions are not the only tool of power available, nor should they be used individually. By combining sanctions with other tools, the United States will be able to reduce the unwanted effects of sanctions, like helping U.S. adversaries profit while reaching the goal faster.

About the Author: Ashley Kelso is a graduate student at Georgetown University’s Walsh School of Foreign Service and a member of the Security Studies program, concentrating in U.S. national security policy. Ashley works as a Proposal Manager at Booz Allen Hamilton and as a Teaching Assistant in the Government department at Harvard Extension School, supporting classes on intelligence and cybersecurity.

1 Berman, Noah and Anshu Siripurapu. “One Year of War in Ukraine: Are Sanctions Against Russia Making a
sanctions-against-russia-making-difference

Berman, “One Year of War in Ukraine.”

Berman, “One Year of War in Ukraine.”


economy-boycott-elites-benefit-ukraine-war.html

economy-boycott-elites-benefit-ukraine-war.html

Berman, “One Year of War in Ukraine.”

Berman, “One Year of War in Ukraine.”

Sonne, “How Putin Turned a Western Boycott Into a Bonanza.”

starbucks-operations.


Sonnenfeld, Jeffrey, and Steven Tian. “The Feckless 400’: These Companies Are Still Doing Business in Russia—
companies-are-still-doing-business-in-russia-funding-putins-war-sonnenfeld-tian/

Sonnenfeld, “The Feckless 400.”

“Companies Are Getting Out of Russia, Sometimes at a Cost.”


Humpert, “Western Sanctions Delay Opening of Arctic LNG 2 Project by One Year.”


critical-examination-pub-84181

Rumer, “Russia in the Arctic.”


Gilchrist, “How Surging Trade with China Is Boosting Russia’s War.”


Adomaitis, “How the EU Ban on Russian Oil Imports Affects Oil Flows.”

imports/which-countries-are-most-reliant-on-russian-energy.
33 “Which Countries Are Most Reliant on Russian Energy,” IEA.
34 Vakulenko, “Shutting Down Nord Stream Marks the Point of No Return for Russian Gas.”
36 Adomaitis, “How the EU Ban on Russian Oil Imports Affects Oil Flows.”
37 Adomaitis, “How the EU Ban on Russian Oil Imports Affects Oil Flows.”
38 Adomaitis, “How the EU Ban on Russian Oil Imports Affects Oil Flows.”
40 Adomaitis, “How the EU Ban on Russian Oil Imports Affects Oil Flows.”
41 Adomaitis, “How the EU Ban on Russian Oil Imports Affects Oil Flows.”
42 Adomaitis, “How the EU Ban on Russian Oil Imports Affects Oil Flows.”
43 Adomaitis, “How the EU Ban on Russian Oil Imports Affects Oil Flows.”
44 Matthews, “The End of Energy Free Trade.”
45 Matthews, “The End of Energy Free Trade.”
47 Said, “Saudi Oil Boost, Yemen Truce Set Stage for Biden Visit to Kingdom.”
48 Said, “Saudi Oil Boost, Yemen Truce Set Stage for Biden Visit to Kingdom.”
49 Said, “Saudi Oil Boost, Yemen Truce Set Stage for Biden Visit to Kingdom.”
52 Tynkkynen, “The Finnish ‘Ruxit.”

Keven Hernandez

In her book, *Overreach: How China Derailed Its Peaceful Rise*, Susan Shirk argues that the People’s Republic of China (PRC) overreached in its foreign policy by needlessly confronting other powers and adopting more strident controls over the economy and society to limit its dependence on foreigners. The PRC’s more aggressive posture, which prompted many countries to adopt defensive measures, may have been apparent to many China watchers since Xi Jinping took control of the PRC’s destiny. However, as she describes, this “old China hand” contends that the PRC’s strategic shift began during the Hu Jintao era.

The author, who served as the U.S. Deputy Assistant Secretary of State in the Bureau of East Asia and Pacific Affairs during the Clinton Administration, objectively catalogs how the Chinese Communist Party (CCP) has changed from when she first met Premier Zhou Enlai in 1971 to today—evolving into a much different China where the CCP seemingly wants to swallow the state whole. *Overreach* is a riveting read for anybody attempting to understand the PRC on a deeper level—how it thinks and how Chinese elite politics influence the foreign policy of a superpower looking to supplant the U.S.-led international order. Shirk’s insight into how the PRC’s “internal fragility” shapes its behavior abroad would benefit U.S. policymakers attempting to make informed decisions about the United States’ near-peer competitor.

Prying open the “black box” of the PRC’s political regime, Shirk outlines how domestic dysfunctionality allowed Xi to slide the PRC deeper into a dictatorship. With the regime already weakened under Hu Jintao, who was General Secretary of the CCP from 2002 to 2012, Xi was able to methodically reverse the constructs of “collective leadership” established by Deng Xiaoping. Intraparty politics soon devolved into corrupt “oligarchic rule” that prevented bureaucratic interest groups from checking and balancing one another, as would be appropriate in a collective leadership structure. Hu failed to continue the reforms instituted by Deng and allowed the “control coalition” to take over. Constituted by public security, propaganda, paramilitary, and military organizations, the coalition took advantage of the CCP’s fear of “social upheaval.” The author explains that this “dynamic drove foreign policies to become more truculent and domestic policies to become more autocratic.”

During Hu’s time in power, the conditions for PRC overreach emerged. Corruption spread. The private sector receded. Internet regulations multiplied. Additionally, aggression in the South China Sea increased. These developments allowed Xi to plot his moves accordingly and drape the PRC with his personalist regime. Such domestic transformations influenced how the PRC conducted foreign policy, one guided to shore up Xi’s popularity and position and enabled by the powerlessness of his rivals. In Shirk’s words:

> So what I’m debunking is the international relations scholars who believe that China’s rise, the reactions of the United States, the Graham Allison view, and the Thucydides trap view is incorrect and that, there’s a lot of human agency here and it reflects the nature
of Chinese politics in this period of collective leadership.\textsuperscript{11}

Shirk provides recommendations on how the United States and the PRC can manage relations. She believes the latter has an obvious path to tread: stop economic coercion, bring back the reform era, constructively engage Taiwan, and westernize the Belt and Road Initiative (BRI). Remodeling the BRI to fit international standards would woo more Western-oriented countries suspicious of Beijing’s coercive economic diplomacy. Avoiding indebting other countries with unsustainable projects and having transparent contracts would improve the PRC’s reputation when providing public goods abroad. Most obviously, Shirk suggests the closure of the “reeducation camps” in Xinjiang, the existence of which she believes has blighted the PRC’s image the most.\textsuperscript{12} That even Joseph Stalin released thousands of inmates from labor camps before he died–she writes–should be a warning to Xi. Her writing is bursting with practical solutions and marked with her trademark optimism, which could have the greatest impact if Chinese audiences could read her tome. But they cannot.

Xi has the human agency to revert the PRC back to those olden days of reform and abundance. But Shirk surely knows that he will not go down such a path. The longer Xi remains in power, the more difficult it will be for the CCP to change course.

For the United States, Shirk argues that it should be wise enough to restrain its worst impulses when combating the PRC challenge. She is short on details about how this restraint could be realized. She must be. Because U.S. self-discipline, like on the PRC side, is a human affair. The type of leadership in power on both sides dictates how the bilateral relationship can change.

Nonetheless, Shirk recommends that the United States work with allies, not give up on diplomacy with the CCP, be forceful but realistic on human rights, avoid an ideological war with Beijing, engage the Chinese people, and reject primacy that could feed into the PRC’s overreach abroad. The most illuminating recommendation is for the United States to embody its democratic principles more fully. The specter of what is occurring on the American right will be obvious for many readers; the Republican Party appears content and sometimes partial to America’s democratic backsliding. But the Democratic Party is put on notice as well. President Joe Biden has continued many of former President Donald Trump’s policies, such as trying to limit the abundant research collaboration between American and Chinese scientists to snuff out China’s technological ambitions. Shirk argues such restrictions are warranted for military technologies but not for the common scientific exchanges that have occurred for decades. “The United States should build on its assets instead of turning itself inside out trying to become more like China,” she cautions.\textsuperscript{13}

If domestic dynamics have locked the PRC into a more aggressive foreign policy on the American side, so has the willingness of Republicans in Congress to accept the high price of not mending ties with the PRC. The PRC and the United States have a history of cooperation in science, trade, technology, education, and personal connections between their citizens. As these unravel, the more trenchant these partings seem to the outside world, to Americans and the Chinese. But again, leaders on both sides always have the opportunity to repair the relationship, as unlikely as that may seem on both sides. Overreach implores us to engender more thinking along these lines to understand the ‘people’ factor.
But does Shirk’s appraisal of the CCP give U.S. policymakers a springboard to think about how to deter or prevent war? The answer is yes. Maybe. The question for American defense strategists should be how to use Shirk’s understanding of the CCP to make better decisions regarding the PRC. Indeed, the CCP will shape the actions of the People’s Liberation Army (PLA) in the run-up to a possible confrontation or conflict. Thus, how can this knowledge about how the PRC thinks filter down to the larger U.S. defense enterprise at the operational or tactical level? Does the commanding officer of a U.S. Navy destroyer performing freedom of navigation operations in the South China Sea know about the kind of political system his PLA counterpart is operating under? Does the pilot flying Boeing P-8 Poseidon know? What happens during a collision that results from aggressive PRC behavior? The U.S. reaction will need to be informed by the logic Shirk spells out. It should be fundamental for American military personnel to recognize the DNA of insecurity coursing through the CCP. In doing so, the United States can be more measured in its response and avoid sparking a confrontation that could lead to war.

The only criticism leveled at the book is in how the author thinks the United States should deploy its military in the Indo-Pacific. Shirk, by the logic of her thesis, implicitly provides backing for a forceful U.S. deterrent model for military challenges, given Beijing’s truculence. But she argues that an arms race in the Indo-Pacific could quickly devolve into an unstable environment that will prompt Xi to worry about the survival of his regime. This contradiction (perhaps less by her and more by the logic of international relations) captures the challenge well: how can the U.S. deter just enough without overreaching and forestalling cooperation in other areas?

Shirk puts the reader in a position to understand how a potent adversary thinks and why it behaves the way it does. She eloquently illuminates the roads that the United States and the PRC can take. Both sides should choose tactfully.

**About the Author:** Keven Hernandez is a graduate of the Security Studies Program at Georgetown University. He was the Georgetown Security Studies Review Associate Editor for the Indo-Pacific from September 2021 to August 2023. He also served as the Editor-at-Large for the Georgetown Journal of International Affairs. Keven most recently was an FMSO-GCKN research fellow for the U.S. Army, TRADOC. He delights in taking in the D.C. music scene one concert at a time. Email: kh1167@georgetown.edu
4 Shirk, *Overreach*, 289.
10 Shirk, *Overreach*, 12.
A Tale of Two Koreas: How the Diverging Korean Language Will Challenge Future Unification

Olivia Schieber

After more than 70 years of division, North and South Korea could not be more different. Not only are their political and economic systems completely opposed to one another, but even the language spoken in North and South Korea has diverged significantly. This paper addresses how this divided Korean language may frustrate future attempts to successfully unify the Korean Peninsula. While the North and South Korean languages are not so divergent that they are completely mutually unintelligible, their vast differences reflect two fundamentally different societies. When it comes to unification, this not only signals profound logistical problems—such as standardizing vocabulary, spelling, and grammar—but also considerable differences in national identity, philosophy, and values. While unification may not be feasible in the near term, it is imperative that policymakers in the United States and South Korea alike understand today how complex such an outcome may be in the future. This paper explains the extent to which the Korean language has diverged; identifies the key factors that have contributed to this divergence; illuminates the relationship between language and Korean identity; and analyzes the political, economic, and societal implications of a Unified Korea. Finally, this paper offers an adaptable policy framework that addresses both the logistical-linguistic and ideological elements of language unification.

Introduction

On January 25, 2018, a bus carrying several female North Korean hockey players arrived in Jincheon, South Korea. There, they met their South Korean counterparts of the joint North-South hockey team for the 2018 Winter Olympics in Pyeongchang. The meeting of the two teams represented a major public relations win for then-South Korean President Moon Jae-in, who supported a pro-unification agenda.¹ To break the initial tension between the teams, journalists instructed players to pose for photographs and shout, “We are one!”² And while the moment succeeded in creating a more relaxed and open environment for the players, the cheer they exclaimed during the photoshoot could not have been further from the truth.

The joint North-South team encountered significant problems during their practice sessions—not because of a difference in hockey skills, but because the two teams had substantially different means of communication. The South Korean players relied heavily on English loan words for basic hockey phrases, such as the words “pass” or “defense.” In contrast, the North Korean players exclusively used native Korean words, like yeollak (연락) for “communication” or bang-ee soo (방어수) for “defense player.”³ As a stop-gap measure, the women’s hockey team created a dictionary to translate words from English to South Korean and then eventually to North Korean.⁴ A three-way translation was likely tedious, but the measure mainly proved sufficient given the team’s time constraints.

While the media flurry over the joint North-South hockey team came and went, the venture is worth remembering, if at least for one reason: language. After more than 70 years of division, the language spoken in North and South Korea has diverged significantly. This paper explores how
linguistic differences between North and South Korea may challenge the unification process. While Korean unification is yet a far-away possibility, it is important to begin the process of evaluating the potential obstacles that could arise in uniting two countries that have been separated for the better part of a century.

As the literature review will demonstrate, there is no shortage of analysis on Korean unification. However, social issues rarely come to the forefront of these discussions, and, glaringly, language hardly factors into this literature at all. At the same time, there is a growing body of evidence that demonstrates how the language barrier negatively affects the ability of North Korean defectors to assimilate into South Korean society and form meaningful personal and professional connections. A South Korean committee tasked with compiling a North-South Korean dictionary (known as Gyeoremal-keunsajeon / 겨레말큰사전) found that there is a 38% difference in regular vocabulary and a 66% difference in specialized vocabulary between the two countries. It should be clear that language will prove a major point of contention when it comes to merging the two halves of the peninsula.

While unification may not be feasible in the near term, policymakers in the United States and South Korea alike must understand today how complex such an outcome may be in the future. The differences between the North and South Korean languages not only signal profound logistical problems—such as standardizing vocabulary, spelling, and grammar—but they also portend considerable differences in national identity, philosophy, and values. A unified dictionary is a welcome start to thinking about this problem, but ultimately, vocabulary differences reflect only one part of the challenge. No modern framework exists for policymakers to prepare for language unification that addresses both the logistical-linguistic elements and the ideological and values component of this problem. This paper seeks to rectify this gap by providing policymakers with a flexible and practical policy framework that can inform policy decisions now and in the future.

Methodology, Scope, and Definitions

This paper will provide original research to establish and synthesize three levels of analysis: a deep dive of how and to what degree the North and South Korean languages have diverged; how language is a core factor that shapes Korean identity; and how these divergent identities, as reflected and reinforced in the language, threaten future unification. The focus of this paper will be on the potential challenges North and South Korea might encounter due to language differences. While readers may be able to extrapolate from this analysis how language will impact the success of unification, defining “unification success” is beyond the scope of this paper. This is because unification will be a highly complex and multifaceted process. Language will impact unification in unexpected ways outside mere linguistics. Therefore, some elements of unification may be more successful than others, but the overall success of unification requires far broader analysis than this paper will provide. Instead, this paper seeks to challenge policymakers and academics alike to view what is traditionally seen as a security issue in a way that incorporates linguistic, social, and cultural elements.

This study assumes a unified Korea in which the Republic of Korea is the victor (henceforth “Unified Korea”) in order to provide a theoretical baseline for empirical analysis. Of course, other alternative futures are possible—but a unified Korea under the North Korean flag and whose government
has complete control over the population is hardly a compelling thought experiment for a study on linguistic challenges. Moreover, this paper will not explore the various scenarios that precede unification, nor will it examine the knock-on effects of a theoretical North Korean collapse (such as nuclear proliferation or refugee crises). Doing so necessitates a certain suspension of disbelief, but it is essential for empirical analysis. Moreover, since unification refers to a theoretical event in the future there are few relevant case studies for comparison. That said, this paper will apply lessons learned from German unification as well as from North Korean defectors residing in South Korea.

While occasionally drawing upon examples from German unification, this paper is ultimately concerned with the Korean experience. To be sure, there are numerous examples of both countries that contend with multiple official languages or dialects in addition to languages that are spoken in multiple countries. While these case studies might have value when considering language and Korean unification, this paper is focused on the uniqueness of the Korean language as it relates to Korean identity. Comparisons to places where countries have multiple languages or dialects (like China or India) or cases where a language is spoken differently in different places (like Spanish or Arabic) are areas for future research but out of scope for this study.

Finally, for the sake of simplicity, this paper will often refer to the North and South Korean variants of the language either as such or as “North Korean” and “South Korean.” Similarly, North and South Korean will correspond to the Pyongyang and Seoul dialects respectively, since these are considered the standard dialects in each country. Naturally, North and South Korea contain several internal dialects, which this paper will address periodically.

**Literature Review**

How a divided Korean language will impact unification is a complex question that must be addressed through the synthesis of multiple disciplines. This challenge draws on the studies of unification, historical language policies of North and South Korea, and the relationship between language and national identity. The role that language may play during unification with respect to Korean identity is a relatively unexplored subject, particularly in English. Therefore, the following review will evaluate each of these components separately and demonstrate how this paper will join these elements in a comprehensive analysis that fills the existing academic gaps.

Korean unification is by no means understudied. Much of this literature is understandably rooted in security issues, such as how to handle North Korea’s nuclear weapons or unify two economies at the extremes of development. These are undoubtedly important elements of unification that should not be dismissed, but on their own, they do not paint a full picture of how social and cultural factors will influence the degree to which North and South Korean can successfully merge. That said, some scholars have attempted to bridge the culture-security gap. For example, Michael Edmonston, a military strategist from the Air Academy, examines how ideology has influenced military culture in North and South Korea in, “The Potential of Korean Unification and a Unified Korean Armed Forces: A Cultural Interpretation.” Edmonston’s contribution is a recent and welcome addition to unification ligature, particularly that which focuses on the military. While Edmonston’s focus is ultimately a security concern—specifically how culture will influence a future joint
North-South force—he acknowledges how the “fixed cultural gap” will make unification challenging.\(^8\)

Paul Chamberlin’s, “Cultural Dimensions of Korean Reunification: Building a Unified Society,” is one of few peer-reviewed pieces that explicitly weaves cultural elements into common points of discussion in unification, such as economics, education, demographics, and the future of political institutions.\(^9\) Chamberlin’s piece does not merely outline potential challenges but also offers several policy recommendations to foster a democratic, free-market, unified Korea. Chamberlin does not acknowledge linguistic gaps (and at the time of publication in 2004, these gaps may not have been as severe then as they are today). Furthermore, this piece was published at the height of North-South cooperation during the “Sunshine Policy” era of South Korean President Kim Dae-jung, who promoted engagement with North Korea. As such, many of Chamberlin’s policy recommendations reflect a less bombastic time in North-South relations and are potentially less applicable today.

Few real-world events offer lessons learned for Korean unification. Consequently, the reunification of East and West Germany is a common point of comparison in academic literature. In “The Challenges of Reunification: Why South Korea Cannot Follow Germany’s Strategy,” Stefan Niederhafner highlights key challenges of German unification and explains why the German model—unify first, reconcile differences later—will not work for Korea.\(^10\) He makes a key observation in pointing out that the sense of common identity between North and South Korea is likely much weaker today than it was between East and West Germany. While Niederhafner does not raise the unique linguistic challenges of the Korean case (nor does he address language issues that occurred in the German case), the question of identity is essential to this project. This paper will build on the notion of shared identity as an integral part of successful unification and argue that language is foundational to this identity.

Finally, the 2018 Winter Olympics in Pyeongchang, South Korea, ushered a renewed sense of optimism for North-South relations, with unification issues coming to the forefront of discussion once more. Perhaps for the first time on the world stage, the linguistic differences between the two Koreas became all too clear with the joint North-South women’s hockey team.\(^11\) Inspired by this event, Leena Hamad offers one of few academic English-language assessments of the linguistic challenges of unification. Her piece, “A Language Split by the Border: What the Division of the Korean Language Means for Reunification,” highlights some examples, such as words that have one definition in the South but carry a completely different meaning in the North.\(^12\) However, this is the extent of Hamad’s linguistic analysis. She provides only a few simplistic illustrations of potential linguistic challenges and offers no recommendations for how to treat these issues. Furthermore, because Hamad offers only a brief, surface-level analysis of the language divergence, she is unable to draw a meaningful connection between language and national identity.

The above contributions are reflective of unification studies that broach the subject from beyond a pure security perspective. Clearly, language does not factor into many of these discussions. It is therefore essential to examine the work of Korean linguists, who have mapped the divergences between the two Koreas. Afterall, it is impossible to provide adequate policy solutions without first understanding the magnitude of the challenge facing the peninsula.
Jaehoon Yeon’s “Standard Language’ and ‘Cultured Language’ explores the language policy decisions that North and South Korea employed through the second half of the 20th century. Yeon methodically outlines three eras in North Korean language policy and illustrates how the regime “nativized” the language. He also provides (Romanized) examples of differences in North Korean syntax, vocabulary, phonology, and spelling. However, Yeon’s work is a purely historical linguistic text and as such does not attempt to extrapolate any relationship between language and a nation’s collective psyche. Furthermore, while Yeon offers a useful primer on North Korean linguistic history, he ultimately concludes that there are “no real difficulties” in communication between the two Koreas despite obvious divergences. It is important to note that Yeon originally published this piece in 2003, and all his sources are from the 1990s. Perhaps at the time of writing, the linguistic gaps were not as great as they are now. It is difficult to imagine Yeon would draw the same conclusion today as he did during the early days of the Internet.

Jae Jung Song, another Korean linguist, provides additional historical context for the language divergences. His chapter, “Language Policies in North and South Korea,” traces Korean language policy to the era of Japanese colonialism, which he cites as the catalyst for both Koreas to elevate the “native Korean” elements of the language (as opposed to words borrowed from Chinese or Japanese). He draws a key distinction between how language developed in the North and South. South Korea largely lacked formal language policies until the 1990s and allowed the language to develop in flux with globalization. In contrast, North Korean language policy fell under the strict purview of Kim Il Sung, who saw language as a tool to shape the minds of his citizens. This is an essential point for this project, as it underscores why language is not simply a matter of vocabulary and syntax but rather a window into how people think. Yeon and Song do not consider the implications of their works in the context of unification, nor do they weigh the broader long-term sociological implications of these language policies.

The link between language and identity is the final major component to the question of unification. Jiyoon Kim offers compelling analysis for the importance of language and national identity in South Korea in, “National Identity and Attitudes Toward North Korean Defectors.” While Kim does not focus on language differences specifically, she provides evidence that when it comes to evaluating Korean identity, South Koreans increasingly prioritize civic factors, including speaking and reading Korean, over ethnic factors. This trend is visible in works that examine North Korean defectors living in South Korea. In, “North Korean Defectors as Cultural Other in South Korea,” Kyung Hyo Chun emphasizes how language differences have contributed to the stigmatization of defectors, who go to great lengths to conceal their accents. Hee Jin Kim et al. quantify this “otherness” in their study, “Social Distance towards the North Korean Refugees in South Korean Society.” They found that South Koreans who reported stronger feelings of Korean identity felt more socially distant from defectors. These three studies demonstrate real-life manifestations of the “fixed cultural gap” to which Edmonston alluded in his joint forces piece. However, these studies do not explain why the North Korean language specifically contributes to the “otherness” of defectors. For example, there are multiple dialects in South Korea, yet only the North Korean accent triggers this “otherness” response.

Each of the above works contributes meaningfully to discussions on unification,
language, and national identity. Yet none combines all three elements to comprehensively address how language both shapes and mirrors national identity and subsequently how this divided identity could threaten unification. While others have broadly argued that linguistic differences will be an obstacle in unification, the prescribed solution is almost always rooted in the reconciliation of spelling, syntax, and vocabulary. The challenge is about more than the use of English loan words or grammatical quibbles. The true test will be reconciling the identities of two groups of people who see themselves and their places in the world in entirely different ways. This is because language is more than a means of communication; it is a reflection of culture, society, history, and identity. To navigate a path through this intellectual quagmire, this paper will build upon the existing literature. It will offer a thorough analysis of the linguistic differences between North and South Korean, illuminate the relationship between language and Korean identity, and provide policy recommendations for mitigating potential problems related to language during unification.

North Korean Language Deep Dive

Understanding the extent to which, and the reasons why, the North and South Korean variants of the language have diverged is paramount. This section will demonstrate that the Korean language has diverged due to three primary factors. First, and most obvious, is that the Korean Peninsula has been divided for over half a century. While South Korea underwent extraordinary economic, social, and political transformation during these decades, North Korea and its language remained essentially frozen in time. South Korea incorporated new technologies, cultural concepts, and a host of foreign loan words, reflecting its globalized nature. In sharp contrast, the Democratic People’s Republic of Korea (DPRK) categorically rejected all the above, which the Kim regime has justified under North Korea’s homegrown Marxist philosophy known as Juche or “self-reliance.”

Second, and related to the first point, North Korea has a multitude of dialects within the country, some of which vary substantially from the standard Pyongyang dialect. While this section will primarily address differences between the standard Seoul and Pyongyang dialects, it is important to keep in mind that in a unification scenario, Unified Korea will not only need to contend with dialect differences between the Seoul and Pyongyang dialects, but also the intricacies found in local dialects. Another complication is the fact that because North Korea does not allow for internal freedom of movement, few North Koreans gain exposure to dialects outside their own or standard Pyongyang speech.

Third, a less obvious reason for the diverging Korean language results from the set of directives set by North Korea’s first dictator, Kim Il Sung. Kim understood that to harness the minds of the North Korean populace, he would need to have control over the media his citizens consumed. This line of thinking is a product of the Marxist philosophy toward language and the building of a socialist society. According to this philosophy, language and thought are one and the same. To do so, he needed a means of effectively distributing propaganda as well as a population that could easily consume those messages. Consequently, Kim enacted a series of language policies aimed at both eradicating illiteracy and shaping the language to suit the ideology of the regime. The legacy of these language policies lives on even today. This context is necessary to appreciate the other factors that have contributed to the language divergence.
Kim’s top-down approach to language and his objective of using language to control North Korean minds underscores that language is more than a set of grammar and spelling rules—it is the vehicle for how we think and communicate. This section will offer this historical context as the precursor for a more in-depth discussion of how the Korean language differs in the North and South today.

A Brief History of North Korean Language Policy

Scholars of North Korean linguistics typically break the DPRK’s language policy into two or three distinct periods, beginning in 1945 and ending in the mid-1960s. Among Kim Il Sung’s early language priorities was a focus on eradicating illiteracy and the beginnings of language purification. The Korean language consists of both native Korean words and Sino-Korean words (i.e., those that correspond to Chinese characters). The creation of the Korean alphabet—known as Hangul in the South and Joseongul in the North—can be traced to the 15th century and is attributed to King Sejong the Great. Hangul offers a straightforward, phonetic writing system designed to complement the previous writing system of Chinese characters (known as hanja or hanmnun) and improve literacy. One of Kim Il Sung’s first moves was to eliminate all hanja. Reverting to the exclusive use of Joseongul would make it easier for the populace to digest written propaganda materials and embrace the ideology of the Kim regime. Additionally, this policy represented the beginning of the “nativization” of the Korean language and the replacement of Sino-Korean words with native Korean ones.

With the rise of the Juche philosophy of “self-reliance”—Kim’s answer for the rapidly declining relationships between North Korea and the USSR and China—so too came new shifts in language policy. Linguistic scholars point to conversations Kim had with linguists in 1964 and 1966, during which Kim shared his vision for the Korean language. In the first speech on January 3, 1964, Kim brings language purification and top-down state management of the language to the forefront. Kim also articulated the need for the government to “manage” the use of hanja, foreign loan words, and vocabulary. Kim argued for new rules regarding the spacing of words, stating that Korean sentences looked like “an unbroken string of syllabics,” making them resemble English and Chinese writing systems. He emphasized the need to hold off on making more drastic changes to the Korean script (Joseongul), arguing that too much divergence between North and South Korea would prevent cultural and technological development. It is interesting to note that as early as 1964, Kim foresaw the difficulties a divided Korean language could cause during unification. Furthermore, this speech reflects Kim’s belief that language could be used to shape North Korea’s sense of identity. He opined, “Many words of Chinese origin that our ancestors used and then discarded are now being revived...With ‘the development of science and technology and the progress of society,’ our vocabulary must also be expanded. We must create a lot of new words. But we should make it a point to form these new words from our own root words.” Clearly, Kim recognized that the Korean language would need to evolve over time but felt that this evolution should take place on his own terms.

In 1966, Kim enshrined the concept of munhwaeo (“cultured language”), the final politicization of the North Korean language. In “On Correctly Preserving the National Characteristics of the Korean Language,” Kim painted the South Korean variant of the language as the antithesis for the
development of a socialist Korea and called for a distinct North Korean language under a new name. Kim called language the “weapon for the realization of revolution,” and called forth further efforts to purify the language. He also reversed course on the ban on Chinese characters—at least where education was concerned. North Korean students, like their Southern counterparts, learn hanja in school, even though Chinese characters are not used anywhere else in the DPRK. With the goal of unification in mind—Kim rationalized that if South Koreans were using hanja, North Korean students would need to learn them as well.

North Korea’s top-down management of language policy reflects the characteristics of an authoritarian regime. Kim saw language as integral to building a socialist, Juche-compliant Korea—one that would be morally and ideologically superior to the corrupt and bourgeois South. In contrast, South Korea’s language evolution mirrored its own journey toward globalization. True, for much of the late 20th century, South Korea was no democracy. From the dictatorships of Syngman Rhee, Park Chung-hee, and Chun Doo-hwan, South Korea is no stranger to authoritarianism. However, unlike the deification of the Kim family in the North, South Korea’s dictators reflected a different character. Though certainly no beacon of human rights, Park Chung-hee’s fixation on growing South Korea’s economy in the 1960s and ‘70s is what ultimately catapulted the country to “Asian Tiger” status. Park embraced an export-driven strategy that would cement South Korea as part of the globalized economy. While not a completely free-market economy (Park favored both private entrepreneurship and state-sponsored industry), the orientation toward a more globalized society helped lay the groundwork for a country that would one day embrace cultural imports and exports alike. It is no wonder then, that language policy in South Korea developed in a much more natural and flexible manner. Not only do the language policies of North and South Korea reflect their given political development, but so do the languages themselves.

Examining the Differences between North and South Korean

North and South Korean vary both in terms of pronunciation, spelling, and syntax as well as vocabulary. Even in cases where North and South Korea use the same word or phrase, the spelling or pronunciation may differ. For example, many words that begin with vowels in South Korea begin with consonants like “n” (ㄴ) or “r” (ㄹ) in the North. Additionally, some words that begin with an “n” (ㄴ) sound in South Korea begin with a soft “r” or “l” sound in the North (ㄹ). The North Korean variant tends to sound more militaristic than its South Korean counterpart, particularly in cases where consonants are duplicated to sound more aggressive. Table 1 in the appendix shows some examples of how North and South Koreans will pronounce the same words differently.

The difference in spelling and pronunciation is not particularly severe, and it is reasonable to assume that North and South Koreans would be able discern the meanings of these words. Both South and North Korea contend with dialectic differences in-country that do not interfere with language standardization. For example, a person from Jeolla (a southern province in South Korea) could still understand someone from Seoul, as with someone from Hamgyeong (a northern province in North Korea) and Pyongyang. These factors could pose logistical challenges during unification, but ultimately, they are smaller hurdles when compared to the vast vocabulary differences.
between the two Koreas. This is especially true if a unified Korea is willing to accept, on principle, multiple pronunciations of the same word. Not only did Kim Il Sung succeed in ridding the country of many of its Sino-Korean and foreign loanwords, but there are also native Korean words that have taken on new, more spirited meanings in the North. That a word meaning “young lady” (agassi / 아가씨) in South Korea becomes “slave of feudal society” in the North, is a compelling example of this. This presents not only a logistical problem, but an ideological one as well. Table 2 in the appendix shows some of these vocabulary differences. While this paper does not attempt to replicate a dictionary, it is important to note that these vocabulary differences are substantial. The “Implications for Unification” section will explore how some of these vocabulary differences will challenge future unification.

**Linking Language and Identity**

* Civic vs. Ethnic Factors

For all the talk of the “one Korea” or “one-blood nationalism” (as in, one ethnicity, one language) sentiment that is popular among unification proponents, the linguistic divergence between the two countries should demonstrate that the real picture is much more complex. What does it mean to be Korean, and to what degree is language a precondition for “Koreanness”? Increasingly, civic factors play a significant role in shaping national identity in South Korea. Studies that explore South Korean nationalism and identity highlight that speaking the Korean language is an increasingly important factor in establishing Korean identity. For example, a 2017 study of national identity in South Korea showed that roughly 87% of South Koreans surveyed indicated that language was either “fairly” or “very important” to national identity. Ethnic factors—including birth, lifelong residency, and ancestry—received lower scores at 82%, 71%, and 71%, respectively. While these results might seem surprising for an ethnically homogeneous country with jus sanguinis citizenship, they are not anomalous: Jiyoon Kim’s aforementioned 2014 publication documented similar results. Some scholars have pointed to this shift toward favoring “achievable factors” over “inherent characteristics” as a sign that South Koreans are becoming more open to foreigners. However, if there are significant differences between North and South Korean, will North Koreans be treated in the same category as foreigners in Unified Korea? Furthermore, if national identity and language are intrinsically linked, which version of Korean is considered the “correct” precondition for claiming Korean identity? For the purposes of this study, we assume a Korea that is unified under the South Korean flag, which would summarily dictate that the South’s version of Korean becomes the standard. It is ironic then, that the “purer” version of Korean—the North’s—is relegated to secondary status. However, as unification scholar Jeon Yeong-sun argues, a nation’s prestige and the influential power of its language are closely related. Even without the assumption of a Unified Korea under the South Korean flag, the South Korean language is the version taught to students around the globe. Further, South Korea’s position as a cultural powerhouse helps cement its version of the language as the standard for unification as well.

Naturally, there are no opinion polls in North Korea, much less those that quiz the general population on their personal definitions of national identity. It is important to consider how North Koreans will value language as a source of their own
identity during unification, particularly if they subscribe to doctrine that suggests their version of the language is the “correct” one. While it is likely that North Koreans emphasize ethnic components of identity more so than South Koreans—the purity of the Korean race is a core part of DPRK propaganda—it would be a mistake to pursue a unification agenda without also considering how North Koreans will receive these policies. Fortunately, North Korean defectors living in South Korea may offer valuable insight into the role language already plays in their lives.

Case Study: North Korean Defectors in South Korea

Using North Korean defectors to better understand North Korean society and culture is by no means a novel approach. While existing studies frequently highlight language as a pain point for North Korean defectors, they do not connect the difference in language to a difference in Korean identity. Unified Korea will not only necessitate the blending of two national identities—it will require a new, cohesive one. That said, North Korean defector testimony can both offer a glimpse into language problems during unification and indicate potential solutions.

North Korean defectors encounter discrimination in South Korea, and many go to great lengths to conceal their accents, less they “out” themselves to their South Korean peers. These fears are not unsubstantiated. A 2015 study on South Korean attitudes toward North Korean defectors showed that while 61% of South Koreans nominally accepted defectors as ROK citizens, few South Koreans would consider forming close, personal relationships with defectors. For example, only 7.5% said they would accept a defector as a significant other, and fewer than 8% said they would marry or allow their children to marry a defector. And while 58% of respondents agreed they would accept a refugee coworker, only 14% would accept them as a business partner. Since North and South Korea are ethnically homogenous, only social, cultural, and language factors can distinguish the two populations. Hence, language plays a significant role in identifying and shaping South Korean perceptions toward defectors. In a 2020 government survey, nearly 62% of defectors cited their “way of communication” as a source of prejudice. While it is impossible to predict how unification will unfold, we can assume that problems seen in a micro-population will only exacerbate when escalated to a grander scale.

A 2018 study that polled North Korean defectors on economic integration found that 45% of respondents felt they were “highly discriminated” against when it came to finding employment and cited linguistic problems—particularly unfamiliarity with English loanwords—and lack of understanding of Korean social norms as key contributing factors. To rectify this, in 2021, South Korea’s Ministry of Unification published a glossary of industry-specific vocabulary to help defectors cross the professional language barrier. While older defectors hailed the publication, citing difficulties acclimating to South Korean workplaces because of language issues, younger defectors found the glossary inadequate for those aspiring toward more technical (and high-paying) professions. While using defectors as a microcosm is a useful strategy, this situation highlights that North Koreans are not a monolith. Linguistic challenges for North Koreans will vary due to age, gender, education, occupation, and familiarity with non-North Korean media. It will be important for policymakers to develop a nuanced approach for managing language issues, as
the South Korean industry-specific glossary demonstrates.

The ability of North Koreans to succeed economically in a free-market Unified Korea is not only a matter of providing translations of loanwords—and the challenges may be especially steep for those in STEM fields. As one defector remarked, he knew conceptually what semiconductors and transistors were, but he knew them as *jipjok soja* (집적소자) and *samguk soja* (삼극소자), instead of their English names, which are used in South Korea. The defector needed to re-learn his entire profession from the ground up in addition to the social terminologies to which he was unaccustomed. Given the alleged 66% difference in specialized vocabulary, it will likely mean that North and South Koreans will be unable to work together in essential fields, like engineering, medicine, and computer science. After all, the stakes in these fields will be significantly higher than those for a joint Olympic hockey team. North Koreans required to essentially re-learn the fundamental knowledge required for specialized industries will likely prove a severe handicap in a unified economy.

While there is some precedence of North-South operations through the Kaesong Industrial Complex (KIC)—a currently suspended joint venture that allowed South Korean companies to use North Korean labor—the KIC is not a true test of language in technical fields. The KIC was designed to provide North Korean laborers with wages to help reform the North Korean economy—it was never an intellectual collaboration.

Furthermore, language may fuel the cycle of discrimination further. For example, if South Koreans in Unified Korea are inherently prejudiced against those with North Korean dialects and mannerisms, it will be harder for them to integrate with South Koreans to learn new words and social behaviors. An inability to freely communicate in their given fields may further fuel negative stereotypes about North Koreans, thus allowing the prejudice among South Koreans to deepen.

**Implications For Unification**

Language is not merely a logistical issue resulting from spelling and syntax but an ideological one. During unification, North Korea will undergo massive societal upheavals like those seen during the unification of East and West Germany. Ideological concepts taken for granted in the figurative West—such as a free and open society with democracy and free-market economics—will be virtually unknown to most North Koreans. Unification scholar Jeon Young-sun refers to language as a “collective common” and remarks that when people speak the same language, they share words and emotions when doing so. This sentiment underscores the idea that language differences signal not only language problems but potentially other issues as well. This section will examine how language may impact three core areas: political participation, economic integration, and social cohesion.

**Political Participation**

A comparative study of historical German and theoretical Korean unification highlighted the difficulties of nurturing political participation in a former authoritarian, communist country. Unlike in South Korea, political participation in the North is restricted to those of the highest *songbun*—the DPRK’s de facto caste system—and therefore, representative leadership and free-and-fair elections and the like are foreign concepts. Stefan Niederhafner reasons that despite having one-third of Unified Korea’s electorate, it would be extremely doubtful that North Koreans would wield significant political
Not only would North Korea’s old guard prove an extremely dubious political option given their implication in the regime’s vast human rights abuses, but the lack of a viable alternative to represent Northern interests is problematic as well. While North Korean defectors are deeply active in the advocacy space, it was not until the 2020 National Assembly elections that a former North Korean defector gained electoral representation in the ROK. Former North Korean diplomat Thae Yong-ho ran for election because of his frustrations with then-President Moon Jae-in’s North Korea policies and his desire to show North Koreans abroad “how democracy works.” Thae’s remarks indicate that most North Koreans are unfamiliar with electoral politics. A lack of vocabulary surrounding basic democratic tenets is only part of the problem, however.

The North Korean language—as orchestrated by Kim Il Sung—centers around the Kim family regime. North Korea’s three dictators—Kim Il Sung, Kim Jong Il, and Kim Jong Un—serve not only as the singular political authorities, but as surrogate parental figures as well. The “parentification” of Kim Il Sung in particular has led to the co-option of the Korean word for “parents” and “father” (oboi / 어버이 and aboji / 아버지) to apply to Kim Il Sung instead. Naturally, the parentification of a political leader necessitates the infantilization of the population over which the leader presides—a common narrative in North Korean propaganda. This distortion of both politics and familiar relationships will likely make the transition to democracy that much more complicated. After all, how does one fully grasp the concept of freely communicating one’s own political preferences—and act on these preferences through voting or running for office—if politics has always been framed in a top-down parent-offspring relationship?

In many ways, North Korean life is inherently political, as the state dictates virtually every facet of North Koreans’ existence. Unsurprisingly, the state has woven propaganda into the North Korean language. One case of this is in the way the North Korean state portrays non-North Korean entities in media. As North Korea researcher Fyodor Tertitskiy writes, North Korean news is either “very hysteric[al] or very neutr[al].” North Korea’s enemies—the United States, Japan, and North Korea hardliners in the ROK—receive the worst treatment. Former South Korean presidents Park Geun-hye and Lee Myung-bak earned the descriptors “ugly female bat disgrace” and “rat-like,” while the South Korean government is simply a “puppet state” of the United States. While it is impossible to know to what degree North Koreans internalize these narratives, it portends poorly for Unified Korea if a third of its populace lacks faith in the political system out of an ingrained belief that its government is illegitimate.

Economic Integration

Workplace life is a source of stress and discrimination for North Korean defectors and will likely continue to be a pain point in the future. Unfortunately, as defectors have alluded, the issue is too complex for a simple glossary to fix. The vast differences between a free-market capitalist system and a state-run communist system pose a major issue for unification. As early as the 1990s, academics theorized there would be a long learning curve for North Koreans when it came to understanding the inner workings of a capitalist market economy (particularly the stock market).

To this point, North Korea lacks the vocabulary for many words associated with
private enterprise in the South, such as equivalents for CEO, *si-i-oh* (시어오) or *dae-pyo (i-sa)* (대표이사). The language difference may also signal challenges for cultural gaps in the professional realm. While some vocabulary differences—such as different words for supplies like “stapler”—will be annoying, if surmountable nuisances, others may indicate a difference in how workers orient and conduct themselves in the workplace. Everything from the language used when giving a workplace presentation to how one refers to deadlines or describes the difficulty of a particular task differs between the North and the South. Furthermore, the ubiquity of English loanwords in the workplace reflects South Korea’s integral role in the globalized economy. North Koreans in Unified Korea will not only need to understand these words as they are used in context but also contend with the idea that the South Korean economy values entirely different principles than a North Korean economy centered on “self-reliance.”

In Unified Korea, it will be the North Korean economy that integrates into the South’s—not the other way around. Therefore, the burden of assimilating will be on the North Koreans. While it is reasonable to assume that South Koreans can generally understand their North Korean counterparts, the reverse may not apply. Additionally, if South Koreans perceive their North Korean counterparts to be less capable or educated based on their word choices or way of communicating, North Koreans will continue to encounter professional discrimination in the workplace. Defector testimony provides ample evidence of this. In an interview, one defector recalled an incident at work where his boss had asked him to bring his “diary” (다이어리), a “Konglish” word that North Koreans would know as *il-jī* (일지). The defector did not recognize the word, thinking instead that the “diary” might be some sort of appliance. The boss confronted the defector in front of his coworkers, who laughed when the boss asked, “No way—you don’t know what a diary is?” The defector felt humiliated to the point of tears and remarked that this incident was only the start of many. Hence, even small miscommunications due to the language difference could result in significant emotional distress for North Koreans who try to keep pace in the Unified Korea workforce.

The divergent language is a challenge in and of itself, but it can also reveal other potential issues. This becomes especially noticeable in more technological fields, like medicine. For example, there are the obvious linguistic differences, such as different words for medical procedures like “drawing blood” or spellings of words like “lung.” However, North Korean physicians would find themselves at a standstill with the words “intern” and “resident,” not only because these words are words borrowed from English but because these positions do not exist in the North. This is because North Koreans are full-fledged doctors after completing medical school, whereas their South Korean counterparts follow a medical training system akin to what is seen in the United States and other parts of the West. How Unified Korea will handle the issue of physicians who have not been trained under comparable conditions is beyond the scope of this paper. However, the language difference demonstrates that economic integration in the medical field will be more than simply informing doctors and patients of vocabulary differences for various ailments and procedures. Rather, the divergence shows a fundamental difference in the way doctors practice medicine, which could have significant implications for public health in the future. It is reasonable to assume that similar chasms will evolve in
other fields that require highly specialized vocabulary, such as law, engineering, and computer science.

Finally, language may impede progress in Unified Korea in ways South Koreans today may take for granted. For example, North and South Korea use different keyboards when typing, which could be hugely disruptive for North Koreans working on computers without a North Korean typing setting. Internet, social media, and texting apps like KakaoTalk play a major role in developed economies in terms of how companies conduct their operations, interface with customers, and brand themselves online. With this comes a set of language norms that differ from those “IRL” (“In Real Life”). How South Koreans communicate virtually—from Internet slang, to emojis, to texting etiquette—could prove yet another professional and social barrier for North Koreans in Unified Korea. While South Korea is one of the most connected countries in the world with nearly 98% of all households connected to the Internet, Internet access in North Korea is heavily restricted. Online communication may prove one of the trickiest grounds to navigate, as virtually no North Koreans will have ever used a smartphone or the Internet.

Interpersonal Relationships and Social Cohesion

Language both reflects a given society’s values and priorities while simultaneously reinforcing these norms. One manifestation of this principle is in the use of direct versus indirect language, which unification scholar Young-sun Jeon argues cites as one of the chief cultural-linguistic differences between North and South Korea. For example, South Koreans might use vague or indirect expressions to anticipate or better understand the intentions of the person with whom they are speaking. In other words, a speaker might have a preoccupation with how his words are perceived among his peers, a principle Jeon refers to as “language etiquette” (eoneo yejeol / 언어예절). For example, a South Korean person would be unlikely to say to a peer over the phone, “I’m busy, so let’s hang up,” because doing so might cause the other person to experience negative emotions. Instead, they might say “I will call you later” or “Something urgent has come up…” to soften the blow. Contextually, the meaning is the same, but the social norm is clear. Even non-native speakers of Korean or tourists in South Korea have encountered the country’s highly liberal use of the phrases “maybe” or “that might be difficult,” which almost always mean “no” and “that is impossible.”

North Korea, on the other hand, values efficiency in communication, both at the interpersonal level and in propaganda. North Korean society is more direct in its communication, and expressions tend to take on a more literal meaning as a result. In the example above, “I’ll call you later” to a North Korean would likely be interpreted as a literal plan of action rather than a polite way of saying, “I’m busy, so don’t wait up.” North Korean defectors have reported experiencing this miscommunication in the workplace, like one defector who interpreted his boss’ statement to “come find him sometime” (to spend time together) in a far more literal fashion than his boss had intended.

While the direct versus indirect language might at first seem like a minor sticking point, all things considered, in reality, it is one of the clearest examples of two divergent societies. South Korea, which emphasizes the individual, tends to favor indirect expression. This in turn manifests in a society with a heightened awareness of how one’s method of communication is perceived socially. In contrast, North Korean society values efficiency when
transferring information (jeongbo jeondal / 정보전달) due to a strong sense of public character. This emphasis on efficiency over social niceties reflects a society where direct communication is essential. There are many reasons for why this might be the case, including its being a consequence of state propaganda that emphasizes directness or the need to reduce ambiguity in one’s speech in a surveillance state. In a low-trust society that requires citizens to confess their sins at regular self-criticism sessions and forces neighbors to report on one another, what is considered etiquette in the South might be considered duplicity in the North. And duplicity, like other “moral failings,” can lead to severe punishments in the DPRK. In any case, this cultural difference in communicating could lead to serious miscommunication and frustration in Unified Korea. Furthermore, this element of language could indicate further issues in attempts to meld two societies with vastly different social norms and priorities. More seriously, in a sense, the way North Koreans speak could reflect mass trauma inflicted by the North Korean state.

Korean linguist Lee Eun-hye argues that language etiquette is a unique feature of the Korean language and reflects human relationships in Korean society. One place where this becomes readily apparent is with the use of formality and language etiquette in North and South Korea. For North Koreans, language etiquette is not merely a way of speaking or expressing formality, but a reflection of the correct way—doubtlessly dictated by the regime—to live everyday life. Both Koreans use tiered language to reflect varying levels of formality and social hierarchy. For example, a high school student would use a different language when speaking with a same-age friend than when speaking with a teacher. However, in North Korea, the highest levels of language formality are reserved for North Korea’s idolized dictators. North Korean grammar rules dictate that when a speaker mentions a North Korean ruler, all other subjects in the sentence, even if more senior than the speaker, may not have honorifics such as -si (سى) or -ggeseo (گجهسى) applied to their names. These rules do not apply in South Korean social situations, as these honorifics are used in various contexts (such as grandparents, teachers, or figures of authority). Furthermore, the deification of North Korea’s dictators represents a clear case in which language both shapes North Korean society and contrasts with South Korea’s societal value system. Unified Korea would need to both reconcile this issue from a grammatical standpoint and also from a societal restructuring one.

While social hierarchy is a feature of both North and South Korean societies, this is especially true in the North. Furthermore, the North Korean language is inherently militarized, and this is reflected in the way North Koreans address one another. In sharp contrast to South Korean society, where terms like “older brother” and “older sister” are ubiquitous (even among non-siblings), these terms are forbidden in the North in public settings. Instead, the regime requires North Koreans to address each other as dongmu (동무) and dongji (동지) or “comrade,” to reinforce Juche ideology and remind citizens that their position in society is principally as members of the socialist state that Kim Il Sung envisioned. Conversely, South Koreans may use the word dongmu to typically mean “friend” though not as a form of address to be added to someone’s name, like they would -ssi (-씨), a commonly used honorific for familiar or casual contacts. In North Korea, dongmu specifically means comrade, or “people who fight together in the revolutionary ranks.”

North Korea’s caste system known as songbun may complicate matters further if
language among and between castes varies. Further research is needed to determine whether this is the case.

The extent to which North Koreans have internalized these social norms could have major ramifications for unification. The Kim regime has endeavored to shape the identity of its citizens chiefly around the state. Before North Koreans are siblings, children, classmates, or coworkers, they are first and foremost, “comrades.” North Korea’s “military first” ideology (son’gun / 선군) has contributed to the influx of militarized vocabulary in everyday speech. Kim Il Sung believed that the North Korean language should express militant emotion and “reflect the necessities of the working class.”

Hence, simple phrases in the South, such as “to try hard” have become “to wage a struggle” in the North. 

**Policy Recommendations**

Unified Korea faces a daunting task. Reunifying the Korean language will require more than a North-South dictionary or new rules regarding spelling and grammar (although all this will also be necessary). Rather, Unified Korea must consider how language shapes and reflects the beliefs, values, and societies of the people within it—and bridging this gap will be a far greater undertaking. While multiple Korean linguists have proposed solutions for unifying spelling, grammar, and, to some extent, vocabulary, there exists no comprehensive framework for policymakers to prepare for language unification at the societal level. Policymakers should adopt a flexible and practical policy framework that addresses both the logistical and ideological implications of language unification.

Furthermore, while unification may not be on the horizon anytime soon, policymakers should still consider proactive measures they can enact today, as well as those that can be implemented during and after unification.

**Policies for Today**

To start, the South Korean government today should begin thinking about ways to bridge the knowledge gap between North and South Koreans. To this end, information operations offer one of the only ways for North Koreans to gain exposure to the outside world. Information campaigns are a preemptive measure that can help introduce North Koreans to vital concepts such as Western systems of governance, economic exchange, and culture. Anecdotal evidence suggests that South Korean media is a highly effective conduit for transmitting cultural concepts into North Korea. For example, a 2019 study by South Korea’s Unification Media Group revealed that more than 90% of defectors had viewed some sort of foreign or South Korean media prior to defecting. In late 2020, North Korean dictator Kim Jong Un ushered in a series of new laws targeting “reactionary ideology and culture” to address the influx of South Korean media flooding the country, in particular the drama sensation “Crash Landing on You.” The drama, which features a South Korean woman falling in love with a North Korean man, partially takes place in North Korea and was a hit among defectors and North Koreans alike. Thanks to the popularity of South Korean dramas such as this, some North Koreans have secretly adopted South Korean accents, slang, and even sarcastic phrases mocking the Dear Leader.

Kim Jong Un, like his grandfather Kim Il Sung, clearly appreciates the connection between language and identity. He responded in-kind by passing laws that promised two years of hard labor simply for speaking in a South Korean accent and 15 years for even possessing South Korean media content. The regime’s desire to tamp down on foreign media showcasing a world much wealthier and freer than North
Korea is an understandable response. However, that the regime specifically passed legislation targeting the use of South Korean speech reveals the power language can have in shaping minds. Information campaigns offer not only a powerful means of arming the North Korean people with knowledge of the outside world but may also be highly effective at preemptively introducing key South Korean cultural and linguistic concepts. Naturally, the North Korean government will likely react negatively to attempts to alert its citizens to the outside world, so policymakers must anticipate the consequences of potential escalation in exchange for providing North Koreans with valuable outside information. That said, several prominent North Korean experts have argued that mass information campaigns may actually deescalate tensions between North Korea and the West in the long run.\(^8\)

Second, policymakers in South Korea should begin thinking about which areas of language unification will be most challenging and urgent. To do this, policymakers should work with linguists, North Korean defectors, and subject-matter experts to identify which parts of the language will need to be standardized immediately and which parts can be unified at a later date. Policymakers should also consider how language unification will align with other unification goals (for example, the unification of public health services). In developing these “prioritization buckets,” policymakers should carve out core areas like medical care, law enforcement, and other places where language standardization might be most urgent post-unification. From there, policymakers can dive in-depth in each area and identify key vocabulary, phrases, and social context that will bridge the linguistic gap and enable North and South Koreans to communicate across essential areas. Policymakers can develop “tiers” of these buckets with appropriate timelines for implementation based on their urgency. For example, it will be far more important to unify vocabulary across essential critical infrastructure and public services than it will be to standardize the use of honorifics.

Since work is already underway for the creation of a North-South joint dictionary, the focus should be less on identifying translation for all words and more on where the language will be functionally different. Additionally, these buckets should address how both North and South Koreans will face challenges in different ways. For example, if South Korean medical professionals need to mobilize in the north to address North Korea’s lack of public health infrastructure, what will doctors, nurses, and technicians need to know when it comes to communicating with North Korean patients and medical personnel? What will North Korean patients need to know when working with South Korean doctors? This dual perspective approach ensures that the needs of both Koreas are incorporated in future language policies.

This policy recommendation is intended to be iterative in nature. Policymakers should repeat this process as time progresses to ensure the prioritization buckets are timely and relevant. The implications of this policy are twofold. First, this policy ensures that South Korea is prepared for a unification that may be sudden or chaotic in nature, such as one resulting from internal implosion of the North Korean government. Second, by working with North Korean defectors, policymakers can use these testimonies to extrapolate what potential sticking points might arise in the future. While vocabulary is likely to be the most significant hurdle for Unified Korea when it comes to day-to-day communication, South Korean policymakers should consider how
to unify grammar, syntax, and spelling as well.

Policies for the Future

Because the future of unification is uncertain, it is vital that policy recommendations be flexible and adaptable to multiple circumstances. The prioritization buckets will hopefully provide future policymakers with some much-needed clarity on which tasks to tackle first, but the nature of unification will dictate much of what is to come. Future policymakers will need to focus on proactive education of school-aged North Koreans; standardize priority vocabulary, spelling, and syntax; and eliminate remnants of the language that overtly or covertly lend credence to the Kim family ideology. That said, future policymakers will need to be adaptable and incorporate flexibility into their models to allow for organic language development among North Koreans.

While it will be important to begin establishing a sense of national identity, policymakers should not neglect the role local identity may play in allowing future North Korean society to develop. The purpose of language policy should not be to force every Korean person to speak with the Seoul dialect. Rather, policies should recognize that there is a place for “official dialect” (such as in government notices, newspapers, textbooks, etc.) and “unofficial dialect.” This will allow North Koreans to organically experience the same kind of language evolution South Koreans have experienced over the past several decades. This will help mitigate negative perceptions—such as one accent being seen as superior to another—while also giving North Koreans a sense of autonomy in developing their own identities. Additionally, local North Korean dialects should be allowed to flourish and develop naturally. Future policies should reflect the very best of democratic societies, chief among these being the empowerment of the average citizen.

Conclusion

Dong Mun Seo Dap (동문서답/東問西答) is a four-character idiom in Korean that roughly translates to, “The East asks, but the West answers.” It is used when one person asks a question, but the listener mishears and responds with a completely incoherent response. In many ways, the unification of North and South Korea will embrace this saying to the extreme degree. Except in the case of Unified Korea, it will take more than a simple clarification for the two sides to truly understand one another. The policymakers of Unified Korea face a daunting task in reconciling the linguistic differences of North and South Korea. North Koreans will likely experience these differences more acutely than their South Korean counterparts will, but both sides will face communication challenges.

As this paper demonstrates, the divergent Korean language is not only a product of North Korea’s isolation and the exacerbation of existing dialectic differences, but also the deliberate attempt of the North Korean regime to use language as a mechanism for societal control. Testimony from North Korean defectors consistently reinforces how language may challenge unification efforts, from vocabulary differences to a mismatch of social norms. While Korean language literature—primarily that of linguists from the early 2000s—lays out some of these differences, the true challenge will be understanding how the Korean language shapes and reflects two wholly different Korean identities.

Policymakers today should consider how language will undermine future unification from both logistical and ideological perspectives. On the logistical side,
policymakers will need to determine how to standardize spelling, grammar, and vocabulary in key sectors. This will require some soul-searching on the part of the South Korean government in terms of the extent to which the South Korean language (pyojun-eo / 표준어) should prevail as the “official” dialect and how much North Korean language (munhwa-eo / 문화어) should be incorporated. Although this paper primarily focused on ideological issues over logistical ones, future research on this subject should explore how policies targeting each should align.

This paper almost exclusively centered on the Korean experience regarding the relationship between language and identity, but further work should explore whether other global cases may apply to Unified Korea. For example, Arabic, Chinese, and Spanish are spoken in many countries worldwide across a range of dialects. The extent to which these various dialects shape the identity of the countries in which they are spoken could reveal interesting lessons for Korea. Similarly, countries with multiple official languages or dialects of the same language could be another point of comparison. However, cases where language and identity are strongly linked will yield more compelling insights.

Finally, there is room for additional language research on the Korean Peninsula itself. Whether the songbun (성분) caste system influences North Korean speech and identity on a more granular level is a particular area of curiosity. Similarly, future research should incorporate interviews from North Korean defectors, who cannot only speak to the linguistic challenges of assimilating in South Korea, but also how their own identities have evolved after defecting.

It is difficult to corral policymakers to mobilize on issues for current events, much less those whose future and timing are uncertain. However, given the colossal challenge of unification, South Korean policymakers must begin making concrete but flexible plans today. At the very least, understanding the language differences between North and South Korea may have immediate benefits in terms of understanding North Korean identity, which could prove useful when crafting other policies toward North Korea. As time passes, the gulf between North and South Korea will only grow larger. Language should be a key priority for unification specialists, or else Unified Korea risks a society that is unified in name only.

About the Author: Olivia Schieber holds an M.A. from Georgetown University’s Security Studies Program and a B.A. from Johns Hopkins University. Her research focus is on the Korean Peninsula, North Korean cybercrime and sanctions evasion, and the intersection of technology and national security. She is grateful for the support of her research advisor, Dr. Kimberley Roberts, and to Georgetown Reference Librarian Ding Ye, who secured much of the Korean-language source material that was necessary for this study.
Appendix

Table 1: Sample of North-South Pronunciation Differences

<table>
<thead>
<tr>
<th>South Korean</th>
<th>North Korean</th>
<th>Example-South</th>
<th>Example-North</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>ㄴ</td>
<td>균</td>
<td>노동 (nodong)</td>
<td>로동 (rodong)</td>
<td>Labor</td>
</tr>
<tr>
<td>ㅗ</td>
<td>균</td>
<td>역사 (yeoksa)</td>
<td>력사 (ryeoksa)</td>
<td>History</td>
</tr>
<tr>
<td>ㅗ</td>
<td>농</td>
<td>여자 (yeoja)</td>
<td>여자 (nyeoja)</td>
<td>Girl</td>
</tr>
<tr>
<td>ㅅ</td>
<td>쌍</td>
<td>원수 (wonsu)</td>
<td>원수 (wonsu)</td>
<td>Enemy</td>
</tr>
<tr>
<td>ㅅ</td>
<td>집</td>
<td>깊발 (gi’bal)</td>
<td>깊발 (gibal)</td>
<td>Flag</td>
</tr>
<tr>
<td>ㅂ</td>
<td>농</td>
<td>손백 (sonbyeok)</td>
<td>손백 (sonbyeok)</td>
<td>Clap</td>
</tr>
</tbody>
</table>

Table 2: Sample of North-South Vocabulary Differences

<table>
<thead>
<tr>
<th>South Korean</th>
<th>North Korean</th>
<th>English</th>
<th>Context</th>
</tr>
</thead>
<tbody>
<tr>
<td>에스컬레이터 (ehseukeolleiteo)</td>
<td>계단승강기 (gyedantseunggang’gi)</td>
<td>Escalator</td>
<td>North Korea shows a preference for Korean words (either native or from Chinese origin) over English loanwords</td>
</tr>
<tr>
<td>핸드폰 (haendepon)</td>
<td>손전화 (sonjeonhwa)</td>
<td>Cellphone</td>
<td>English loan word vs. Korean word</td>
</tr>
<tr>
<td>캠페인 (kaempein)</td>
<td>람빠니야 (ggambbaniya)</td>
<td>Campaign</td>
<td>In some cases, North Korea will adopt the Russian pronunciation of an English loanword.</td>
</tr>
<tr>
<td>컵 (keop)</td>
<td>고뿌 (gobbu)</td>
<td>Cup</td>
<td>In other cases, North Korea will adopt the Japanese pronunciation of an English loanword.</td>
</tr>
<tr>
<td>한복 (hanbok)</td>
<td>조선옷 (Joseonot)</td>
<td>Hanbok (traditional Korean clothing)</td>
<td>For words that reference &quot;Korea&quot;, South Korea uses 한 (han) while North Korea uses 조선 (Joseon). This corresponds to how each country refers to itself.</td>
</tr>
<tr>
<td>기억 (giyeok)</td>
<td>기록 (gieuk)</td>
<td>Memory</td>
<td>Korean words with similar, but different pronunciation</td>
</tr>
<tr>
<td>이익 (i’ik)</td>
<td>밥을알 (meokeulal)</td>
<td>Profit</td>
<td>While North Korea still uses words derived from Chinese, it prefers Native Korean words, as seen here. The North Korean phrase roughly translates to &quot;eggs to eat.&quot;</td>
</tr>
<tr>
<td>오징어 (ojingeo)</td>
<td>낙지 (nakji)</td>
<td>Squid</td>
<td>North Korea uses the South Korean word for &quot;octopus&quot; to mean squid</td>
</tr>
<tr>
<td>낙지 (nakji)</td>
<td>오징어 (ojingeo) or 서해낙지 (seoheae...)</td>
<td>Octopus</td>
<td>...and they use the South Korean word for &quot;squid&quot; to mean &quot;octopus&quot;!</td>
</tr>
<tr>
<td>폴란드 (Pollandeu)</td>
<td>뽀스카 (Bbolseugga)</td>
<td>Poland</td>
<td>While South Korea often adopts the English transliterations of country names, North Korea typically adopts the transliterations of country names as they are pronounced in that country.</td>
</tr>
</tbody>
</table>


6 “南北한 언어 차이 심각… 일반어 38%, 전문어 66% 달라,” [“Serious language difference between North and South Korea…38% difference for everyday language, 66% for specialized language,”] VoA Korea, November 1, 2016. https://www.voakorea.com/a/3574287.html.


14 Yeon, “‘Standard Language’ and ‘Cultured Language,’” 42.


24 Scholars have different interpretations of the exact years for these periods. For example, Min-su Kim breaks these periods into 1945-1954 (Period of Unified Orthography), 1954-1966 (Period of Revised Orthography), and 1966-present (Period of Prescriptive Orthography). In contrast, Chin-wu Kim and Hui-sook Kim recognize two periods (the exclusive use of Hangeul and the munhwa-policy and linguistic trends these scholars identify remain roughly the same. For details, see, Jae Sun Yeon, “‘Standard Language’ and ‘Cultured Language,’” in Korean Language in Culture and Society, ed. Ho-min Sohn (University of Hawai’i Press, 2006), 32-33.

25 For the sake of clarity and consistency, this paper will use “Joseongul” when referring to language policy in North Korea and “Hangul” in all other instances.


27 For example, South Koreans tend to use the word hongsu (홍수), which comes from the Chinese characters 洪水, to mean “flood.” In contrast, North Koreans use the native Korean word keummul (큰물), literally “big water.” In Korean Language in Culture and Society, Yeon cites several interesting examples of this nativization, which also applies to foreign-loan words common in South Korea. South Koreans use the word p’ama, adapted from the English word “perm,” whereas the North Korean equivalent, bokkeum meori (볶음 머리), is simply “fried hair.”


32 Strand, “Questioning North Korea’s Digraphic Inflection Point”, 266.
With a 38-66% difference in vocabulary, it would be impossible to showcase every instance where vocabulary between the two countries differs. For additional reading, see Gyeoremalkeun Dictionary North-South Joint Compilation Project. 한눈에 들어오는 남북생활용어 I [North-South Everyday Terms at a Glance 1], (Seoul: 맵씨터 [Mapseater], 2017); Gyeoremalkeun Dictionary North-South Joint Compilation Project. 한눈에 들어오는 남북생활용어 2 [North-South Everyday Terms at a Glance 2], (Seoul: 맵씨터 [Mapseater], 2019); and Yeong-choel Yang, 남북한언어 탐구생활 [North-South Language Exploration] [Seoul: 지식의 숲 [Forest of Knowledge], 2018].


46 For an in-depth look at North Korean propaganda narratives that center on ethnic purity, see B.R. Myers, The Cleanest Race (New York: Melville House, 2010).


50 “하나원, 탈북민을 위한 기초직업훈련 총이 해설집 발간” [“Hanawon publishes a glossary of basic vocational training terms for North Korean defectors"], ROK Ministry of Unification, June 29, 2021.
Internet offline, researcher says

https://datareportal.com/reports/digital

Need to Unify the Different South Korea, the word for lung is spelled

South Korea's version is derived from Chinese characters

For drawing blood, South Koreans use プロック (pibboggi), while North Koreans typically use 피뽑기 (피뽑기).

South Korea’s version is derived from Chinese characters (采血), while North Korea uses native Korean words.

In South Korea, the word for lung is spelled 北 (pye), while in North Korea, lung is spelled 北 (peh).

Gyeoremalkeun Dictionary North-South Joint Compilation Project, 한눈에 들어오는 남북 생활용어 1 [North-South Everyday Terms at a Glance 1], (Seoul: 백석터 [Mapseater], 2017), 49.

For drawing blood, South Koreans use Chaehyeyol (체혈), while North Koreans typically use pibboggi (피뽑기).

South Korea’s version is derived from Chinese characters (採血), while North Korea uses native Korean words.

In South Korea, the word for lung is spelled 北 (pye), while in North Korea, lung is spelled 北 (peh).

Gyeoremalkeun Dictionary North-South Joint Compilation Project, 한눈에 들어오는 남북 생활용어 2 [North-South Everyday Terms at a Glance 2], (Seoul: 백석터 [Mapseater], 2019), 94.

For drawing blood, South Koreans use Chaehyeyol (체혈), while North Koreans typically use pibboggi (피뽑기).

South Korea’s version is derived from Chinese characters (採血), while North Korea uses native Korean words.

In South Korea, the word for lung is spelled 北 (pye), while in North Korea, lung is spelled 北 (peh).

Gyeoremalkeun Dictionary North-South Joint Compilation Project, 한눈에 들어오는 남북 생활용어 2 [North-South Everyday Terms at a Glance 2], (Seoul: 백석터 [Mapseater], 2019), 96.

Noh, Dong-gyun. “kept in the dark by the North’s secret service.” Chosun Ilbo, October 26, 2018. K

https://it.chosun.com/site/data/html_dir/2018/10/26/2018102601977.html

Kemp, Simon. “Digital 2023: South Korea.” Data Reportal, February 13, 2023,


Conversely, the honorifics -시 and -께서 must always be used when speaking or writing about North Korea’s leaders. Moun Kyoung Shin, “A Comparative Study of Honorific Systems in North and South Korea: Shifts since 1950,” (PhD Dissertation, University of York, 2017), 112-114, Proquest 27679000.


Shin, Moun Kyoung. “A Comparative Study of Honorific Systems in North and South Korea: Shifts since 1950,” (PhD Dissertation, University of York, 2017), 118, Proquest 27679000. https://www.proquest.com/dissertations-theses/comparative-study-honorific-systems-north-south/docview/2307312484/se-2. It is common for women to address older female and male friends respectively as unni or oppa (언니/오빠) and for men to use nuna and hyung (누나/형). While these terms directly translate to “older sister” or “older brother,” they are common endearments for boyfriends and girlfriends. The South Korean drama, Crash Landing on You, a romantic comedy featuring a woman who becomes stranded in North Korea, was immensely popular in North Korea, where residents watched it illegally. The drama was highly influential, prompting North Koreans to (illegally and privately) adopt the South Korean endearments. See, Olivia Schieber, “A culture war is brewing in North Korea. It shows Kim Jong Un’s deepest fear,” The Washington Post, February 2, 2021. https://www.washingtonpost.com/opinions/2021/02/02/culture-war-is-brewing-north-korea-it-shows-kim-jong-uns-deepest-fear/.


Gyeoremalkeun Dictionary North-South Joint Compilation Project, 한눈에 들어오는 남북 생활용어 2 [North-South Everyday Terms at a Glance 2], (Seoul: 밀씨터 [Mapseater], 2019), 185.


Yeolsimhi noryeokhada (열심히 노력하다) in the South can mean “to zealously put in effort,” while tujaeng-eul beolida (투쟁을 벌이다) in the North translates roughly to, “to wage a struggle.”


The Decaying Superpower: A Review of the Russian Navy

Bora Balçay

The Russian Navy has undergone profound changes since the dissolution of the USSR, reflecting its reduced role on the global stage. The Russian Navy has undergone three distinct stages in the post-Soviet era. In the first stage, the focus of the country was on restoring operational capability to a vast and expensive fleet from the Soviet era; therefore, the improvement of the Russian Navy was greatly neglected during this time. The second stage saw the creation of a deliberate strategic vision for the Russian Navy that sought to develop power projection capabilities. During this time period, Russia made a concerted effort to acquire and field new platforms that were smaller but more modern in kit, protecting the country’s strategic nuclear deterrent capabilities. Finally, after the invasion of Crimea in 2014, Russia’s Navy entered a third stage, where a comprehensive sanctions regime profoundly affected its economy and thus constrained its ability to spend on military assets, especially the Russian Navy. Consequently, modernization efforts of the Russian Navy have either slowed significantly, been scrapped, or postponed indefinitely. These developments indicate that Russia’s future naval strategy is less ambitious about projecting power beyond its immediate borders through the seas. It also means that the naval medium will likely be confined to a leg of the Russian nuclear triad.

Introduction

During the Cold War, the Soviet Union sought to compete with the capitalist West in all domains. As the single largest country on earth, the Soviet Union and its predecessor, the Russian Empire, were primarily land powers. However, as a superpower, the Soviet Union pursued an ambitious naval policy, particularly during the tenure of Admiral Sergei Gorshkov between 1956 until the collapse of the USSR in 1991. The Soviet Union spent considerable resources to develop its naval forces in a manner that not only sought to defend the expansive territorial waters of the Soviet Bloc but also to project power wherever needed. To this end, the USSR constructed a large and advanced surface fleet that included large surface combatants and aviation assets with a considerable blue-water presence, in addition to a considerable fleet of nuclear-powered submarines that secured the USSR’s second-strike capability.

The collapse of the USSR in 1991 left the Russian Federation with an immense fleet with an undersized role. Between 1991 to 2024, the Russian Navy has seen three distinct stages, which can be traced to its national security posture at the time. In the initial stage, which lasted from the early 1990s to the mid-2000s, the Russian navy solely sought to maintain the former Soviet fleet while the new country focused on stabilizing its domestic and regional decline. In the mid-2000s, Russia devised a consistent strategic vision to incorporate a naval component which led to considerable naval building plans, which this paper refers to as a second stage. Finally, Russia’s navy entered a third stage following the 2014 invasion of Ukraine, where the modernization of the Russian Navy was stifled due to sanctions and economic downturn. Whereas other branches of the Russian military receive considerable attention as the country takes a regionally aggressive posture to impose its territorial claims by force, the standing of the Navy
seems to be regressing, except for its nuclear deterrent role.

Russia’s economic advancement following the collapse of the Soviet Union enabled it to plan for a large naval force that could project Russia’s influence abroad, while recent Russian aggression has limited the country’s economic prospects and its ability to build and sustain a large navy. The paper concludes with a brief analysis of the future of the Russian Navy using Russia’s investments in the Russian Navy as an indicator of the direction of Russian security policy and strategy. The future of the Russian Navy mirrors that of the Russian nation, one whose grand ambitions are constrained by its behavior abroad. Unless Russia invests in naval modernization, its Navy is likely to struggle with the development of new platforms and the modernization of its large surface combatants; but it will likely continue experimenting with asymmetric platforms and keep up its naval nuclear posture.

**Background: The Soviet Fleet Before the Collapse of the USSR**

**Vessels**

The naval forces of the Soviet Union were an expansive undersea component and an impressive ensemble of surface combatants of all sizes. At the center of the Soviet naval forces were the nuclear ballistic missile submarines (SSBNs) that carried a core leg of the Soviet nuclear triad, ensuring the Soviets secure second-strike capability throughout the Cold War.1 Accompanying these SSBNs were various types of attack submarines, nuclear and conventional, that sought to protect the fleet and hunt American SSBNs.

For its surface fleet, the USSR possessed a wide variety of assets that reflected the complex geographic and strategic realities that the country faced. To operate in the Baltic and Black Seas, the USSR had fast attack craft, missile boats, and corvettes, which contested the NATO naval presence in these relatively small seas. The defense of these seas posed a fundamental strategic necessity since both the Baltic and the Black Seas offered pathways to major population and industrial centers of the USSR.2 Thus, these small but well-armed and numerous vessels represented a core component of the defense of the Soviet heartland.3 However, the small surface combatants were not the most potent assets of the Soviet Navy, for they also possessed larger vessels and sophisticated capabilities like nuclear cruisers, cruisers, and destroyers. These ships adopted an approach that assumed direct roles in countering American vessels, notably anti-surface and anti-submarine warfare duties. In addition to these major types, the naval forces of the USSR had a collection of other vessels ranging from amphibious transports and frigates to spy ships and arctic ice breakers as well that endowed the country with a robust range of capabilities.

There were, however, certain types of combatants that the USSR lacked in its fleet. Notably, it lacked the large amphibious assault capabilities that the United States possessed for expeditionary tasks and power projection. This indicated that the Soviet Navy did not seek to project power over long distances as it sought to develop countermeasures to NATO naval assets rather than developing distinct power projection capabilities. Until the late 1980s, the Soviet Union did not have large flattop carriers for operating fixed-wing aircraft, analogous to the supercarriers Americans had throughout the Cold War.4 Only after the continuing issues of STOVL aircraft, which struggled to perform as well as catapult-launched aircraft, did the USSR finally seek to develop an aircraft carrier comparable to American carriers.5 While the
USSR had the industrial capability to construct vessels of larger sizes and scope, its choices were constrained by bureaucratic concerns and doctrinal preferences. An outstanding limitation that the USSR faced in constructing naval aviation assets was the Montreux Convention which forbade the presence of aircraft carriers in the Black Sea. No Soviet aircraft carrier was ever designated as such since the USSR circumvented the restrictions on the transit of aircraft carriers through the straits through clever naval taxonomy. Even the Kuznetsov-class vessels, which displace 55,000 tons, are classified as “aviation cruisers” as they also integrate other capabilities such as anti-ship missiles.

**Bases**

There were four major centers of the Soviet Naval Forces. Murmansk hosted the Northern Fleet and the majority of the USSR’s SSBNs, the most important component of the Soviet Fleet. The northern port was strategically significant because it allowed access to the Arctic Sea and did not get frozen during the winters. Its location also meant that it was difficult to blockade and sufficiently far away from NATO countries to prevent a surprise attack. Another significant naval base was Sevastopol, which in addition to housing the Black Sea fleet, also possessed some of the largest shipyards in the USSR; notably, the Soviet aircraft carriers were constructed there. The naval base in Vladivostok hosted the Soviet Pacific fleet and served as another station for Soviet SSBNs. The Baltic Fleet, despite its prestige and historical significance during the Tsarist Era, was less significant for the USSR due to the geographic constraints of the Baltic Sea; access to the sea was through the Danish straits, and Denmark was a member of NATO.

**First Phase: The Collapse of the USSR**

*The State of the Russian Fleet*

The aftermath of the USSR’s collapse was rather unexpected in terms of naval planning. The Russian Federation inherited the vast majority of the Soviet Fleet, and it broadly continued the Soviet Navy’s shipbuilding plans throughout the 1980s. In fact, when the USSR’s dissolution took place, it was in the process of building two aircraft carriers in Sevastopol, now situated in the newly independent Ukraine. Thus, the mere maintenance of the Soviet Navy was a financial burden for the new Russian state. There was also the issue of Sevastopol, as the important naval base ended up outside of Russia’s borders. The Sevastopol base and its shipyards continued to serve as a major naval hub, and its facilities made it indispensable for the Russian Navy. While this was not an issue at the time, as the Russian government obtained a favorable agreement from the Ukrainian state that secured the base's lease, it would become a vulnerability as Ukraine progressively sought to affiliate with the West through an affiliation agreement with the EU.

The Russian Navy attempted to form an operational core from the remnants of the Soviet Fleet, a grave shortfall that defined this era. During this time, Russia sought to consolidate the seaworthy assets left over from the USSR, which required major upkeep, given that the Soviet fleet had focused much of its resources on constructing new ships. Furthermore, the supply chains necessary to upkeep much of the Russian Navy laid outside the borders of Russia, especially because many of the shipyards and docks essential for the maintenance of vessels were in Ukraine. However, during this era, Russia also began making the difficult decisions of which modernization proposals to advance and which ones to axe. This process resulted in
much of the former Soviet fleet being rendered inoperable as ships were left to rust in Murmansk and other Northern ports.  

Making Hard Strategic Choices  

During the second half of the Cold War, under the leadership of Admiral Sergei Gorshkov, the Soviet Navy had adopted a distinct and decisively offense-oriented blue-water approach aimed at threatening American carriers to deter them from threatening the USSR.  

This approach posed a problem for the modern Russian Navy since most Soviet ships were single-mission oriented and unsuitable for the new mission sets that modern battlefields required, such as advanced anti-access and area-denial. The offensive-oriented boats of the Russian Navy were ill-suited to counter the modern threats as aircraft and cruise missiles went on to pose potent threats to ships.  

To accommodate the challenges posed by the modern battlespace, such as an increased threat posed by aircraft, Russia embarked on projects to construct some major surface combatants, like Udaloy II-class destroyers and Neustrashimy-class frigates. However, these were rare instances of commissioning new vessels for the surface fleet, as most ships under construction were either scrapped or sold to other customers. Consequently, for the first 15 years of the Russian Federation, its surface fleet was severely under-maintained and ill-equipped to face the challenges of the 21st century as it suffered significant defense cuts.  

Contrary to the state of the Russian surface fleet, the Russian sub-surface fleet fared better regarding upkeep and investment in new vessels. Not only were many nuclear ballistic missile submarines under construction completed during the second half of the Cold War, but the sustained development of new classes continued relatively uninterrupted. Russia’s decision to invest in the Soviet Navy’s already robust submarine force demonstrated strategic continuity between the USSR and the Russian Federation, considering the submarine force the premier branch and an essential component of the country’s nuclear deterrent. However, Russia reduced its diesel-powered and propelled attack submarines with the decommissioning of many older boats and the sale of newer ones.  

Russian policymakers prioritized the maintenance of their nuclear deterrent capability through a robust second-strike option, demonstrated by the relative lack of decline in the Russian subsurface fleet.  

However, the fatal *Kursk* incident of August 2000 affected the Russian fleet’s morale. The tragic sinking of the advanced submarine and especially the loss of the crew represented significant failures that became characteristic of the Russian military; the accident was attributed to a faulty weld on a torpedo. Throughout the 1990s and the early 1990s, there were also broader challenges that affected morale, such as systematic corruption, which disturbed the integrity of the inventory and supply chain, and payroll issues.  

A broader challenge, though, was the conscript dependence of the Russian Armed Forces. While appropriate for the land forces due to the lack of required advanced training, the extensive conscription system proved problematic in roles that required advanced training. The resources spent on the training navy personnel became wasted as soon as the conscript left the service.  

Russia also inherited a successful arms export business and cultivated a robust export-oriented military-industrial complex. For example, the Russian Navy sold numerous vessels, including submarines and frigates, that it could not afford to keep operating. Beyond existing ships, many uncompleted hulls destined for the Soviet
Navy were also sold. Notably, the hulls of the aircraft carriers *Varyag* and *Admiral Gorshkov* were sold to Chinese and Indian Navies, respectively. While these sales hurt the future modernization plans of the Russian Navy, they were nonetheless necessary from a financial point of view.

Whereas this era was one of decline for all branches of the Russian Military, other branches retained more relevance in Russian security perceptions. Despite lagging in developing advanced technologies, the Russian Air Force could sustain itself as the Russian aerospace industry continued to be a successful producer of airframes. On the other hand, the Color Revolutions in the former Eastern Bloc and the wars in the former Yugoslavia influenced Russian thinking during this era. This prompted a revival in Russian interventionism as the country became increasingly involved in its neighbor's affairs. During this time, Russia developed a new doctrine, yielded new tactics to dominate the grey zone. The Russian Army developed many tenets of modern asymmetric warfare; these tactics were utilized effectively in the invasion of Georgia in 2008 and the invasion of Crimea in 2014. Dominated by asymmetric warfare, the Navy was reduced to a secondary role for much of the late 1990s and the early 2000s, rendering it unable to challenge the NATO navies in the blue-water environments or project power far from the Russian mainland. Despite operational challenges like force discipline and equipment issues, the Russian Army regained considerable capability after Putin acceded to office.

**Second Phase: The Russian Naval Revitalization**

**A New Fleet**

While the sub-surface fleet of the Russian Navy maintained a central role in the strategic calculus of the Russian state, the surface fleet gained increased significance in the mid-2000s. With successive years of economic growth and a greater political appetite for a large military, Russia displayed comprehensive and ambitious plans for a naval redevelopment. This new doctrine would not only seek to regain many capabilities lost to the decline of the Soviet Navy, such as a fleet air arm, but would also seek to facilitate the development of many new capabilities that had come to form the core of many Western navies.

Implementing a naval construction program only became feasible towards the end of Vladimir Putin’s first term as president, ramping up with Putin's inauguration as prime minister. While the Navy was not the recipient of the largest budget, the injection of capital profoundly impacted the future of the branch. Unlike the other branches that spent their increased budgets on growing their ranks and investing in iterative defense projects, the Russian Navy sought to become a potent blue-water force. Putin’s increased spending on naval plans also coincided with a more internationally active Russia, as the country again started using its military abroad. Especially with the escalation of the Syrian Civil War in 2011, the Russian government recognized the importance of naval strength in supporting Russian allies, which culminated in an operational deployment in 2016.

**Projecting Power Abroad**

The new power projection strategy brought with it a slew of new plans for large surface vessels, including the procurement of Mistral-class amphibious assault ships and the desire to build a new class of destroyers, the *Lider*. The desire to incorporate a new large-format amphibious
assault crafts represented a new form of power projection. The large amphibious assault ships enabled the delivery of large numbers of troops and equipment to shores through landing crafts and aircraft. While during the Cold War, advanced amphibious assault capabilities had been mostly exclusive to the United States, their platform’s flexibility in responding to varied challenges and cost efficiency compared to traditional aircraft carriers led to the development of many new classes in France, Spain, and the United Kingdom. Thus, the Russian Ministry of Defense sought to acquire advanced amphibious assault vessels in the mid-2000s to develop the country’s ability to project power abroad. However, the complexity of the platforms and the unfamiliarity of the Russian industrial base with this kind of vessel meant it would be much easier for the country to explore the export market. Therefore, Russia sought an off-the-shelf solution by ordering two Mistral-class amphibious assault ships from France.

The prospect of procuring an amphibious assault ship from a foreign country was unprecedented for the Russian military-industrial complex, although the plan ultimately did not come to fruition. The idea was met with mixed reactions at home and abroad. In Russia, the significant price tag of the project was hard to justify. On the other hand, the United States criticized France for attempting to sell to Russia, stating it violated the alliance spirit of NATO since it amounted to. Despite the criticisms, the project would have yielded major opportunities for the Russian military and its defense industrial base. In terms of military capability, adding new Mistral would have given the Russian troops similar reach and capability to the United States Marine Corps, who had been operating helicopters for flat-deck amphibious assault ships for decades. The purchase of Mistral would have benefited the Russian industrial base as it would have required the integration of Russian subsystems into the French vessel while also giving Russia a role in constructing a second similar vessel at home. However, after Russia’s 2014 invasion of Crimea, France canceled the contract for the Mistral-class amphibious assault ships, instead selling them to Egypt. Failure to procure these capabilities foreshadowed Russia’s intentions to become more involved in regions far away from the Russian homeland; the logistical and operational capabilities that these vessels would have introduced to the Russian Navy would have provided an immense boost to the operations of the Russian military in Syria and Africa, where the number of Russian troops and the intensity of the operating environment would have benefited from the additional cargo space and command capabilities of an amphibious platform.

Another major project during this time was the development of a new class of destroyers, designated Lider. The Russian Navy had inherited several different classes of destroyers, all designed around specialized missions aligned with the Soviet naval doctrine of targeting the American carrier strike groups. However, these vessels were ill-suited for the new power projection strategy that the Russian Navy had intended to pursue. They neither had the advanced combat systems needed to canvas large swathes of airspace nor the advanced long-range missiles that would be needed to engage over long distances. In addition, advances in shipbuilding that enabled stealth characteristics in new hulls were entirely missing from these Cold War designs, making them highly visible targets for anti-ship cruise missiles. Therefore, Russia set out to develop a new class of anti-access/area-denial platforms to catch up to new projects in the West, such as the
American Arleigh Burke-class, the British Type 45, and the Franco-Italian Horizon-class. The *Lider* project was designed to have a tall mast to incorporate a new phased-array radar and accompanying long-range anti-air missiles launched from vertical launch system tubes, an advanced feature many Russian ships lacked. The ship was designed to form the air defense backbone of a task group or a carrier strike group, as its armament would be capable of range that far exceeded what is considered to be self-defense interceptors. Thus, the design and development of this platform indicated a doctrinal shift from countering the U.S. fleet and their NATO allies to producing a comprehensive task group that could operate far from Russian shores and defend itself in addition to projecting power.

**Conventional Expansion**

The Russians also developed two new classes of frigates to better counter modern threats and incorporate new capabilities, the Gorshkov-class and the Grigorovich-class. For most modern navies, frigates constitute the backbone of the fleet as they combine relevant capability with adequate range and sensors to form a compelling value proposition. In general, they tend to be more mission-specific, with specialized classes focusing on missions such as anti-submarine warfare or anti-air warfare. The Gorshkov-class is more advanced as it has a larger integrated mast that features phased array radar sensors and a stealthier silhouette, while the Grigorovich-class is a cheaper, more conventional design. Initially, the Russian Navy sought to develop Gorshkov-class frigates. However, as the project’s original target slipped by a decade due to the complexity and cost of the project, the Russian Navy developed the Grigorovich-class to export abroad since it was a proven and available design. For instance, the Indian Navy has purchased ships of this class with varying modifications to accommodate its national needs. This platform is modeled after the Krivak-class, an older Soviet design that integrates many modern sensors and weapons systems. In this context, to enable the construction of these platforms, Russian shipyards had to conduct major overhauls in facilities to modernize their production and design processes to simulate Western shipyards.

Another major initiative by the Russian navy was the continued improvements of their corvettes to improve their lethality. These vessels, smaller than frigates and limited in range, were a core component of the Soviet Navy as they were cheaper to produce but fielded lethal anti-ship missiles. The USSR produced them in large numbers, as they were quite cheap, and they made it difficult for NATO navies to operate in the Baltic and Black Seas. In fact, there were more than a hundred corvettes in service at the time of the USSR’s collapse in the Eastern Bloc. Although the Russian Navy sought to develop blue-water capabilities during this era, the ability to dominate the smaller seas that Russia bordered remained a priority due to the geopolitical significance of these areas. The progressive admission of Eastern European states into NATO meant that Russia perceived an enduring threat in the small seas it bordered.

In 2001, Russia embarked to build new corvettes, unusually heavily armed, featuring large cruise missiles, vertical launch system cells for air defense, and capable sensor suites. The heavy armament of these ships limited the range of the corvettes even further, limiting their operation to the Baltic and Black Seas. In terms of doctrine, the development of heavily armed corvettes to especially operate in the smaller seas indicates a recognition of the threat Russia saw in the NATO dominance of these seas, as the types
were designed rationally to threaten NATO ships that prioritize long-range/endurance and usually operate in task groups.\textsuperscript{26}

\textit{The Shipbuilding Industry}

A major issue affecting the pace of the Russian Navy’s expansion programs was the state of the nation’s shipbuilding infrastructure, which had broadly been unable to keep up with the novel practices of Western shipbuilders. Shipyards in the West fully embraced digitalization and modular construction practices in the early 1990s; these revolutionary methods cut the cost of construction and time of design by considerable margins.\textsuperscript{27} Furthermore, the alliance system among the Western allies facilitated technology transfers for sensitive sensors and parts to remain interoperable, thus making the development of many platforms much easier and cheaper across NATO. Whereas the Russian shipyards were stuck with outdated shipbuilding methods, which take more time to construct vessels but also compound the cost of construction.\textsuperscript{28} Additionally, the country experienced shipyard accidents that caused significant damage to military vessels.\textsuperscript{29}

The Russian government identified these efficiency issues, and after extensive scrutiny of the competitiveness of Russian shipyards, the government established the United Shipbuilding Corporation (USC) in 2007.\textsuperscript{30} While the creation of the USC streamlined the industry considerably, it nonetheless failed to address the increasing dependence on Western technology, especially imports of gas turbines. These turbines require extensive development, so the Russian government had to import them. However, the turbines became inaccessible following the successive rounds of sanctions in 2014 and 2022 and USC’s status as a sanctioned entity.\textsuperscript{31} Russia’s endemic corruption also affected the military industrial base. Several scandals have occurred around the misallocation of state funds awarded for naval development programs. The most significant of these scandals was the arrest of the director of the shipyard tasked with repairing and renovating the \textit{Kuznetsov} aircraft carrier.\textsuperscript{32} Despite the government’s efforts to resolve these issues, the problems with military-industrial productivity endure, especially in the light of diverting resources and attention to the war effort in Ukraine.

\textbf{Third Phase: The End of Russia’s Naval Ambitions}

\textit{The Impact of Sanctions}

The invasion of Crimea in 2014 and the following international sanctions affected all segments of the Russian economy, including the country’s military. The devastating economic effect of sanctions after the invasion of Ukraine in 2022 has also severely limited the budgets allocated for many military projects, particularly those requiring large research and development budgets, such as aviation and naval projects. Economic contraction meant that the previously allocated budget for projects such as the Lider-class dried up almost immediately, and the platform was left in perpetual research and development limbo, although a model was unveiled to the public in 2017.\textsuperscript{33} Even more ambitious projects, such as the planned replacement of the \textit{Kuznetsov}-class aircraft carrier, also appear to be stalled.\textsuperscript{34} Therefore, the novel, advanced power projection capabilities that Russia had sought to acquire have become out of reach.\textsuperscript{35}

The construction of the relatively simple platforms also became difficult. The production of the Grigorovich-class, which was supposed to be a simpler design, slowed down considerably because the propulsion plants of the vessels were to be imported from Ukraine, a plan immediately foiled...
after the invasion. Although there were domestic options available, the proposal was not financially viable. As a result, only three of the six planned hulls were commissioned into the Russian Navy, while the remaining three were sold to India in an incomplete state. Presently, for the Gorshkov-class, although plans for procurement remain standing, the delays have pushed the commissioning of these ships by years, significantly deteriorating Russia’s air defense capabilities.

The Impact of Antiquated Equipment

Procurement issues following the 2014 sanctions did not stop the Russian Navy from attempting to implement its doctrine of power projection by carrier aviation with the tools already at its disposal. Notably, the Russian Navy deployed the aircraft carrier Admiral Kuznetsov to Syria to take part in the 2016 Aleppo offensive, accompanied by a relatively robust task force that included two destroyers, a nuclear battlecruiser, and a tugboat. This voyage was a relative failure even though the Russian Navy had never attempted such a mission before. For one, although the carrier did successfully transit from the North Sea to the Mediterranean as soon as combat operations commenced, the neglected state of the boat became apparent. While aircraft from the ship partook in combat operations over Syria, various aircraft were lost at sea because of technical failures on the ship, and the lack of experience of the Russian naval aviators. To add insult to injury, the aging propulsion system of the Kuznetsov failed during the mission, which forced the boat to be tugged back to port. This experience was a painful demonstration of the unpreparedness of the Russian Navy personnel and the inadequate state of their equipment, which could not project power away from Russia’s shores. Regarding power projection abroad, the Russian military still relies heavily on ground forces supported by airlift and locally based fighter jets. This greatly limits the theaters in which the Russian forces can operate.

The Impact of Conscription

Since the beginning of the Ukrainian War, the Russian Navy has played a relatively minor role in the conflict. While the navy’s involvement has been limited to blockading Ukrainian ports, laying mines, and launching cruise missile attacks on specific Ukrainian targets, the sinking of the Moskva, the flagship of the Russian Black Sea Fleet, further illustrates the lack of training in the Russian Navy. The facts surrounding the sinking of the vessel remain disputed, as the Russian claims that it sank as the result of an accident, but most likely, Ukrainian action sunk the vessel. The Ukrainian Armed Forces state they struck Moskva with a Ukrainian Neptune anti-ship missile. However, the Moskva, a Slava-class cruiser, should have had the necessary armament and sensors to defend itself against the Ukrainian missiles, given it possessed advanced radars, self-defense interceptors, and close-in weapons systems that would have been able to meet such a challenge. The fact that the ship sunk despite possessing the necessary countermeasures indicates likely operator or command error in the absence of any other reasonable explanation. In the aftermath of the disaster, reports suggested that the ship’s crew was mainly made up of conscripts. Therefore, the failure of damage control indicates a crucial failing in core competency, for the ability to conduct damage control, is one of the most fundamental skills and exercises for any crew. The significance of this loss cannot be overstated, as the last time a warship of this size was sunk was during the Falklands War.
Future of the Russian Navy

Competition Among Branches

The tightening sanctions regime has only made Russia’s military development plans increasingly unfeasible, potentially prompting competition for resources across the different branches for funding.\(^{44}\) Whereas some modernization plans of the other branches, such as the continuing development for the SU-57 stealth fighter jet has continued, and the Navy’s ambitions for procuring new platforms seem to be becoming increasingly difficult. The disproportional lack of attention to the branch in the context of failed or delayed procurement programs reinforces the adage that Russia’s strategic disposition leans toward being a land power.

It appears that other branches of the Russian military were less affected by these final challenges. Due to its continued involvement in active conflicts such as Syria and Ukraine, the army has seen continued financial support to grow its ranks and procure some equipment. Although the ground forces have not seen a complete overhaul, with a great deal of their equipment still belonging to the Soviet era, they have seen a considerable degree of modernization. In this context, the army has seen new camouflage patterns, main battle rifles, and other pieces of kit to bring it up to par with Russia’s Western peers. The spending on the land forces, however, continues to be taken up by efforts to increase manpower due to the large losses suffered in Ukraine and the ambitious scale of operations in the country.\(^{45}\)

While the Russian Air Force has also faced considerable issues, the gravity of its woes does not match the Navy either. The fundamental doctrine of the force has not changed since the Cold War, and it is unlikely to do so until the country acquires sufficient numbers of stealth jets. Despite the continued issues with its development, the Russian fifth-generation fighter jet is slowly being adopted in the service. This program will endow Russia with an advanced interceptor that will be better able to compete with Western jets and would likely be an attractive proposition for export since fifth-generation fighters are very sensitive pieces of technology. However, the Russian Air Force and the military-industrial base surrounding it represent considerable financial value; Russian-made aircraft are still popular among many legacy operators of Soviet aircraft. Thus, some of the financial burden created by research and development is offset by exporting jets to other countries. However, the War in Ukraine has waned the export performance of some Russian defense products.

Continued Issues with Procurement

The greatest challenge in conventional terms that the Russian Navy continues to face is its inability to procure new hulls and develop new platforms. Endemic issues with the country’s lack of industrial capability that were significant in peacetime have become only more exasperated. The import of technologies and expertise needed to modernize the country’s aging industrial infrastructure has become near impossible due to the comprehensive nature of the sanctions that the country came to face in the aftermath of the invasion of Ukraine in 2022. The older manufacturing, design, and maintenance methods that the country employs, notably not using modular shipbuilding and lack of sophisticated design practices, keep the costs of developing new platforms high and make it difficult to maintain existing vessels. With the increasing complexity of large new surface combatants, as new weapons systems and sensor suites are designed, the difficulty of integrating advanced systems
into new platforms becomes exponentially more complicated, complicating the Russian Navy's development process.

To this end, the delivery projections for the new ships of the Gorshkov-class have slipped back further to the end of the 2020s. The shortfall in the numbers of this crucial vessel that forms the backbone of the Russian naval anti-access/area denial system leaves the fleet vulnerable to aerial threats and anti-ship missiles, thus limiting the Russian Navy's contested environments. The Lider-class appears to have been put on indefinite hold as the information shared with the public has dried up. Although the platform was in the late 2010s, relatively early in its development cycle, it required integrating new technology, making it unlikely that the class will be fielded in 2030s. However, it may be canceled altogether despite Russian claims that the development is ongoing. The feasibility of the plan to develop a domestic amphibious assault platform seems increasingly dubious as well; as the country already experiences challenges with the maintenance of its existing fleet, the development of a new, complex large hull platform that requires many specialized personnel and systems is unlikely.

An area that has not suffered much is the construction of new submarines. The significance of SSBNs in terms of maintaining Russia’s nuclear deterrence remains a core component of the Russian Navy. Thus, boats of the Borei-class have been progressively commissioned into service, relieving the Soviet-era SSBNs of the Russian Navy. The development of new nuclear delivery platforms displays an awareness in the Kremlin that maintaining Russia’s secure second-strike capability takes precedence over developing the country’s power projection capabilities. However, the Russian submarine force’s role as a component of the nuclear deterrent compared to the surface fleet has allowed it to continue developing its boats, equipping its nuclear attack submarines with new capabilities. The new generations of Russian SSNs have integrated cruise missiles, which expands their mission set from just a hunter-killer submarine to one that can also conduct land-attack and anti-ship warfare. Thus, these new tools represent another cost-efficient way for the Russian Navy to perform a limited format of power projection.

**Making Do with Existing Platforms**

The lack of new hulls has prompted the Russian Navy to adopt a more creative attitude to address the gaps in its capabilities. The Navy has started a program to modernize older hulls with newer sensors and weapons suites. For instance, the modernized Udaloy-class features new cruise missiles and, through removing the second turret, numerous vertical launch system cells, accompanied by an upgraded radar system. While it is not unusual for many navies to perform mid-life upgrades, this kind of extensive refit is uncommon for ships as old as the Udaloy-class. Modernization packages are expected to be applied to all of the country's Udaloy-class destroyers. Similar packages may also be developed for other large hull platforms, such as the Slava-class cruiser. Russia has taken a similar approach to the Kirov-class nuclear battle cruisers; these ships have great symbolic significance for the Russian Navy as they are the only remaining nuclear battlecruisers in the world, and their towering stature enables them to carry approximately 200 missiles of varying sizes without having to ever stop for refueling. While vessels of this class are extremely capable, ships of the class have proven difficult to maintain, and half of the class have been decommissioned already,
suggesting the continued operation of these hulls is not cost-efficient.\textsuperscript{50}

The Navy has also committed to the upkeep and modernization of its sole carrier. What shape this modernization will take remains suspect since the vessel can only operate older navalized aircraft such as the Mig-29K and Su-33; these airframes, already fielding outdated avionics and weapons systems, will continue to decay in capability against advanced air defenses and newer, more advanced fighters. The modernization of large older vessels represents a dubious financial proposition. Each modernization package must be uniquely developed for each, even if they use existing systems and weapons, and the refit of vessels themselves can often take years. Because there are few hulls of each class or just one in the case of the Kuznetsov, such packages can be expensive per the number of hulls modernized.

\textit{A Novel Approach}

The Russian Navy, in the absence of the funding necessary for the construction of new hulls, has switched to a strategy of asymmetric development.\textsuperscript{51} In the naval context, asymmetric capabilities often mean adopting novel ammunition that can overcome existing defense systems, developing small unmanned boats that can swarm large ships, and small advanced submarines that can evade detection. Such capabilities are particularly effective because there are no effective countermeasures against them, and they can be very destructive despite their often lower cost.

The Russian Navy has started developing a new generation of hypersonic cruise missiles aimed squarely at the American carriers, just as the USSR had done during the Cold War. Such missiles are especially potent since they pose a potent threat to high-value vessels. These new missiles would integrate the existing ecosystem of sensors and ships in a manner that would make their adoption easy and facilitate their extensive proliferation across the fleet. Similarly, the development of advanced land attack cruise missiles would mean Russia would gain advanced land strike capabilities, similar to the American Tomahawk missile, thereby seeking to address the power projection ambitions of the Russian Navy to an extent. The success of these projects is rather likely given the existing levels of expertise among the Russian military-industrial complex in building such missiles and an existing customer base that could subsidize their development if exported.

Furthermore, the Russian Navy has considered addressing the lack of new submarines with unmanned subsurface missiles to streamline much of Russia’s nuclear deterrence strategy.\textsuperscript{52} These unmanned vessels would roam the seas constantly without the need to stop for resupplying, thereby theoretically addressing one of the greatest weaknesses of traditional nuclear submarines; it would also be much cheaper to operate as it would not need a crew for operation.\textsuperscript{53} Given the nascent state of this technology, despite the Russian expertise in submarines, commissioning these boats in the near future is highly unlikely. The platform’s success would also profoundly destabilize the nuclear equilibrium, as no analog platform exists in the West. Investing into these technologies demonstrates that Russia may be unable to afford the upkeep of a sufficiently large at-sea deterrent force in the future and thus may be preparing for a future where it can rely on cheaper platforms to execute a similar task.

\textit{Conclusion}
The Russian Navy inherited a difficult predicament with the collapse of the USSR; it was left with a massive fleet that was difficult and expensive to maintain, with vessels that were not fit to counter the challenges of the 21st century. The improvement of Russia’s economic situation enabled the Navy to put into motion plans to develop a large surface navy that could project power far from Russian shores; however, the invasion of Crimea and the following sanctions regime killed the plans for most procurement projects.

The future of the Russian Navy will continue to be affected by the consequences of Russia’s aggression as the country will likely struggle to access the necessary technology and find the funds to expand its fleet. The preliminary success of the early projects of the Navy was only possible due to the country’s access to the global markets where it could purchase technology and expertise to modernize its poor industrial base. Engaging in belligerent behavior across Europe and the world resulted in the exclusion of Russia from the global markets; therefore, the political standing of the naval force to support strategic priorities appears to be second behind the other branches. Nuclear deterrence will remain at the centerpiece of the Russian Navy. However, Russia’s desire to expand its power projection capabilities will require innovative stop-gap solutions that combine its existing fleet with new capabilities. These solutions will only enable Russia to project limited levels of power in regions where the Russian military has a presence.

About the Author: Bora Balçay is a student in the Masters in German and European Studies Program at the School of Foreign Service at Georgetown University. His focus is on Transatlantic Security and Naval Strategy in Europe. He is from Istanbul, Turkey.


Miller, “Assessing the Soviet Navy.”


Renz, “Russia’s Military Revival.”


Manaranche, “Russia Postpones Leader Destroyer Construction – Part 1.”


LaGrone, Sam. “‘Doomsday’ Submarine Armed with Nuclear Torpedoes Delivers to Russian Navy.”  


Sharing Secrets: Why Do States Publicly Share Intelligence?

Peter Kerkhof

In 2022, the Biden Administration publicly released an unprecedented amount of intelligence, which primed the world for Russia’s invasion of Ukraine. This paper seeks to examine why states periodically release intelligence publicly. It explores this question by focusing on the decisions that different administrations made in the 2014 and 2022 invasions of Ukraine. Through examining memoirs, interviews, and reporting on the invasions, the paper uncovers three factors that explain the sharing of intelligence as a tool of statecraft. First, Biden’s experience likely translated to a more effective National Security Council. Second, the U.S. Intelligence Community’s (IC) greater focus on Russia before 2022 gave the Biden Administration more options for releasing intelligence. Third, the more advanced open-source community and technology in 2022 amplified the impact of the intelligence that the administration released.

Introduction

States spend vast sums of money to collect intelligence. Agents risk their lives to steal secrets, and scientists develop unique capabilities to divine other states’ intentions. As a result, the repercussions of individuals sharing this information without permission are severe, including potentially compromising agents or rendering expensive technical capabilities useless. For example, after the United States shared satellite photos of India’s nuclear facilities with Indian officials, India was able to disguise its preparation for a nuclear test, thereby robbing the United States of the ability to anticipate the test.1

In some instances, however, leaders accept risk and publicly release intelligence. Why do states take this risk? What reasons or conditions lead a state to risk a future advantage by potentially revealing sources and methods? Further, even if decision-makers think releasing intelligence is worth the trade, how do they overcome resistance from their intelligence communities?

Prior to the 2022 Russian invasion of Ukraine, the Biden Administration released a stunning stream of intelligence, including a satellite imagery-based assessment of the Russian build-up, details of false flag operations, and a brief on Russia’s invasion plan. The administration’s actions did not deter a Russian invasion, but many commentators believe that they “succeeded in revealing Russian plans and aligning major Western powers behind measures to isolate Russia economically and diplomatically.”2 This success may make publicly sharing intelligence a more regular part of the diplomatic toolkit.

This paper will examine two U.S. administrations’ responses to Russian aggression to determine the motivations and factors influencing the decision to share intelligence. Whereas in 2014, the Obama Administration did not share intelligence before the Russian invasion of Ukraine, in 2022, the Biden Administration did. Through an investigation of memoirs, interviews, press releases, and contemporary documents, this paper examines the thought process of each administration and which factors played a part in their decision-making.

Initially, this paper hypothesized that states release intelligence to influence their allies rather than signal to their adversaries. This hypothesis was not supported in the case studies, as officials from both
administrations discussed the importance of messaging multiple audiences, whether information was classified or not. Based on a review of the cases, several other factors instead impacted the Biden Administration’s decision to share intelligence. First, most of the Biden team was in government during 2014 and understood what would unfold in Ukraine in 2022. On top of that, Biden’s experience allowed him to delegate and manage the national security bureaucracy more effectively to release intelligence. Next, the United States likely had better intelligence in 2022 because it prioritized Russia as a threat to U.S. interests. The Russian interference in the 2016 election, among other factors, drove a gradual shift in the U.S. intelligence community’s (IC) collection priorities, which gave the United States more intelligence to consider releasing in 2022. Finally, technology changed dramatically from 2014 to 2022. The Biden Administration found it easier to establish credibility and impact public opinion since the increase in satellites, social media proliferation, and open-source investigators could support the intelligence released.

This paper begins by reviewing the literature on information sharing, secrecy, influence on public opinion, and intelligence sharing between states. Next, the hypothesis, methodology, and definitions section reviews possible motivations for states to release intelligence, what meets the standard of public intelligence sharing, and which viewpoint this study will focus on. An assessment of the strengths and weaknesses of the 2014 and 2022 case studies follows in the case selection section, which provides an overview of available data and a relevant summary of the 2014 and 2022 crises. The analysis and findings section finds that bureaucratic experience, threat levels to U.S. interests, and technological advancements influenced the United States’ decisions to release intelligence. Finally, the paper concludes with recommendations for policymakers, limitations of the study, and areas for future research.

Literature Review

Introduction

This literature review discusses four topics that contribute to why states publicly share intelligence as a tool of statecraft: the impact of sharing information, secrecy, influence on public opinion, and sharing intelligence with other states. First, some researchers found that selectively leaking information can be effective. Second, there are advantages and disadvantages to keeping secrets. Third, intelligence scholars find that U.S. administrations sometimes selectively release intelligence to influence domestic public opinion, which can negatively affect a government’s credibility. Finally, there are obstacles to sharing intelligence with other states related to the transfer of sharing intelligence publicly.

Leaks and Leverage

While not intelligence focused, academics Matthew Castle and Krzysztof Pelc highlight how selectively releasing information can influence an international policy outcome. Specifically, they investigate the cause and effects of leaks in international trade negotiations. By examining over 120 leaks regarding trade negotiations between the European Union (EU) and Canada, they found that leaks effectively mobilized domestic audiences towards a preferred outcome and that most leaks were from actors focused on preserving the status quo.

Additionally, scholars Rochelle Terman and Erik Voeten examine how public information can influence other states, which impacts credibility and norms. Using the United Nations (UN) as a data set, they
find that states subject to name-and-shame campaigns are more likely to reject criticism from their adversaries and accept recommendations from their allies. Related to intelligence release, states are much more likely to trust the intelligence of a friendly state than an adversary, and the subject of the intelligence release is often not the target of the release. Additionally, Terman and Voeten highlight how adversaries are likely to shrug off an opponent’s criticism as an attempt to ruin their reputation. This may translate to intelligence release, given states usually try to influence another actor or set of actors with the secrets they release.

Why States Keep Secrets: Tactical Advantages, Norms, and Public Opinion

States benefit greatly from keeping secrets. Michael Colaresi, a professor at the University of Pittsburgh, argues that secrecy allows democracies two main advantages: anticipation and deception. Anticipation enables states to foresee an opponent’s plans and act accordingly. For example, in 1967, Israel used intelligence to launch a preemptive war and defeat Egypt, Syria, and Jordan in six days. Deception is the ability to hide plans and intentions from an adversary. In WWII, deception was critical for the success of Allied landings on D-Day. While Colaresi highlights the concrete advantages of anticipation and deception, he avoids decision-maker viewpoints.

David Gibbs, a professor at the University of Arizona, uses three models to examine the motivations of government and decision-makers to keep secrets. The first model is the “external threat approach,” which posits that states keep secrets to protect them from other states. This model implies that states will release intelligence when the sources and methods are obsolete or when a foreign government already knows the information. The second model, “bureaucratic politics,” asserts that secrecy results from bureaucratic inertia and interagency competition. According to this model, a state will keep secrets because of standard operating procedures and feuds between agencies. The banality and frustration of the bureaucratic politics model may resonate the most with many practitioners, but Gibbs found the least evidence for it. The final model is the “internal threat approach,” where states “use secrecy as a device to mislead the public and to ensure elite control over foreign policy.”

In this model, leaders selectively release information to influence domestic opinion toward the “right policy options.” Using the United States covert intervention in the Congo as a case study, Gibbs finds that the internal threat model has the most explanatory power for why states release information domestically: to influence public opinion. His argument does not address why states would release intelligence beyond the domestic arena.

Influencing Public Opinion with Intelligence: Always Bad for Business?

Existing literature suggests that many intelligence studies scholars hold negative views on using intelligence to influence public opinion. Robert Mandel, a Lewis & Clark College professor, highlights numerous examples showing the leaking of “bad news” as a common practice by administrations to build domestic support for their policy choices. For example, he highlights how the Clinton Administration released intelligence assessments about refugees from Haiti to build public support for military intervention in the country. Mandel further discusses how intelligence can be politicized from the bottom up. Mandel asserts that analysts politicize intelligence by adjusting their assessments to counteract decision-makers’ preferences. Michael Rubin, a resident scholar at the American Enterprise Institute, also provides evidence for this idea. He highlights the
2007 National Intelligence Estimate (NIE) as an example of how analysts changed standards to give Iran a “clean bill of health” on their nuclear program because of their concerns that the hawkish Bush Administration would attack another Middle Eastern country. However, Rubin’s claim is contentious given that other scholars and the National Intelligence Council Chairman defend the Iranian NIE as the most analytically rigorous ever conducted. Mandel and Rubin both assume that releasing intelligence assessments dramatically impacts public opinion.

Richard Betts, a professor at Columbia University, provides a nuanced discussion on the inevitability of intelligence politicization. He discusses two cases from Vietnam that highlight how a critical policy question became politicized. In each case, the assessment was released publicly to influence public opinion. One of Betts’s conclusions is that the intelligence community must ensure multiple points of view on key issues are represented to avoid inevitable politicization. Further, Betts focuses on how contentious policy debates should be conducted secretly to ensure intelligence is not corrupted to support one side.

Credibility can also impact the effectiveness of publicly releasing intelligence. Jonathan Pinkus, in his paper “Intelligence and Public Diplomacy,” argues that the ultimately disproved allegations of nuclear weapons ahead of the 2003 Iraq War undermined the future ability of the government to influence domestic opinion by sharing intelligence. Pinkus compares the lead-up to the U.S. invasion of Iraq in 2003 to the response to Syrian chemical weapon use in 2014. Pinkus asserts that the politicization of intelligence by the Bush Administration in 2003 undermined public trust in the intelligence community in the United States and the United Kingdom (UK). Therefore, citizens in both countries rejected the countries’ cases for intervening in Syria in 2014. Pinkus makes the case that credibility matters regarding the effectiveness of intelligence sharing by states.

**Intelligence Sharing Comes with Costs**

James Igoe Walsh, an associate professor at the University of North Carolina, establishes a framework for intelligence sharing between nations based on contracts and relative power. Walsh seeks to explain why and how the United States developed intelligence-sharing relationships with certain partners and not others. He applies the economic logic of relational contracting to explain how states overcome barriers to sharing intelligence. As part of his investigation, he examines why states choose not to share intelligence. He identifies three principal barriers: the inability to negotiate an acceptable agreement, passing shared intelligence to unauthorized third parties, and the secret nature of intelligence.

While the first two are inapplicable to public intelligence sharing, for the secret nature of intelligence, Walsh highlights how intelligence is a non-tangible good, meaning it cannot be inspected or validated by other states or the public. In this way, receiving intelligence is like buying a used car without looking at a vehicle history report—it is hard to know whether you are buying a lemon. Walsh’s model provides a valuable understanding of intelligence sharing. However, the world has changed significantly since his analysis, making his conclusions less applicable. Many tools states have to collect intelligence are now widely available, and open-source investigators can now validate intelligence in ways that were not previously possible.
This literature review highlights a gap in why states would share intelligence as a tool of statecraft. Scholars document how releasing information can impact negotiations and influence allies more than adversaries. Intelligence scholars highlight the concrete advantages of secrecy and how it can protect decision-makers. Many assume that intelligence releases can impact public opinion, but Pinkus highlights how damaged credibility decreases the impact of shared intelligence. Walsh highlights the obstacles to sharing intelligence with other states. However, the world has changed significantly since most of this research was published. Therefore, this paper will examine states’ motivations to share intelligence as a tool of statecraft and evaluate how they may have evolved.

Hypothesis, Methodology, and Definitions

To test the hypothesis that states release intelligence to influence their allies or partners, this paper compares two instances during the ongoing war in Ukraine. Although the foreign policy challenges are similar, the United States only released intelligence in one instance. Through examining the U.S. responses to the 2014 and 2022 Russian invasions of Ukraine, the paper will contrast decision-making scenarios and identify the causes for why the administrations in power made different decisions about releasing intelligence. To do this, this paper will examine each crisis and then determine why leaders made specific choices, using memoirs and interviews as key sources of data.

*Intelligence* includes secrets that are collected by spies, communications intercepts, and satellites, and the analysis of all those secrets mixed with open-source information.15

*Publicly released intelligence* is intentionally released, not leaked. This study determines why states use intelligence to achieve foreign policy goals, so unintentional leaks—such as Wikileaks—are out of purview.

Case Selection and Data Collection

This section will examine the U.S. responses to the Russian annexation of Crimea in 2014 and the lead-up to the Russian full-scale invasion of Ukraine in 2022. These two cases have strengths and weaknesses. They are a good comparison because they share similar characteristics: level of aggression, region, U.S. administration composition, and Russian leadership (i.e., President Vladimir Putin).

Most players involved in the response to both Ukraine conflicts were similar, from individuals to international organizations. Both conflicts were about Ukraine’s future as either Eastern or Western facing— involving institutions like the EU, the North Atlantic Treaty Organization (NATO), the G7, and the UN. Both events have the same three primary state players: the United States, Russia, and Ukraine. The individuals involved are also similar. Vladimir Putin was the instigator of both crises, and most of the officials in the Biden Administration played vital roles in both 2014 and 2022. Finally, in both 2014 and 2022, the Russians tried to take Ukrainian territory.

Both cases have good data on leaders’ perspectives. Many Obama administration officials wrote memoirs about 2014, including the National Security Advisor, CIA Director, and UN ambassador. Additionally, *Politico* conducted in-depth interviews about the administration’s thought process in 2022, including quotes from the same players (see Table 1 for a complete list of sources).

It is worth noting potential weaknesses in comparing the two conflicts, including differences in the size and timing of the
Russian operations, as well as differences in threat prioritization. Some argue that the 2014 Russian annexation of Crimea was an intelligence failure. The Obama Administration was surprised by the invasion of Crimea, such that the United States was only able to respond after the fact. The Russian operation was conducted exceptionally quickly with a much smaller force than in 2022. The contrast between the Obama Administration releasing intelligence after the downing of MH17 – five months after the annexation – and the Biden Administration’s efforts in 2022 emphasizes how the administrations leveraged the public sharing of intelligence in different ways.

In 2014, the U.S. government did not prioritize Russia as a threat. The United States was still actively negotiating with Russia over several issues, such as the Joint Comprehensive Plan of Action (JCPOA). What is notable from the viewpoints of National Security Advisor Susan Rice and Deputy National Security Advisor Ben Rhodes is the number of phone calls between Putin and Obama. John Brennan, the CIA director, also visited Moscow to meet with his Federal Security Service of the Russian Federation (FSB) counterparts and even invited them to the United States in November 2014—after the annexation of Crimea, the downing of Malaysian Airlines Flight 17 (MH17), and concurrent unrest in eastern Ukraine. The Annexation of Crimea in 2014

In late November 2013, Ukrainian President Victor Yanukovych canceled talks on a trade association with the EU in favor of closer ties to the Russian Federation. The decision unleashed a wave of protests across Ukraine that escalated over the next several months. On February 22, 2014, Yanukovych fled Ukraine while Russian special operations forces prepared to seize Crimea. Numerous social media reports showed Russian airborne troops and special forces leaving their bases around Moscow on the 22nd and 23rd. Local Russian media outlets quickly reported news of the mobilizations.

On February 26, Vladimir Putin declared a snap exercise in the western military district. Unidentified gunmen seized government buildings in Crimea the next day. Despite initial ambiguity about the identities of the gunmen, it became increasingly clear that they were elite Russian soldiers. Putin and Russian spokespersons denied that there were any Russian troops in Crimea. Occupied Crimea quickly held a referendum to join Russia, and the Russian parliament voted to absorb Crimea on March 18, 2014.

The United States did not publicly release intelligence immediately before the annexation or as a response in the following month. There are two potential explanations: either the United States decided there would be no benefit, or it did not have the relevant intelligence to share.

The Russian annexation of Crimea was well-timed and quickly executed, while the Ukrainian government was in shambles in late February and early March. Even if the United States had foreknowledge of the event, publicly releasing intelligence may not have been enough for the fractured Ukrainian government to increase military readiness and rebuff the Russian operation. More likely, however, the United States did not have a clear intelligence picture of what Russia was about to do in Crimea. The memoirs from Obama Administration officials do not mention Ukraine or Crimea prior to the annexation. It seems likely that it was an intelligence failure. James Clapper, the Director of National Intelligence, mentions that despite warning signs, the annexation was a surprise: “In fact, we’d been warning for several days in February about Russian soldiers without insignia...
positioning themselves around Crimea and Russian troops massing near the border, but we never expected Russia to actually seize control, much less formally annex the peninsula.” Clapper’s comments suggest that the intelligence picture of Russian intentions was unclear at best.

After the annexation of Crimea, the United States response focused on sanctions, which took several weeks to enact and were targeted at specific Russian leaders, limiting their impact. The Obama Administration spent significant time and effort calling European leaders and coordinating a response following the invasion. Russia’s consistent denials of activity further muddied international perceptions of what happened. If the United States had released intelligence after the annexation, it may have increased the effectiveness of the international response by providing a clear and compelling picture of what happened in Crimea.

**The Delayed Response: MH17**

In July 2014, four months after the invasion, the Obama Administration finally released intelligence about Russian involvement in Eastern Ukraine after Russian-controlled forces shot down MH17 and killed 300 civilians. Russia predictably denied any involvement. The following day, Obama made a circumstantial case that Russia was responsible for shooting down the airliner, stating how “evidence indicates that the plane was shot down by a surface-to-air missile that was launched from an area that is controlled by Russian-backed separatists inside of Ukraine” and “ever the last several weeks, Russian-backed separatists have shot down a Ukrainian transport plane.” In his speech, Obama did not invoke intelligence or provide further specific details about the incident. A week later, however, in response to a Russian press conference about MH17, the State Department released satellite imagery and analysis, which showed that Russian artillery was firing from Russia into Ukraine. The presentation was underwhelming. Open-source investigator and Bellingcat founder Elliot Higgins described the release: “US officials merely tweeted a satellite image with the supposed launch site of the missile. This poor-quality graphic cited no source and earned ridicule online.”

Rhodes and Clapper offer conflicting views about this intelligence release. First, Clapper’s description of the U.S. intelligence related to the downing of MH17 highlights the IC’s progress: “We had the Russian’s dead to rights in just a few hours, fusing data from our so-called ‘national technical means satellites,’ intercepts, and open-source reporting—particularly social media.” and exclaimed that the U.S. IC “spoke intelligence truth to the world.” By contrast, Rhodes has described his efforts as trying to single-handedly counteract Russian disinformation. He recounts the lackluster results of the intelligence release:

A little over a week after MH17, we got the intelligence community to declassify overhead imagery that showed Russian military equipment pouring over the Ukrainian border, but it was still framed by objective international media outlets as a “he said, she said” story, and Russia just issued more denials in response.

The contrasting viewpoints of Rhodes and Clapper about the publicly released intelligence highlight the gap between what the U.S. government knew and what was released. Although it was not publicly released, Clapper cites precise intelligence that confirmed within hours, through multiple intelligence sources, that Russian forces shot down the planes. Conversely, Rhodes laments that the public received some imagery weeks later that was only
tangentially related to the shootdown. Based on their comments, the IC had much more compelling information that could have made a better case to the international community. Still, the administration decided not to release it—possibly due to bureaucratic friction or an assessment that the sources were not worth risking.

In his memoir, Rhodes laments that his efforts to counter Russian propaganda were inadequate, especially considering his team was only five people. He also felt the restrictions of classification: “Because we had to be fact-based [and] because we had to be mindful of sensitive intelligence, we were slower than the Russians, and could not be as definitive in our statements or as far-reaching in our social media presence.”

Despite the invasion of Crimea and eastern Ukraine, and the death of 300 civilians, the National Security Council could not overcome the friction of classification to release more intelligence, which would have built a better case for the international community to respond.

In the aftermath of the conflict, in memoirs, administration officials spent significant time discussing the events leading up to the election in 2016. Clapper points out the connection between how the Russian propaganda machine supported the invasion of Crimea in 2014 and its later use against the United States in 2016. He identifies how Russian propagandists used the “new” technology of the Internet to sow doubt in the international community. It seems that the Russian interference in the 2016 election was the decisive point for the administration to prioritize Russia as a threat—a few years too late.

Russia’s 2022 Invasion of Ukraine

The Biden Administration suspected Russia was planning to invade Ukraine in October 2021. General Paul Nakasone, the National Security Agency director, stated, “by the 11th of October, I’m convinced the Russians are going to invade Ukraine. The preponderance of intelligence was different than anything we’d ever seen before.”

Over the next month, the Biden Administration shared information privately with allies and communicated directly with Russia. Biden talked with the prime ministers of France, Germany, and the UK at the G20 Summit in October, and Secretary of State Antony Blinken met with Ukrainian President Volodymyr Zelensky. CIA Director Bill Burns traveled to Moscow to warn the Russians that the United States knew what they were doing. The Director of National Intelligence, Avril Haines, traveled to Brussels to brief the North Atlantic Council—NATO’s leadership body—on the U.S. assessment. The next step was to go public.

On December 3, 2021, the administration released a one-page intelligence product that showed satellite imagery of the Russian forces on the border of Ukraine. The document showed satellite images of recent Russian troop movements and assessed the size of the troop buildup. The document focused on Russia’s capabilities with very granular detail but did not assess Russia’s intentions. Biden and Putin exchanged phone calls and agreed to talks in January. However, the talks went poorly, and Russian representatives continued to deny they had any intentions to invade Ukraine. Over the next several months, the Biden Administration would make a series of statements—based on intelligence—to prove Russian intentions to invade.

On January 14, 2021, White House spokesperson Jen Psaki and Defense Department spokesperson John Kirby both stated at press conferences that the United States had intelligence that Russia was preparing to fabricate a pretext to invade
Ukraine. The press statements were not accompanied by hard evidence, such as the satellite images from December or recorded phone calls— and Russia quickly denied the allegations.

Next, the UK released a memorandum that detailed Russia’s plan to form a new government after deposing Zelensky— including the names of Ukrainians to lead it. On February 3, State Department spokesperson Ned Price went a step further than Psaki and Kirby by conveying that Russian agents were making a fake video as a pretext for invasion. During the press conference, a reporter aggressively questioned Price about his evidence.

Their exchange highlights several of the challenges with releasing intelligence. First, the reporter rightly identified that the U.S. government offered no proof that Russians were preparing a video:

No, it’s an action that you say that they have taken, but you have shown no evidence to confirm that. And I’m going to get to the next question here, which is: What is the evidence that they – I mean, this is – like, crisis actors? Really? This is like Alex Jones territory you’re getting into now. What evidence do you have to support the idea that there is some propaganda film in the making?

There were no audio recordings, images, or tangible evidence for an observer to independently evaluate. In that sense, it was hard to accept the U.S. government’s assertions. Price argued that the United States would not release the details to back up his claim because doing so would compromise the sources that allowed the United States to collect the information. In the exchange, Price rests his case on the credibility of the U.S. government and Russian behavior since 2014. This exchange highlights the interplay between state credibility, past history, and protecting intelligence sources. Two weeks later, Russia released a blatantly staged video of a car bomb in Donetsk.

On February 17, 2022, Blinken delivered a short speech at the UN Security Council about Russia’s imminent invasion of Ukraine. He detailed the steps of the Russian plan: mobilizing hundreds of thousands of soldiers on the border, manufacturing a pretext for invasion, holding emergency meetings in Russia to authorize intervention, conducting cyber-attacks, and bombarding Ukrainian cities, followed by tanks and soldiers crossing the border, and kill lists of specific Ukrainians. Even after its plans were exposed, Russia proceeded with a full-scale invasion of Ukraine on February 24.

Data Analysis

In 2014, the United States did not release intelligence before the invasion or the following month. In 2022, by contrast, the United States took a much more proactive approach and released various intelligence products. What is different between the two cases, and what does the difference reveal about a state’s motivations to release intelligence?

Initially, I hypothesized that states were focused on releasing information to influence their allies and not their adversaries. The results of the case study comparison are inconclusive with regards to this hypothesis. In both cases, officials said their public messaging targeted multiple audiences. In 2014, then-Ambassador to the UN Samantha Power gave several speeches to the UN Security Council denouncing Russia’s actions in Crimea. She did not reference classified intelligence but targeted her message to both allies and the Kremlin.
In 2022, administration officials discussed how releasing intelligence would hopefully both convince allies and deter the Kremlin.

In 2022, the Biden Administration was clearly trying to alert allies of the potential Russian invasion. In November 2021, Michael Carpenter—the U.S. Ambassador to the Organization for Security Cooperation in Europe (OSCE)—was concerned that European allies were not convinced of the gravity of the situation. Liz Truss—at the time the UK Foreign Minister—was also concerned that other NATO members did not share the view that Russia was planning to invade. Many U.S. decisions on releasing information were evidently driven by a desire to convince allies about the imminence of a Russian invasion. Biden’s Deputy National Security Advisor Jon Finer even said, “We eventually brought people around by bombarding them with information that you could not ignore.”

As shown by Haines’s brief to the North Atlantic Council in November 2021, sharing classified intelligence was part of that bombardment.

The Biden Administration likely went public in December due to the intransigence of specific partners. UK Ambassador to the United States, Karen Pierce, reinforced this point vis-à-vis NATO allies when she said: “We knew that the French and Germans had the same reports that we had. We were puzzled by their insistence that he would not invade.”

Amanda Sloat, former National Security Council (NSC) Senior Director for Europe, had a similar view: “It got to the point where we had to say to the Europeans, “Fine, we can agree to disagree analytically, but let’s start planning as if we are right.” Sloat’s view, coupled with those of other U.S. administration officials, highlights the pushback about the invasion from allied governments. The shift from the private intelligence sharing campaign in October and November to the public campaign in December, January, and February was likely a response to that allied intransigence.

The release of intelligence was also clearly aimed at Russia. When Price briefed intelligence about the Russian video, he hoped that merely discussing the video would stop the Russians from employing that tactic. Many other administration officials echoed this sentiment about disrupting Russian plans. For example, Kirby said, “We were beating Putin’s lie to the punch, and we know that by doing so, we got inside his decision-making loop.” Burns concurred, “I do think it made a difference, not only in putting Putin on his back foot, but also in shoring up the solidarity.”

Blinken credited his speech at the UN with delaying Putin’s invasion by a week – a commonly held position within the Biden Administration.

In sum, publicly releasing intelligence served two purposes. First, the Biden Administration wanted to convince and pressure allies to be more proactive by “bombarding them with information.” Second, the administration wanted to dissuade or at least disrupt Russian plans to invade. Evidently, states conduct messaging campaigns to influence adversaries, allied governments, and allied populations—intelligence released publicly reaches all groups. Several other interesting factors differed between the two cases and facilitated the release of intelligence: the Biden Administration’s experience, the increased Russian threat, and technological changes.

The Effect of Experience: Building a Trusted Team

Experienced leaders are more likely to release intelligence publicly because they can effectively manage their national security bureaucracy. This result is likely
generalizable to future administrations and possibly other countries. Leaders that do not have secure bases of support or are concerned about the support of their national security apparatuses will be unwilling to override the security establishment’s concerns about classification and release intelligence to the world.

Interviews of Biden officials highlighted that their prior experiences in the Obama Administration dramatically affected their response to the full-scale invasion of Ukraine in 2022, including their use of intelligence. Sullivan notes how he “was working in the White House when Crimea unfolded and the ‘little green men,’ the early hours of confusion and fog of war. We had the benefit of being able to learn from that experience.”\footnote{50} Jake Sullivan held several positions in the Obama Administration, including as the Vice President’s National Security Advisor.\footnote{51} Likewise, Victoria Nuland, Under Secretary of State for Political Affairs for Biden and Assistant Secretary of State for European and Eurasian Affairs under Obama\footnote{52} notes how “A lot of us were veterans of 2014, ’15, and ’16, and felt that if we had done more faster then to help Ukraine, we might have had a better result.”\footnote{53} It is common for political appointees to work with multiple administrations, but notable that recent experience with Russian in 2014 carried through numerous administration officials.

However, experience level is best examined at the Presidential level. The gap in foreign policy experience between Biden and Obama may explain why the Biden Administration chose to release intelligence. A study conducted by Elizabeth Saunders, a professor at Georgetown University, highlights how foreign policy experience can impact decision-making. Specifically, Saunders compares the experience of the two Bush Administrations before the Persian Gulf War and the 2003 Iraq War, both of whom used many of the same national security officials.\footnote{54} Whereas the Persian Gulf War is widely recognized as a successful application of American power, the 2003 Iraq War is often regarded as one of the United States’ most significant foreign policy mistakes. Saunders highlights how H.W. Bush’s significant foreign policy experience allowed him to delegate and supervise his national security team more effectively.

It is clear that Obama’s lack of experience and intelligence together contributed to an inefficient and hesitant NSC, embodying signs of ineffective delegation. As former Obama National Security Advisor Susan Rice notes, “Obama’s queries, taskings, and decision-making had the effect of reminding all involved that, for better or worse, he was consistently the smartest guy in the room.”\footnote{55} Being the smartest guy in the room is great, but it can have a downside. The Obama Administration’s NSC was notorious for their long and painful decision-making process. Rice sums up the issues in her memoir: “1) the Principals Committee met too frequently; 2) meetings went too long; papers were late; and discussions went too much in the weeds; 3) we got too deeply into the agencies knickers; and 4) we were overly directive.”\footnote{56} Despite his lack of prior foreign policy experience, Obama was extensively involved in foreign policy as evidenced by his dictating and editing of the surge orders himself after the extended Afghanistan review process.\footnote{57} Saunders would describe this as his inability to effectively delegate—a pitfall that she notes is common among inexperienced principals. Leaders without foreign policy experience—and therefore those who do not effectively delegate—might be more unwilling to take large risks, such as sharing intelligence. One prerequisite of publicly sharing intelligence
is deciding that risking sources and methods is worth advancing some political goal. Intelligence officials will almost uniformly oppose releasing intelligence because it cuts against the spirit of their profession. That political risk invites blowback from political opponents and the intelligence services.

Obama also struggled to navigate the bureaucracy and build a trusted team—a vital aspect of a president’s foreign policy success. In his book *Presidential Command*, Peter Rodman discusses NSC performance from Truman to the second Bush Administration extensively, focusing on how presidents control the national security bureaucracy. One of his main observations is that cabinet secretaries need to be “strong and loyal.” Specifically, they need to be “trusted by the president to carry out policies in harmony with the president’s wishes and... see themselves as the president’s agent in the department, not the spokesman of the department.”

Obama was new to Washington and did not have a strong relationship with his officials. Rhodes described Obama’s relationship with his cabinet, stating, “Most of the people who filled the top positions at Obama’s State Department or Pentagon were people he had never actually met.” Obama was frequently at odds with his appointees, which likely drove his over-involvement to ensure agencies executed his policies. During Obama’s Afghanistan review, he felt that his military advisors were trying to limit his options through leaks and public statements. Obama fired his Director of National Intelligence, Dennis Blair, for picking too many fights with the CIA. In an NSC meeting on the Libyan intervention, Obama asked all of the principals to state their opinions and then—in frustration—asked for the views of all of the backbenchers. Robert Gates, Obama’s Secretary of Defense, said the Obama national security staff “took micromanagement and operational meddling to a new level.” Obama compensated for disagreement among appointees in his administration by overly involving himself in the NSC.

By contrast, Biden has extensive experience with foreign policy and navigating the bureaucracy, which enables him to delegate more effectively. He is the oldest U.S. president, served in the Senate for over three decades, and chaired the Senate Foreign Relations Committee for almost a decade. In addition, he served as Vice President for two terms under Obama. As Vice President, Biden played an active role in foreign policy issues. A few examples illustrate the contrast between Biden and Obama. During Obama’s review of Afghanistan, then Vice President Biden was intimately involved in strategy discussions and challenged the "surge" widely favored by other cabinet members. After the invasion of Crimea, Biden took on the primary role of communicating with Ukraine’s leaders.

Biden’s experience has enabled him to build a trusted team of advisors who can implement his foreign policy agenda. Biden has worked with almost all of his key appointees in the Senate or as Vice President. Blinken was the Democratic staff director of the Senate Foreign Relations Committee, where Biden was the senior Democrat. Blinken continued to work for Biden as Vice President. Next, Sullivan was Vice President Biden’s National Security Advisor. Burns worked on the JCPOA with Sullivan during the Obama Administration. Haines served as an Obama NSC advisor and Obama-era Deputy Director of the CIA.

The 2022 Ukraine crisis demonstrated Biden’s effective delegation of authority. Sullivan convened a small group focusing on Ukraine that met daily at the start of November 2021. That small group devised...
the idea to release intelligence, and then Biden approved it in early November. With that clear guidance, Sullivan, Burns, and Haines aligned the bureaucracy and created a system that could effectively declassify and release relevant intelligence over the next four months before the Russian invasion. This execution stands in stark contrast to Rhodes in 2014, who struggled with the IC to release some grainy, partially-relevant satellite photos after the downing of MH17. Biden effectively appointed the right people to the right seats to overcome bureaucratic friction and release intelligence supporting his foreign policy.

Russia as a Threat: Releasing Intelligence Requires You to Have It

The U.S. IC had better intelligence on Russia’s intentions in 2022 than in 2014, which impacted the Biden Administration’s decision to release intelligence. Collecting intelligence is dangerous and expensive, making it impossible to gather information everywhere. Furthermore, national security threats are evaluated based on intent and capability, the former of which is much harder to track. Simply put, intent is an actor’s desire to conduct an attack, and capability is an actor’s ability to execute the attack. Threats are identified as actors with both intent and capability.

Even with the full subpoena power of a congressional committee, the public cannot understand what the IC knew and did not know. Based on the memoirs and documents examined for this paper, the Obama Administration did not have clear intelligence that Russia would invade Ukraine in 2014. Before the invasion of Ukraine in 2014, Russia was not considered one of the United States’ top threats. In Obama’s first National Security Strategy (NSS), Russia was referred to as a key center of influence, and a goal for the United States was to build a strong and multi-dimensional relationship with Russia. At that time, the United States sought to cooperate with Russia on multiple issues: reducing nuclear proliferation, countering violent extremism, and creating the JCPOA with Iran. During that time, Brennan visited Moscow to meet with his Russian counterpart multiple times at the CIA. After the Russian invasion of Crimea, the cordial nature of the U.S.-Russian relationship changed. The tone of the second Obama NSS marked a dramatic shift in focus vis-à-vis Russia. Instead of seeking to build a strong multi-dimensional relationship with Russia, the United States was “mobilized and is leading global efforts to impose costs to counter Russian aggression.” The emphasis on Russia as a threat continued through the Trump and Biden Administration National Security Strategies, which has impacted how the IC sets its priorities.

Public announcements of the CIA’s changing priorities highlight how the IC’s focus shifted with the change in national focus. In 2015, Brennan announced that the CIA would create 10 new mission centers to better fuse intelligence and operations. Three years later, former CIA director Gina Haspel announced that the CIA would return to its roots and focus on nation-state rivals. In 2021, Burns announced that the CIA would create a new mission center for China to continue to focus on strategic competition with great powers. These progressive announcements illustrate how changing national priorities affect intelligence priority collection and suggest that the U.S. intelligence collection efforts on Russia were lacking in 2014.

This change of focus painted a clearer picture of Russia’s capabilities and intentions in 2022. Unlike in 2014, the Biden Administration had enough intelligence to convey both Russia’s intent...
and capability to invade Ukraine. The details on Russia’s intent were critical to convince skeptics that Russian troop movements were more than just an exercise. Additionally, the administration and IC had enough intelligence on Russian intent that they were able to release some without gravely compromising sources and methods. If a state does not have enough details to make a compelling picture, they are unlikely to release intelligence.

**Technology: Unleash the Open-Source Investigators**

One of the most surprising elements of the 2022 war in Ukraine is the amount of granular information that is widely available to the public. Anyone with access to the Internet can log in to Twitter or Telegram and access high-quality intelligence assessments that are remarkably similar to what senior U.S. military commanders receive. This marks a significant change in who has access to intelligence and information – something officials in the Biden Administration have noted.

In *Spies, Lies and Algorithms*, Amy Zegart describes the explosion in the quality and quantity of information available to the general public as the “democratization of intelligence.” Zegart focuses on nuclear threat intelligence as a “tough test” of how emerging technologies fundamentally challenge U.S. intelligence. She found three trends driving democratization: increased quantity and quality of satellite imagery, increased connectivity, and automated analysis. These factors contributed to the reality that “several non-governmental nuclear intelligence groups have amassed a breadth of expertise that rivals or, in some cases, exceeds the capabilities of the U.S. Intelligence Community.” These factors were beginning in 2014 and were in full force by 2022.

The growth of the open-source firm Bellingcat captures the changing environment. Elliot Higgins started Bellingcat three days before the downing of MH17 in July 2014. He quickly gathered a global eclectic group of volunteers to investigate the downing of MH17. Using open-source techniques, his investigators used satellite imagery and social media to identify who shot down MH17. In a few weeks, Higgins and his team tracked down the location of the Russian missile launcher when it fired, the Russian unit that operated it, and assigned responsibility to the Russian Armed Forces. They used a mix of openly available information, dashcam videos from YouTube, satellite images for geo-location, a Russian license plate database, and social media posts from Russian soldiers. MH17 was Bellingcat’s first investigation. The firm continued to grow, build, and train open-source investigators around the world. As open-source researchers gained capability, they changed how government claims and publicly shared intelligence were received.

By 2022, the information environment had shifted. As discussed previously, the Biden Administration was aware of the Russian playbook and better prepared for the changed information environment. When NSC spokesperson Emily Horne discussed the Biden Administration’s decision to release intelligence, she highlighted three critical factors that gave them the confidence to proceed: satellite imagery, citizen journalists with smartphones, and an increased public understanding of misinformation and disinformation. Horne’s description of the information environment is starkly different from that of Rhodes, who in 2014 was “going alone” trying to fight Russian disinformation with a single Twitter account. By 2022, open-source investigators were actively countering Russian disinformation in the news media and therefore adding weight to
the Biden Administration’s claims by verifying information.

The open-source industry served as a feedback loop for the Biden Administration. The Biden Administration publicly released intelligence—which was greeted with skepticism—and then open-source investigators would vet their assertions and discover more details. This played out with tracking the troop build-up and Russian provocations. The administration released satellite imagery once, on December 3, 2021. After that, Maxar, a satellite imagery company, released numerous images with exquisite detail. On February 9, 2022, Maxar released images of helicopters at staging bases close to the Ukrainian border and a field hospital.\(^{82}\) Maxar then identified specific weapons systems such as the Iskander theater ballistic missile and SU-25 ground attack aircraft. That precise information was paired with social media videos that showed the progress of Russian military convoys toward Ukraine.\(^{83}\)

Next, the administration warned about a fake video that would be used to instigate the conflict. A few weeks later, several videos appeared. They were treated with ready skepticism by news media outside of Russia. Researchers quickly tore apart the blatant fabrications. A man whose leg was “blown off” already had a prosthetic. The metadata was from weeks ago. The gun noise was taken from a ten-year-old video.\(^{84}\)

Due to the Biden Administration’s decision to warn by releasing intelligence preemptively, the videos had little impact, unlike Russian claims in 2014. Better yet, the administration did not have to counter Russian disinformation in real time.

The change in technology and the corresponding democratization of intelligence increased the returns on releasing accurate intelligence. Now, when a government releases accurate intelligence, the vast array of open-source investigators will do their best to confirm or deny it.

While the Obama Administration released intelligence with little impact, the Biden Administration released intelligence that primed a well-developed open-source community. Despite some rocky press conferences and skeptical reception in January 2022, open-source evidence continued to support what the administration was saying and increased the impact of its allegations. The administration could count on the open-source community to fill in the gaps—usually faster than the government.

**Conclusion: Limitations, Future Research, and Policy Implications**

This paper investigated why states release intelligence publicly. Initially, the paper hypothesized that states release intelligence primarily to influence their allies. This would allow states to build stronger coalitions or convince allies to support a specific cause. Based on examining Russian attacks on Ukraine in 2014 and 2022, however, that hypothesis was not comprehensive. In 2022, the choice of audience did not drive administration officials to release intelligence publicly. Administration officials were focused on both Russia and European allies. Instead, three factors stood out when comparing the two different invasions. First, Biden’s experience contributed to a more cohesive national security team, which, in this instance, likely contributed to releasing intelligence publicly. Second, the U.S. IC simply had more intelligence to release due to the shift in collection priorities. Third, the Biden Administration benefited from the expansion of open-source research and technological change between 2014 and 2022.

Studying intelligence has inherent limitations. Due to classification, it is hard to assess who knew what, when, and how
any changes would play out differently based solely on publicly available information. This paper assumes that decision-makers choose to risk exposing sources and methods, but it is hard to determine how much risk decision-makers assume. The Biden Administration may have had so many sources of intelligence that confirmed what they released that there was limited risk to any one source or method. Similarly, it is difficult to assess what the cost of releasing intelligence was.

The next limitation is effectiveness. This study purposefully chose not to focus on the effectiveness of the intelligence releases. The intelligence that the administration released may have shaped the narrative and created a more effective response, but it may have exposed sources and methods that would give the United States foreknowledge of a critical future event. It is also extremely hard to measure the effect of “shaping the narrative.” From the accounts of decision-makers and Western media consulted for this study, however, it does appear that the Biden Administration successfully shaped the narrative in the Western market.

There are several avenues for future research. This study identified the open-source community as a factor in intelligence release. It would be worth examining how publicly releasing intelligence shapes the news media and open-source community interaction, which may present an opportunity to counter authoritarian disinformation efforts. Additionally, existing literature on releasing intelligence is focused on the downsides of politicization and impact of credibility. This is worth re-investigation because the change in the information environment has made government claims more verifiable and secrets harder to keep.

There are several implications for policymakers. First, publicly releasing intelligence is likely not going away. The decision-makers used intelligence to both signal to Russia and convince European allies. It is easy to imagine a similar situation playing out in the Indo-Pacific. The United States will have to both deter the People’s Republic of China (PRC) while bolstering resolve among a coalition of Pacific nations. While researching this paper, the Biden Administration released more intelligence about PRC intentions—hoping to dissuade the PRC from sending lethal aid to Russia.85

With that in mind, the U.S. IC should begin to look into collecting or analyzing intelligence with the intent to release it. This will require a careful assessment of sources and methods that are not easily compromised or can be easily recovered. There is also potential for increased informal or formal partnerships with the open-source community. U.S. IC leaders should seek opportunities to steer the open-source community toward appropriate problems. Taiwan is a great example. There are legions of open-source investigators who track the People’s Liberation Army (PLA) capacity to invade Taiwan closely, such as the daily movements of their amphibious landing ships. Under the public eye, it will be extremely difficult for the PLA to invade Taiwan without warning from the open-source community.

However, the rise of open-source cuts both ways. Firms like Bellingcat will turn on the United States if they suspect the U.S. of hypocrisy. Investigators can just as easily expose the United States’ secrets as it can those of Russia or the PRC. Finally, intelligence is a continuous competition between those trying to discover secrets and those trying to hide them. Releasing intelligence publicly may have unintentional long-term effects that dampen the U.S. IC’s ability to collect intelligence and inform...
decision-makers. The U.S. should proceed cautiously.

About the Author: Peter Kerkhof is an active-duty army officer with experience in the Middle East, Europe, and Asia. He received an M.A. in Security Studies as a General Wayne Downing Scholar from Georgetown University (‘23). He also received a B.S. in Mathematics and Economics from the United States Military Academy (‘11).
### Table 1. Administration Officials in 2014 and 2022

<table>
<thead>
<tr>
<th>Role</th>
<th>2014</th>
<th>2022</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSA</td>
<td>Susan Rice</td>
<td>Jake Sullivan</td>
</tr>
<tr>
<td>CIA Director</td>
<td>John Brennan</td>
<td>Bill Burns</td>
</tr>
<tr>
<td>NSA-Deputy</td>
<td>Ben Rhodes</td>
<td>John Finer</td>
</tr>
<tr>
<td>UN Ambassador</td>
<td>Samantha Powers</td>
<td>Linda Thomas Greenfield</td>
</tr>
<tr>
<td>DNI</td>
<td>James Clapper</td>
<td>Avril Hanes</td>
</tr>
<tr>
<td>State Department</td>
<td>John Kerry*</td>
<td>Antony Blinken</td>
</tr>
<tr>
<td>JCS Chairman</td>
<td>Martin Dempsey*</td>
<td>Mark Milley</td>
</tr>
<tr>
<td>NSA Director</td>
<td>Keith Alexander to Michael Rogers*</td>
<td>Paul Nakasone</td>
</tr>
<tr>
<td>DIA Director</td>
<td>Michael Flynn*</td>
<td>Scott Berrier</td>
</tr>
</tbody>
</table>

* Comments not available or not used for this study

### Table 2. Intelligence Releases by the Biden Administration prior to Ukraine 2022

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
<th>Type of Intelligence Release</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec. 3, 2021</td>
<td>Biden Administration releases intelligence assessment with satellite imagery</td>
<td>Satellite Imagery</td>
</tr>
<tr>
<td>Jan. 14, 2022</td>
<td>Psaki and Kirby say that U.S. intelligence indicates that Russia is preparing to fabricate a pretext for invading Ukraine</td>
<td>Press Briefing</td>
</tr>
<tr>
<td>Feb. 3, 2022</td>
<td>Price accuses Russia of making a fake video showing an attack on Ukraine</td>
<td>Press Briefing</td>
</tr>
<tr>
<td>Feb. 16, 2022</td>
<td>Senior administration officials say intelligence does not indicate any pull back</td>
<td>Press Briefing</td>
</tr>
<tr>
<td>Feb. 17, 2022</td>
<td>Blinken lays out invasion plan at the UN Security Council, saying “Putin has decided to go to war”</td>
<td>UNSC Speech</td>
</tr>
</tbody>
</table>


7 Gibbs, 214.


14 Walsh, 7–13.


18 Howard and Pukhov, 160.

19 Howard and Pukhov, 160.


25 Clapper and Brown, Facts and Fears, 266.

26 Clapper and Brown, 266.

27 Rhodes, 381.

28 Rhodes, 382.

29 Clapper and Brown, Facts and Fears, 265.
39 “Department Press Briefing – February 3, 2022.”
45 Banco et al, “Something Was Badly Wrong”, Politico.
46 “Department Press Briefing – February 3, 2022.”
49 Banco et al, “Something Was Badly Wrong”, Politico.
50 Banco et al, “Something Was Badly Wrong”, Politico.
56 Rice, 424.
An official publication of the Center for Security Studies at Georgetown University’s Edmund A. Walsh School of Foreign Service