GEORGETOWN SECURITY STUDIES REVIEW

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

EDITORIAL STAFF
Freddie Mallinson, Editor-in-Chief
Kara Joyce, Deputy Editor
Caroline Hecht, Associate Editor for Sub-Saharan Africa
Ona Spreenberg, Associate Editor for the Western Hemisphere
Arjan van Tongerlo, Associate Editor for Europe & Central Asia
Sojung Ha, Associate Editor for Gender
Keven Hernandez, Associate Editor for the Indo-Pacific
Sandra Marcushamer, Associate Editor for the Middle East & North Africa
Simon Handler, Associate Editor for Defense
Nathaniel Rome, Associate Editor for Emerging Technology
Becca Ebert, Associate Editor for Terrorism & Transnational Threats
Ashley L. Morefield, Associate Editor for Race & Security
Amelia Dal Pra, Associate Editor for Environment & Health
Xiang Chi, Associate Editor for Geoeconomics and Resource Competition

PEER REVIEW STAFF
Alexander Chang, Chief Peer Reviewer
Raj Dhillon, Senior Peer Reviewer James McGee, Peer Reviewer
Megan Johnson, Senior Peer Reviewer Mano Harada, Peer Reviewer
Aaron Mikell, Senior Peer Reviewer Laura Barela, Peer Reviewer
Samanvya Hooda, Peer Reviewer Logan Duarte, 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
Follow the Georgetown Security Studies Review on Twitter @gssreview
Contact the Editor in Chief at gssr@georgetown.edu
# TABLE OF CONTENTS

An Analysis of Israel’s Counter-Hezbollah Strategy ............................................................ 1
*Luke Baumgartner*

**Loyal Principles** ................................................................................................................. 16
The Logic of a French Return to NATO After the Cold War
*Paul Cormarie*

**Reframing Proxy War Thinking** ....................................................................................... 32
Temporal Advantage, Strategic Flexibility, and Attrition
*Amos C. Fox*

**When an Intelligence Agency Loses Its Mind** ................................................................ 51
An Analysis of the KGB’s Operation RYAN and What It Means for Russia’s War against Ukraine
*Patrick Hutson*

**Turning Grievance and Vulnerability Into Violence** ....................................................... 62
Lessons on Identity-Based Conflict from the United States and Overseas
*Jordyn Iger*

**Gender Roles and Military Necessity** .............................................................................. 84
Women’s Inclusion in the Liberation Tigers of Tamil Eelam
*Kara Joyce*

**Spies in the Skies** .............................................................................................................. 96
Analyzing the Development of the U.S. Commercial Satellite Industry vis-à-vis the Hermit Kingdom
*Sue Kim*

**International Legal Accountability and the World Order** ............................................. 118
Book Review of Geoffrey Robertson’s *Crimes Against Humanity: The Struggle for Global Justice*
*Miriam Kvaratskhelia*

**Soviet Nationality Policy and the Forgotten Periphery** .................................................... 123
Book Review of Kate Brown’s *A Biography of No Place: From Ethnic Borderland to Soviet Heartland*
*Miriam Kvaratskhelia*

**Alternative Perspective** .................................................................................................... 126
Can Gang Classification Mitigate the Threat of Violent Extremism?
*Michael P. Losacco*

**The Worst Menace to Society** ......................................................................................... 140
Turkey’s Approach to Digital Censorship
*Jessica Maksimov*

**Lessons Learned and Neglected** ..................................................................................... 151
Germany and the Viability of the Offensive Before World War I
*Christian Trotti*
An Analysis of Israel’s Counter-Hezbollah Strategy

Luke Baumgartner

Israel’s four-decade struggle against the Shi’a terrorist group Hezbollah has largely failed to produce positive results for the Jewish state. Since Hezbollah’s emergence in the early 1980s, their broad cross-sectarian appeal within Lebanon and vast financial support from the Islamic Republic of Iran continue to present a challenge to the Israeli state security apparatus. Having consolidated support of the historically disenfranchised Lebanese Shi’a population and with vast amounts of funding from Iran and a worldwide Shi’a diaspora, Hezbollah solidified itself as the vanguard of the 1979 Islamic Revolution and the primary bulwark against Israeli aggression. Following the failed assassination attempt of an Israeli diplomat in 1982, Israel launched a full-scale invasion of southern Lebanon as a means of finally defeating the Palestinian Liberation Organization, providing the stimulus for the various Shi’a militant groups simultaneously fighting sectarian militias as well as the Lebanese Army in the midst of a bloody civil war.

During the course of Israel’s subsequent 18-year occupation of Southern Lebanon, Hezbollah’s terrorist tactics evolved, eventually transforming the group into an insurgent army whose sole aim was to expel Israeli troops from Lebanon and implement traditional Islamic law. The Israeli Defense Forces’ (IDF) strategy to confront Hezbollah during the course of their occupation experienced persistent change; frequent changes in military leadership strained Israel’s civil-military relations and failed to produce a coherent strategy that adapted to Hezbollah’s changing tactics.

In the years since Israel’s unilateral withdrawal from Southern Lebanon in 2000, the lessons learned from their nearly twenty-year war as well as the shorter armed conflict against Hezbollah in 2006 failed to materialize. With domestic political turmoil and a resolute Iranian source of funding to Hezbollah, Israel’s strategy to counter Hezbollah’s political influence in Lebanon and its formidable military wing face continued difficulties for the foreseeable future.

Introduction

For nearly four decades, Israel has been engaged in a perpetual—albeit sometimes sporadic—war of attrition against Hezbollah, or the “Party of God.” Hezbollah first appeared as a loose confederation of disgruntled, politicized Shi’a fighting against the Palestinian Liberation Organization (PLO) and other militias for control of southern Lebanon in the early 1970s. Following Israel’s invasion of southern Lebanon in June 1982, Hezbollah established itself as the preeminent Shi’a militant group fighting the “Zionist occupiers,” beating out the likes of the secular, former co-belligerent from which they split, Amal. Now a significant force in Lebanese politics as part of the March 8 Alliance, Hezbollah primarily exists to provide essential services to its constituents while simultaneously serving the regional interests of the Islamic Republic of Iran through a three-pronged approach, encompassing military, political, and religious pretexts and representing a perpetual thorn in the side of the Jewish state.

The cornerstone of Israel’s strategic doctrine—strength through deterrence—is preventing large-scale conventional wars
with its Arab neighbors, like those in 1948, 1967, and 1973. Relying on preemptive offensive operations and leveraging defensible borders, intelligence superiority, and air force supremacy, the IDF utilized an “indirect approach.” Consequently, the IDF categorized its security threats into “fundamental” and “routine” security. The former involved countering conventional threats from neighboring Arab armies, while the latter focused on border infiltrations and continual terrorist and guerrilla incidents. For the IDF, threats posed by Palestinian militants from Jordan, Lebanon, and other Arab states fell into the latter. At the same time, they relied on punishment, or the threat thereof, as the primary form of deterrence for any perceived or actual acts of aggression by several hostile actors. This line of thinking shaped the IDF’s earliest efforts against Hezbollah throughout the 1980s and the early 1990s, with little credence given to diplomatic efforts. Since Hezbollah’s formal inception in February 1985 and entry into the political arena in 1992, the success of Israel’s continued counterterrorism efforts directed at the Party of God is limited, ebbing and flowing with the tides of regional politics and punctuated by periodic flare-ups and “small wars,” both within its borders and beyond. While Israel maintains a significant conventional military compared to Hezbollah, to this day, the group remains a significant terrorist threat with a vast network of international training, funding, and sympathizers. Consequently, Israel’s strategy to counter Hezbollah’s political and military influence within the region and reduce its threat to the Israeli state has largely failed.

This essay discusses and analyzes Israel’s countermeasures against Hezbollah and their impacts in four parts. Part one provides a historical background of Hezbollah’s ideology and the group’s formation amid Lebanon’s civil war. Part two examines Israel’s counterterrorism efforts through distinct phases: phase one covers from 1982 through 1992; phase two, from 1992 through 2000; phase three, from 2000 through 2006; and phase four, from 2006 to the present. Part three addresses a few lessons from Israel’s fight against Hezbollah that apply to the United States’ campaign against groups such as al-Qaeda and the Islamic State. Finally, part four summarizes this essay’s findings and comprehensively reviews Israel’s counterterrorism strategy against Hezbollah.

Part I: History of Hezbollah

For even the most well-rounded historians, political scientists, and counterterrorism analysts, attempts at accurately portraying the authentic character of Hezbollah continue to prove difficult, engaging in semantic debates surrounding the subtleties and differences between Hezbollah’s political and military lines of effort. However, participating in such discussions is beyond the scope and purpose of the following analysis. Despite the absence of an agreed-upon definition of terrorism in academic and policy-making circles, Israel and the United States classify Hezbollah’s military activities as terrorism. In doing so, “by definition, any act of violence that it commits or seeks to commit is an act of terrorism, and so there are no gray areas of justifiable behavior.” Therefore, working within the confines of Israel’s designation of Hezbollah as a terrorist organization satisfies the conditions required for properly analyzing its countermeasures against the group and their corresponding outcomes.
Geopolitical Landscape

After the Six-Day War in June 1967, a massive influx of Palestinian refugees fled to the surrounding countries of Egypt, Jordan, and Lebanon. Nearly 300,000 refugees settled into camps around Amman and other cities, igniting tensions between the indigenous populations and the Palestinian newcomers.\(^9\) Notwithstanding the 1982 Israeli invasion, the most significant event transforming various factions of discontent Shi’a Muslims in Southern Lebanon into Hezbollah was the expulsion of PLO fedayeen from Jordan following the Jordanian Civil War (also known as “Black September”) in 1970. The PLO and other radical Palestinian groups, such as the Popular Front for the Liberation of Palestine (PFLP), used the Karameh refugee camp in Jordan as a forward operating base, conducting frequent cross-border raids into Israeli territory and attacking IDF patrols.\(^10\) Inevitably, heavy-handed reprisals by the IDF were commonplace, and tensions between Jordan’s King Hussein and the Palestinians boiled over, primarily due to their regular border skirmishes jeopardizing Hussein’s aspirations for peace with his Jewish neighbor. After three failed assassination attempts by Palestinian militants and the events at Dawson’s Field, King Hussein declared martial law on September 16, 1970, and the Jordanian Armed Forces (JAF) launched a brutal campaign against the refugee camps, killing as many as 15,000 Palestinian militants and civilians.\(^11\) As fighting continued through 1971, the predominantly Sunni Palestinian refugees continued flowing into southern Lebanon, raising the potential for disturbing the country’s already delicate sectarian balance that was decades in the making.

Shi’a Politicization

During the 1950s and 1960s, the Lebanese Shi’a became politicized through the efforts of the prolific cleric Musa al-Sadr. Al-Sadr began his clerical journey in the elite Shi’a seminary of Najaf in Iraq and returned to Lebanon in 1959 to unite the Shi’a youth under a shared ethnic and sectarian identity, encouraging his followers to speak out through their religion and overcome their condition.\(^12\) Al-Sadr’s aspirations came when the Lebanese Shi’a were known for their poverty, underdevelopment, and the growing temptation to adopt leftist and nationalist ideologies such as Ba’athism, Marxism, and Nasserism.\(^13\) Al-Sadr rose to prominence in 1969 as the elected leader of the Lebanese Supreme Islamic Shi’a Council, a representative body authorized by the Lebanese parliament that provided the Shi’a political autonomy independent of their Sunni countrymen.\(^14\)

Shortly thereafter, the PLO—heavily concentrated in the suburbs of Beirut and southern Lebanon—established a “state-within-a-state” following their expulsion from Jordan, challenging the authority of the Lebanese government and undermining the progress of al-Sadr’s Harakat al-Mahrumin (“Movement of the Deprived”) and its armed militia, Amal.\(^15\) Founded in 1974 by Musa al-Sadr, Amal sought to challenge “the stifling and often brutal domination of the Palestinian guerrillas whose public support plummeted in the late 1970s and early 1980s for bringing southern Lebanon into the crossfire with Israel.”\(^16\)

As intense clashes between Amal and the Palestinian militias increased in frequency during the early 1970s, so did those between the Maronite Christian-led Lebanese Front and the Palestinians.\(^17\) This fighting eventually metastasized into an all-out civil war in 1975, and the various militias divided
along sectarian lines, further sinking Amal into relative obscurity. In March 1978, Israel invaded southern Lebanon in Operation Litani to eliminate its Palestinian foes and establish a security zone along its northern border. Consequently, the introduction of Israeli troops to Lebanon combined with Musa al-Sadr’s disappearance the same year and the Iranian Revolution in 1979 resulted in a resurgence of sectarian fervor among the Shi’a population and, thus, a second wind for Amal.

Part II: Israel’s Counterterrorism Strategy

Phase 1A: 1982-1988, Hezbollah’s Early Resistance

On June 6, 1982, the IDF invaded southern Lebanon after, according to Israel, the PLO violated an eleven-month cease-fire through the failed assassination attempt of the Israeli ambassador to the United Kingdom, Shlomo Argov. In response, Israeli Defense Minister Ariel Sharon ordered the commencement of Operation Peace for Galilee. According to Schulze, the operation’s objectives were “to destroy the PLO’s military and political infrastructure, strike a serious blow against Syria, and install a Christian regime that would sign a peace treaty with Israel.” In this phase, “utterly preoccupied with the PLO when it invaded, Israel paid little attention to the Lebanese Shi’a community, which predominates in southern Lebanon, the northern Biqa valley, and Beirut’s southern suburbs.” For Iran and Ayatollah Khomeini, the deployment of approximately 1,500 military advisors from the Iranian Revolutionary Guard Corps (IRGC) to Lebanon shortly after the invasion “represented the realization of the revolutionary state’s zealous campaign to spread the message of the self-styled ‘Islamic Revolution.’” The Ayatollah’s order followed the meeting of a limited contingent of nine Shi’a Islamist movement leaders, calling themselves the “Shura Council.” This Shura Council became a reference point for the Lebanese Shi’a resistance to Israeli occupation until May 1984, when the name “Hezbollah” was officially adopted.

Following the opening phase of high-intensity combat operations, Israel had no intention of disengaging from the Lebanese theater. The IDF’s operational focus after 1983 was low-intensity warfare against the various groups hostile toward their occupation. This period of 1982-1985, according to Hezbollah’s deputy secretary-general, Naim Qassem, “was foundational for the crystallization of a political vision, the facets of which were harmonious with faith in Islam as a solution and for the establishment of an effective jihad operation as represented by the Islamic Resistance forcing Israel’s partial flight from Lebanon in 1985.” It was during this time that the nascent Party of God pioneered the tactic of suicide bombings, or “martyrdom operations,” against Israeli targets in Lebanon beginning in November 1982 when an attack against an Israeli government building in Tyre claimed the lives of 91. In doing so, Hezbollah demonstrated the primary purpose of suicide terrorism: “to use the threat of punishment to coerce a target government to change policy, especially to cause democratic states to withdraw forces from territory terrorists view as their homeland.” These suicide operations also suggested that Hezbollah’s religious fervor could compensate for the gross imbalance in capabilities between the competing forces, therefore laying claim to a devotion from its adherents that no other militia—not even Amal—could match.
Over the next year, Hezbollah continued using high-profile suicide bombings, most notably against the United States Embassy and Multinational Force barracks in Beirut in April and October 1983, respectively. The “use of suicide bombing was based on clear cost-benefit calculations and was adopted because of early lessons learned of its propaganda value and its effectiveness in achieving both political and military goals.” In response, the IDF launched massive bombing raids deep into Lebanon “to ensure that everyone … gets the message … a successful attack like the car bomb is likely to become a model for imitation.”

While the IDF was practically defenseless against suicide tactics, they experienced greater success repelling Hezbollah’s more conventional assaults. Marcus states:

“Hezbollah carried out large-scale (but ineffective) ‘hill-storming’ operations, which entailed fighters openly charging the IDF’s hilltop fortifications and were partially inspired by Iran’s tactics during the Iran-Iraq War. An incident in February 1987—a frontal assault on IDF positions at Beaufort Castle in which Hezbollah was easily targeted by IDF forces and suffered eight casualties—exemplifies Hezbollah’s ineffective early military strategies and tactics.”

Moreover, the IDF’s success in countering Hezbollah’s primitive conventional tactics corresponded with an increase in the group’s use of suicide bombings, reaching their peak in the late 1980s.

Israel’s response to Hezbollah’s religiously inspired suicide bombings and human-wave attacks throughout this phase was comparable to other military powers blinded by conventional asymmetries, blissfully unaware of the growing guerrilla insurgency in their occupied territories. Similarly, even the February 1985 release of “An Open Letter Addressed by Hezbollah to Downtrodden in Lebanon and the World” failed to signal to the IDF Hezbollah’s commitment to driving them from Lebanon. Hezbollah’s open correspondence with the world clearly stated the group’s objectives: 1) to force “the final departure of America, France, and their allies from Lebanon and the termination of the influence of any imperialist power in the country;” 2) to use “Israel’s final departure from Lebanon [as] a prelude to its final obliteration of existence and the liberation of venerable Jerusalem from the talons of occupation;” and 3) implement an Islamic state modeled after that of 1978-1979 Iranian revolution under *velayat-e-faqih*.

By 1984, the rate of attacks against the IDF was so intense that an Israeli soldier was dying every third day. The climax during this period of the conflict between Hezbollah and the IDF came in June 1985, when two members of Hezbollah hijacked TWA flight 847 from Athens, demanding the release of more than seven hundred combatants held in Israeli prisons. Frequent suicide bombings, including those against American and Multinational Force (MNF) targets in Lebanon, combined with the TWA hijacking would prompt the IDF to withdraw a majority of its forces from Beirut in 1985 and consolidate into an enlarged “security zone” encompassing nearly 10 percent of all Lebanese territory.

**Phase IB: 1988-1992, From Civil War to Abroad**

As the Israeli occupation entered its sixth year, training and funding from the IRGC continued to flow into Hezbollah’s coffers. Its fighters received advanced training in explosives and demolition, field intelligence, and reconnaissance. One of the earliest recipients of the IRGC’s training in the
Baalbek region of southern Lebanon was Hezbollah’s second Secretary-General, Abbas al-Musawi. Al-Musawi recognized the efficacy of this training, utilizing it to pass on to other Hezbollah fighters for use against their rival Amal and the IDF in the late 1980s. As a result, Hezbollah conducted more complex ambushes using rocket-propelled grenades (RPGs) and small arms during day and nighttime raids against IDF positions, precipitating significant IDF reprisals. In May 1988, the IDF conducted combined-arms combat operations outside the security zone for the first time to target the Hezbollah strongholds of Maydun and Luwayza, killing sixty militants. These operations reflected “the IDF’s increased entrenchment in southern Lebanon and its ... willingness to confront Hezbollah in reactive, higher-tempo operations.” Considering its previous tactical shortcomings, Hezbollah adapted to the operational environment again; it began using roadside improvised explosive devices (IEDs) against the IDF, carrying out 178 attacks in 1988 alone. In response, the IDF improved the protective capabilities of its defensive fortifications, further highlighting the IDF’s hubris and static and passive mode of operation.

Violence between Hezbollah and Amal culminated from 1988 to 1989 following the kidnapping of United States Marine Corps Lt. Col. William R. Higgins, resulting in an eruption of clashes in the southern suburbs of Beirut and Hezbollah, ultimately winning the Shi’ā heartland and significantly degrading Amal’s influence in the conflict. In October 1989, the Taif Accords ended the civil war, calling for all militias to disarm. However, with Iranian support, Hezbollah circumvented these requirements and justified the continued existence of its armed factions by referring to them as “Islamic resistance” groups committed to ending Israel’s occupation.

The IDF continued operating under the flawed assumption that the Taif Accords would end most of the attacks directed their way. Despite this formal peace process, under the direction of al-Musawi, Hezbollah’s rocket attacks into northern Israel continued, prompting his assassination in 1992 by the IDF. Succeeded by current Secretary-General Hassan Nasrallah after al-Musawi’s death, Hezbollah retaliated with increased rocket attacks, forcing the IDF to implement a more liberal targeting policy. According to Marcus, “this tit-for-tat exchange of firepower marks the origins of the deterrence equation that would evolve throughout the next decade.” Although more liberal targeting policies resulted in nominal battlefield results in Lebanon, Hezbollah struck Israeli targets abroad. In 1992 and 1994, Hezbollah attacked the Israeli Embassy and a Jewish cultural center in Buenos Aires, Argentina, highlighting Israel’s continual failure to adapt to its evolving threat landscape.


In light of the continued violence, the Oslo Accords in 1993, and domestic political pressures calling for an end to the occupation, the IDF remained engaged with Hezbollah in Lebanon and attempted to secure peace for Israel’s northern communities. Former Israeli defense officials classify this period as one in which the IDF operated with “one hand behind its back,” unable to change the strategic considerations for its involvement in Lebanon, succumbing to political restraint, and culminating in “the rules of the game.” Restricted to “containment operations” against Hezbollah in Lebanon, the IDF’s tactics had a limited impact. They often resulted in significant drops in morale and disproportionate responses to Hezbollah’s attacks, thereby violating the “rules of the
Following an attack that killed seven IDF soldiers in July 1993, the IDF conducted its first large-scale conventional operation of the decade with “Operation Accountability.” The underlying goal was to inflict a “heavy price” on Hezbollah by pressuring the Lebanese population and government to cease supporting the group. After action reviews (AARs) revealed much of the same: Hezbollah fired more Katyusha rockets into northern Israel, hoping to deter the IDF from conducting future operations of this kind. Still adhering to the “rules of the game,” the IDF’s slow and sporadic operations under General Mordechai proved inadequate to quell the violence from Hezbollah, commanding a significant paradigm shift in the IDF’s counterterrorism strategy beginning in 1994.

Under Major General Amiram Levin, the IDF refocused their efforts by conceptualizing Hezbollah as a “guerrilla army” instead of a “terrorist group.” Opting for a more offensive approach to the group, the IDF formed an elite specialized commando unit—the Egoz. Bearing some resemblance to the U.S. Army’s Special Forces, the Egoz’s primary mission was to train others in the necessary tactics of “counterguerrilla” or “counterinsurgency” warfare. Utilizing “pinpoint” operations, the Egoz launched missions deep into Lebanese territory, away from the security zone, to selectively strike prominent Hezbollah fighters. Complementing the Egoz’s debut in 1995, the IDF launched “Operation Grapes of Wrath” in April 1996, resulting in significant Lebanese civilian casualties and once again incurring the wrath of Hezbollah’s Katyusha rockets. The operation intended to undermine popular support for Hezbollah among the Lebanese populace and quell their support from Syria. It ultimately failed due to the massacre at Qana, a United Nations (UN) base housing Lebanese civilians seeking refuge from IDF ground and air attacks. The 1996 offensive failed to achieve its strategic goals, demonstrating that the reconceptualization of the group from a terrorist to a guerrilla threat operating as a commando force proved too little too late. Although IDF casualties continued to decrease until its withdrawal in May 2000, “the small and intimate nature of the IDF and the sacrosanct cultural importance of the IDF in Israel as the ‘guarantors of Jewish survival’ ensured that even limited casualties reverberated deeply in Israeli public discourse and civil society and impacted the thinking at the top and bottom of the military.”

Phase III: 2000-2006, Beyond Occupation

Contrary to predictions of mayhem following Israel’s withdrawal, from 2000 to 2006, an “uneasy quiet” fell over the region. Hezbollah limited its attacks to the Shebaa Farms area in the Golan Heights—a territory disputed by Israel, Lebanon, and Syria that eventually became Hezbollah’s justification for maintaining its arms. From 2000 to 2003, Hezbollah initiated more than 250 anti-aircraft, mortar, and rocket attacks, which Hassan Nasrallah described as “reminder operations.” In this context, Nasrallah compared Israel’s stoic and militaristic public image to a spider web, describing it “as a formidable military power, but one that is rooted in a civil society that has become materialistic and lazy, with self-satisfied, comfortable, and pampered citizens who have gone soft.” Despite the frequency of these small-scale attacks, only 27 Israeli soldiers were killed.
between 2000 and 2006, compared to the annual rate of 25 throughout the IDF’s 18-year occupation of Lebanon. With Syrian forces continuously supporting Hezbollah and being the only foreign power remaining in Lebanon, the IDF adapted its deterrent posture “from the ‘routine security’ realm vis-à-vis Hezbollah to the realm of conventional state-on-state deterrence vis-à-vis Syria.” Thus, Syria received the brunt of Israeli retaliation in response to Hezbollah's attacks on the IDF.

During this phase, the most telling aspect of the Israeli counterterrorism strategy against Hezbollah was the perceived erosion of IDF deterrence, punctuated by Israeli’s willingness to participate in large-scale prisoner swaps and failure to follow through on its avowal to respond to attacks and kidnappings. The reasons for this relatively weak response were twofold. First, the risk-averse and casualty-sensitive Israeli public would not tolerate further entanglement in Lebanon after nearly two decades of war. Second, the IDF failed to capitalize on the lessons learned during its occupation and establish a coherent deterrence strategy against Hezbollah’s hybrid threats. As a result, Hezbollah operated on the expectation that Israel would continually endure intermittent incursions and minor border skirmishes as part of the long-held tit-for-tat “rules of the game” that defined the post-withdrawal environment.

Adherence to the “rules of the game” would only last until July 12, 2006, when Hezbollah launched an attack on Israeli soil and ambushed an IDF patrol on Lebanon’s southern border with Israel. In the attack, Hezbollah captured two Israeli soldiers and killed three more. Drawing on the IDF’s prior inaction, Hezbollah took on the role of a rational actor and operated within the confines of what it perceived as Israel’s “red lines.” While initially successful, the daring cross-border raid to fulfill its wa’d al-sadiq (“faithful promise”) to secure the release of Lebanese prisoners in Israeli custody was a grave miscalculation. According to Lieutenant General Dan Halutz, Israel’s retaliatory offensive readily displayed the IDF’s commitment to fighting terror “by inflicting continuous, painful blows that inflict on the other side a much higher price than he ever expects ... one that produces deterrence.”

Relying on air power and heavy artillery bombardment, the IDF attempted to achieve Prime Minister Ehud Olmert’s war aims: 1) returning the kidnapped soldiers to Israel; 2) a complete cease-fire; 3) the expulsion of Hezbollah from the area in line with United Nations (UN) Resolution 1559 (which calls for Hezbollah’s disarmament), and 4) the deployment of the Lebanese Army in southern Lebanon to maintain security. However, in an eerie repeat of 1996’s “Operation Grapes of Wrath,” the Israeli bombing of Qana during the 2006 war that killed 28 civilians quickly shuddered its base of Arab support from Saudi Arabia, Jordan, and the United Arab Emirates. The Israeli operation lasted 34 days, and despite killing approximately 500 fighters and using psychological operations to discredit Hezbollah, the IDF failed to achieve its strategic goal of restoring the credibility of its deterrence.

Phase IV: 2006-Present, Syria and the “Calm Between Wars”

Israel envisioned its survival through a dichotomous lens of peace or war for nearly its entire existence. In the wake of the 2006 Lebanon War’s strategic failure, the IDF concluded that a “total victory” over a highly-capable, well-funded guerrilla group such as Hezbollah was not possible in the near term. High-level leadership recognized that non-state actors conduct a long, twilight
struggle—a slow-motion, slow-burning war against them. Rather than accumulating “tactical” victories that would lead to an overall strategic win, the IDF’s strategic goal in its fight against Hezbollah is to degrade its military capabilities—sometimes dubbed “mowing the grass.” Through sharp, decisive blows that restore Israel’s deterrent position and deliver a period of “relative” quiet, the IDF anticipates draining the motivation from enemy fighters to harm Israel and eventually cause the movement’s collapse. Referred to as the “campaign between wars,” Israeli counterterrorism operations during these periods of “relative” quiet are used to “foil enemy military buildup, and led to the deaths of figures in Hamas, Hezbollah, and the ranks of the Syrian military.”

Conversely, “Israel understood that its enemies use the years of quiet on its borders to build up their ranks and capabilities, incrementally raising the threat to Israel’s population and encroaching on its freedom of operation.” According to the Atlantic Council, Hezbollah’s active manpower is approximately 30,000 fighters, with a reserve force of another 10-20,000. Additionally, Israeli estimates place Hezbollah’s surface-to-surface rocket and missile arsenal at 130,000 to 150,000, signaling a marked increase in the group’s strength during the “campaign between wars” and suggesting a significant weakness in the IDF’s strategy of gradually degrading Hezbollah’s military capabilities.

With the outbreak of civil war in Syria in 2011, Hezbollah’s military priorities and Israel’s countermeasures against them took on a new dimension, entrenching themselves in regional politics and great power competition. Initially justifying its intervention in the conflict to protect the Syrian-Lebanese border from spillover, Hezbollah’s role transformed from “advising” and “assisting” to direct combat operations against the Islamic State, al-Qa’ida, and other Sunni militant groups. In the broader context of Iran’s regional interests, Tehran’s vested interest in the Assad regime’s stability and ultimate survival is twofold. First, Syria serves as the physical and symbolic land bridge through which weapons, funding, and other forms of support reach Hezbollah in Lebanon. Second, as a physical and ideological extension of the Islamic revolution, Hezbollah is an external expression of Iran’s perpetual war of annihilation against Israel. Given the implications of the drastic transformation in the geopolitical landscape over the last decade, alterations in the Israeli strategy for countering Hezbollah are hardly surprising. Consequently, Israeli air strikes against Iran, Hezbollah, and the Assad regime in Syria reflect this. Recognizing attempts at holding the Lebanese government accountable for Hezbollah’s actions are futile, “IDF strikes are designed to erode Hezbollah’s capabilities as a part of a long-term attritional campaign to cumulatively deter Hezbollah.”

Part III: Evaluations

Israeli Strategic Evaluation

In nearly four decades of asymmetric competition accented by periods of intense violence, numerous opportunities for the IDF to adapt its strategy to that of Hezbollah presented themselves but failed to materialize. While successful during the opening phase of “Operation Peace for Galilee” and subsequent occupation of southern Lebanon in the early 1980s, the IDF’s strategic doctrine historically emphasizing conventional military power to counter existential threats left it ill-equipped to defend against the nascent Hezbollah’s...
guerrilla and terrorist tactics. Moreover, the formation and early success of the Egoz and strained civil-military relations during the Oslo Accords peace process in the early 1990s hamstrung the IDF’s efforts to assimilate its strategy to Hezbollah’s evolving threat profile in the aftermath of the Lebanese civil war. Lastly, sustained external support from Iran and Syria amplified Hezbollah’s effectiveness.

With mounting pressure from a war-weary Israeli public, the looming realization within the Ministry of Defense and Knesset that the status quo was untenable, and the IDF’s false hope for Hezbollah’s continued adherence to “the rules of the game,” Israel unilaterally withdrew from southern Lebanon in May 2000. The subsequent “calm between wars” prompted an uneven recalibration of the mutual deterrence equation defining the Israel-Hezbollah conflict. As Hezbollah continued to gain domestic and international support in the early 2000s, Israel’s focus on the Second Intifada detracted from its ability to focus on launching an effective campaign targeting Hezbollah’s leadership, counternarratives against its propaganda efforts, and its primary sources of funding and training. The 2006 war is the last example highlighting the strategic failures of the IDF in countering Hezbollah. Israel’s conventional retaliation after Hezbollah’s deadly attack on IDF soldiers in the Shebaa Farms area killed approximately 600 fighters but took the lives of nearly twice as many civilians, prompting massive international condemnation and an UN-brokered cease-fire after 34 days of fighting. More importantly, the war resulted in Hezbollah’s military capabilities and domestic popularity emerging more potent than before. Pew Research polls in 2007 indicated Secretary-General Hassan Nasrallah was among the most popular leaders in the Middle East and North Africa, completely negating one of the IDF’s goals of the 2006 war of diminishing Hezbollah’s popularity.81

Lebanon: Israel’s Afghanistan?

On August 30, 2021, the last U.S. troops left Afghanistan, marking the end of the longest war in American history and evoking comparisons to the withdrawal from Vietnam in 1975. In the aftermath of 9/11, the initial U.S. invasion of Afghanistan aimed to remove the Taliban regime from power, destroy al-Qa’ida’s training camps, and prevent them from carrying out future attacks against the U.S.82 In March 2003, the United States invaded Iraq; it toppled the regime based on intelligence suggesting Saddam Hussein harbored al-Qa’ida training camps and continued its nuclear weapons program in direct contravention of multiple United Nations decrees.83 The American military engaged in perpetual nation-building and counterinsurgency operations for nearly twenty years as part of the broader Global War on Terror. Since the American national security posture pivoted back to great power competition, “over-the-horizon” capabilities remain the centerpiece of American counterterrorism policy, as demonstrated by the drone strike killing al-Qa’ida leader Ayman al-Zawahiri in August 2022.84 Furthermore, several lessons emerge when comparing Israel’s fight against Hezbollah to that of America’s against al-Qa’ida.

First, the United States and Israel heavily relied on conventional military strength to defeat adversaries utilizing guerrilla and terrorist tactics. In its germinal stage, Hezbollah was primarily a religious terrorist organization that sought to establish an Islamic state in Lebanon. Turning to suicide bombings to impose its will and drive the Israeli occupying force from Lebanon, Hezbollah’s use of such tactics “suggested...
that religious fervor could compensate for small numbers.” Subsequent Israeli countermeasures to Hezbollah’s suicide attacks were overwhelmingly subservient to the military deterrent posture that emerged from the IDF’s previous existential wars against its Arab neighbors. Drawing on conclusions from the previous analyses in this essay, attempting to counter Hezbollah’s religiously motivated extremist violence “by pursuing qualitative superiority” did not produce the intended effects and, in turn, further emboldened Hezbollah’s resistance activities. Similarly, when responding to the asymmetric threats of al-Qaeda, the U.S. strategy fell victim to the same short-sightedness.

Second, involvement in the Second Intifada in September 2000, less than four months after the withdrawal from Lebanon, had two detrimental effects: 1) allowing Hezbollah to reconstitute and increase its military capabilities and 2) degrading the IDF’s ability to make doctrinal changes to a deterrent posture better suited to combatting Hezbollah’s transformation into a hybrid threat. Analogous to the Second Intifada, America’s shift in priorities to Iraq—an [near] immediate engagement in a new area of operations—detracted from developing a coherent strategy conforming to the conditions on the ground in the nation’s preceding military contest. Afghanistan, America’s “original” combat theater in the Global War on Terror, inevitably played second fiddle to the deteriorating situation in Iraq at the expense of allowing al-Qaeda to regain its strength and preventing the U.S. from harmonizing competing priorities.

Part IV: Conclusion

With its heavy human and financial capital investment in nearly all sectors of Lebanese society, Hezbollah presents a particularly unique challenge to Israel against the backdrop of more “traditional” terrorist organizations in its immediate vicinity, such as Hamas and the Palestinian Islamic Jihad (PIJ). As the most formidable opponent to Israel’s occupation of southern Lebanon in the early 1980s, Hezbollah’s evolution from a Shi’a terrorist group hellbent on Israel’s destruction to a legitimate player in Lebanese politics with a visibly armed resistance wing constitutes a national and regional security threat extending beyond its home in Beirut. Conversely, the Israeli countermeasures against the group failed to adapt to Hezbollah’s Janus-faced nature. Instead, its perpetual reliance on conventional military superiority projected through its deterrent posture against its regional neighbors left the IDF woefully unprepared to meet the challenge posed by a militant organization as multifaceted as Hezbollah.

Since the 2006 war’s conclusion and Hezbollah’s involvement in the Syrian civil war beginning in 2011, Israeli countermeasures continue to reflect a strategy that ignores various alternative mechanisms. Both sides understand that the next round of conflict invites mutually devastating consequences, and neither can eliminate the threat of retaliation in the event of a first strike. Therefore, the current solution presupposes that each side is deterring the other through the promise of a full-scale retaliation. The accumulation of sharp, decisive blows directed against Hezbollah’s military capabilities may give Israel a comparative advantage in a future conventional conflict. However, participation in electoral politics, its deep integration into the Lebanese state, and the
maintenance of its armed wing to serve the interests of Iran and vanguard of the Islamic revolution illustrate Hezbollah’s complex nature and ability to withstand Israeli countermeasures.

Finally, shifts in Israel’s domestic political stability present similar challenges to the IDF and political leadership as during the 1990s. Last December, Benjamin Netanyahu’s return as Prime Minister of Israel and leader of the country’s most far-right government in history risks reigniting tensions between Israel and Hezbollah.\(^88\) Similar to the IDF’s involvement in the Second Intifada distracting the Israeli government from Hezbollah’s rearmament between 2000 and 2006, the domestic fallout from protests over Israel’s proposed changes to its Supreme Court could strike a blow to its focus on combatting Hezbollah’s rocket attacks and border incursions.\(^89\) Hezbollah’s Secretary General, Hassan Nasrallah, echoes these sentiments, stating that Israel’s domestic crisis will prevent them from “responding too harshly.”\(^90\)

Moreover, mixed responses from the IDF in confronting Hezbollah’s activities on its northern border with Lebanon further obfuscate any resemblance of a coherent strategy to deter the Party of God. Continued provocations by Hezbollah fighters and activists, such as damaging Israeli surveillance equipment, establishing small outposts, and firing anti-tank missiles at the fence demarcating the UN-recognized Blue Line, are met with limited responses from the IDF, thus testing their resilience.\(^91\) However, at the IDF’s insistence, the force with which the IDF retaliated against Hezbollah during the 2006 War openly deters the group from undertaking any serious military operations. With domestic political upheaval, increasingly brazen attacks on Iran\(^92\)—Hezbollah’s primary source of funding—and a muffled response to Hezbollah’s probes into Israeli territory, Israel’s campaign of mixed messages to Hezbollah threatens the current “uneasy peace” and risks further bloodshed.

**About the Author:** Luke Baumgartner is a fourth-semester graduate student at Georgetown University’s Walsh School of Foreign Service and a member of the Security Studies Program, concentrating in terrorism and sub-state violence. Luke is a US Army veteran, serving from 2017-2021. He was commissioned as an active duty Field Artillery Officer from Southern Illinois University Carbondale after earning a B.A. in Political Science. Luke currently serves as a Program Manager for We the Veterans and Military Families, a 501(c)(3) nonprofit dedicated to strengthening American democracy and combating the spread of targeted violence and misinformation among the military veteran community.
3 Ibid.
4 Ibid.
5 Ibid.
9 Bruce Riedel, “Fifty Years after ‘Black September’ in Jordan,” *Studies in Intelligence* 64, no. 2 (June 2020): 35.
12 Ghaddar, *Hezbollahland*, 17.
14 Ibid, 10.
15 Ibid, 8.
16 Ibid, 12.
17 Imad, “Republic of Lebanon,” 351.
19 Imad, “Republic of Lebanon,” 351.
21 Ibid, 21.
26 Ibid.
32 Ibid, 45.
33 Ibid, 42.
34 Ibid, 43.
36 Ibid.
38 Ibid, 31-32.
40 Marcus, *Israel’s Long War*, 44.
41 Ibid.
42 Ibid, 45.
43 Ibid.
44 Ibid.
45 Ibid.
49 Marcus, *Israel’s Long War*, 50.
51 Marcus, *Israel’s Long War*, 54.
52 Ibid, 54-55.
53 Ibid, 55.
54 Ibid, 56.
55 Ibid, 56-58.
56 Ibid, 57-58.
57 Ibid, 61.
59 Ibid, 68.
60 Ibid, 70.
64 Norton, *Hezbollah*, 103.
67 Marcus, *Israel’s Long War*, 98.
68 Ibid, 99.
70 Ibid, 123-126.
74 Ibid.
76 Ibid, 3.
77 Ibid.
81 Karoç et al., “Hezbollah’s Popularity,” 86.


Loyal Principles: The Logic of a French Return to NATO After the Cold War

Paul Cormarie

It is curious that France decided to return closer to the NATO Alliance after the Cold War, in a notably calm and stable period for European security. It is even more striking to note that, 10 years later, France has become popularly perceived as a reluctant NATO ally and a staunch defender of the European Union’s autonomy in strategic affairs. As such, this paper presents the case that the same pursuit of strategic autonomy is much older than commonly perceived. In fact, the quest for autonomy was the prime motivation for France in returning closer to NATO Alliance after the Cold War. This paper explores the historical facts related to the French rapprochement to NATO and explains the negotiations made at the time related to European autonomy. Altogether, it concludes that French foreign policy has selected different regional organizations in the post-Cold War period as avenues to build a common defense in Europe, be it through NATO or the EU.

Introduction

French President Emmanuel Macron’s recent declarations after his state visit to the People’s Republic of China have provoked significant pushback from the United States and his Eastern European partners. In an interview, he defended his goal of Strategic Autonomy on the Taiwan issue, which he qualified as “American escalation and Chinese overreaction.” This statement came as a surprise to many observers who believed President Macron’s vision of European autonomy to be “brain dead” – as the President had once labeled the North Atlantic Treaty Organization (NATO) – since the Russian invasion of Ukraine. With Finland’s accession into the Alliance and Sweden’s bid to join, there are few EU members left that are not members of NATO.

While looking at France’s historic relationship with NATO, Macron’s statements may come as a surprise. France has been, with other members such as Spain, one of the few member countries not party to NATO’s integrated command ever since its dramatic departure from NATO under Charles de Gaulle in 1966. While France reassured allies that it would still contribute its forces to a contingency against the Soviet Union during the Cold War, it meant that French forces would be led independently from Allied command and that France could not sit at the table of NATO’s military command. This would only change officially in 2009, under Nicolas Sarkozy’s presidency, after a long process of negotiations that started at the end of the Cold War. After French reintegration into NATO Military Command, the newly elected Francois Hollande was publicly skeptical of the reintegration process but decided not to flip the decision despite pressures from his party. Ultimately, his successor, President Macron, proved to be more confrontational towards NATO in public declarations than his predecessor, yet has not left either.

In just ten years, France went from upending a 50-year-old stalemate and reintegrating NATO, to calling it “brain dead.” As such, it might seem that the French return was a mistake and contrary to modern interpretations of Strategic Autonomy.
While most of the current literature centers around the question of whether Europe must become more autonomous from the U.S. defense umbrella, I argue that Strategic Autonomy is an age-old goal of French foreign policy that does not oppose membership in NATO. Further, this paper examines whether the French decision to rejoin NATO was because of, rather than despite, Strategic Autonomy.

To do so, I present a historical study of French foreign policy since the end of the Cold War which demonstrates a clear linear path: in developing a European defense structure that is separate from the United States by examining past declarations and analyses, we gain a broader understanding of why France rejoined NATO’s integrated command, and its boasting of “autonomy” once inside.

By examining the French position historically, we can not only better understand the decisions made by President Macron’s predecessors, but also the French meaning of the Alliance and European defense building. According to this study, Strategic Autonomy and membership in NATO are not opposed to the construction of a European defense but are more so a continuation of post-Cold War French foreign policy by other means.

**Mitterrand’s Shy Attempt**

When the Berlin Wall fell in 1989, French President François Mitterrand believed that the United States would leave Europe, based on phone calls with Soviet Union General Secretary Gorbachev. During those calls, President Mitterrand thought that an alternative security arrangement with the Soviet Union was possible. Not long after, European leaders agreed on the Charter of Paris and transformed the Conference on Security and Co-operation in Europe into the Organization for Security and Co-operation in Europe (OSCE).¹

However, some decision-makers inside Mitterrand’s administration believed that there was a change in strategic thinking in the United States after the fall of the Berlin Wall. According to some historians, the United States was looking into ways to make NATO “more European” as the Cold War was coming to an end. This belief led Mitterrand’s foreign policy advisor, Admiral Lanxade, to hold discussions with Brent Scowcroft and then-National Security Council (NSC) staffer David Gompert.² During those talks, the United States had hoped that France would become a key part of NATO’s transformation, whereas France, on the other hand, mostly remained passive to U.S. advances. The negotiations were parallel to those made under the “quad format” in the Autumn of 1990 between German, French, American, and British diplomats, which concluded that Europeans had to become “more responsible for their own defense.”³

At the end of December 1990, the American Permanent Representative to NATO William Taft proposed to his French counterpart, Gabriel Robin, the “emergence
of a European Pillar in NATO, French reintegration into NATO, and the adaptation of structures." At the time, the French foreign ministry was much more skeptical about American advances than its delegation to NATO, especially because of the preeminence of the U.S.-led Supreme Headquarters of the Allied Powers in Europe (SHAPE) in NATO’s politics. Contrary to the Defense Ministry and Admiral Lanxade, the Foreign Ministry believed that all efforts in creating any European force or defense entity would remain de facto “platonic.”

Instead, the Foreign Ministry advocated for the Western European Union (WEU) – Europe’s own military alliance – to become the European pillar of NATO. Meanwhile, the U.K. and the United States wanted to give WEU a subsidiary role. As this disagreement became prevalent during negotiations, the United States grew less interested in the prospects of reintegration in early 1991. Specifically, disagreements grew over the possibility of a European pillar or entity having forces separate from NATO’s Integrated Command. While the US and UK proposed a separable, yet not separate force, France saw any such ideas as an insincere European pillar. The window for a possible agreement closed at the end of January 1991 when a French Foreign Ministry report mentioned that the “U.S. and U.K. positions do not respond to current challenges.” The lack of substantive convergence on what a transformed NATO would look like finally halted all discussions.

The Treaty of Maastricht in 1992 truly kickstarted the changing French-NATO relationship. Immediately after signing the treaty, the Franco-German Summit of La Rochelle in May 1992 led to the creation of the Eurocorps, a Franco-German military unit, and outlined common goals of the French-German strategic and defense cooperation and integration. These goals provoked discussions regarding Eurocorps’ operational status under NATO – given that the corps was composed of NATO-integrated Germans and non-integrated French. This difference led to a new series of discussions between Admiral Lanxade, his German counterpart, and the Supreme Allied Commander of Europe (SACEUR).

Additionally, despite the Treaty of Maastricht’s promise to build a long-term vision of European defense, President François Mitterrand recognized that NATO would remain the dominant defense organization for the foreseeable future. “European defense,” he told U.S. President George H.W. Bush in the spring of 1991, “would remain ‘virtual’ for some time, whereas NATO would remain durably ‘real.’” At the same time, Admiral Lanxade believed that it was impossible to substitute NATO with a European organization – and instead, the alliance would have to evolve into a European-U.S. partnership.

The Emergence of a Pro-NATO Bloc

In 1993, the right-wing Rally for the Republic party won the legislative elections, placing the French government in a state of ‘cohabitation’ under Mitterrand’s left-wing presidency. The new right-wing government made several changes to French national strategy and its relationship with NATO. By citing their “Gaullist” heritage, the government argued that changes in national security were a mere continuation of Cold War doctrine rather than a shift. The consequent Defense White Paper of 1994 green-lit the participation of the Defense Minister in NATO’s North Atlantic
Additionally, in the Spring of 1993, a French representative was present at the Military Committee for all matters concerning “non-Article 5” activities in which France was involved – a first since 1966. All of these movements represented significant changes: while diplomats had guarded a strict policy of French exclusion since French withdrawal from NATO’s Integrated Command, the military was relatively more “Atlanticist” and desired closer links with other Western militaries. Following these developments, Jacques Chirac won the Presidency in 1995 and resumed efforts to deepen ties with NATO. Chirac was President Pompidou’s protégé and popularly perceived as a champion of modern Gaullism – a perception he particularly pushed during his electoral campaigns. While Presidents Giscard and Mitterrand, the two Presidents following Pompidou, were perceived as a technocratic reformer and a socialist, Chirac on the other hand was the first right-wing Gaullist to be elected since 1974. This success provided Chirac with considerably more credibility in the pro-Gaullist electorate, allowing him to renegotiate the French position in NATO and reformulate Gaullist doctrine.

However, President Chirac was pressed to move ahead with French reintegration into NATO’s command early on due to changing strategic circumstances. Since the beginning of the War in Yugoslavia, the French military proposed a multilateral force that would take the form of the WEU, yet this proposal failed to materialize due to opposition from the U.K. and Germany. The U.K. saw any European force as a French-dominated force, while France saw the U.K. as being too U.S.-aligned. Meanwhile, Germany was too reticent in disrupting the status quo, as it was prioritizing reunification with the East. Thus, President Chirac decided that the question of European defense could only be resolved within NATO as the end of the Cold War changed the geopolitical landscape. According to a senior German diplomatic officer in June 1996, the reason why the Chirac government suddenly became a proponent of NATO reintegration could be traced back to these adaptations in French foreign policy: “French President Jacques Chirac found a way to meld Gaullist national glory with transatlantic cooperation. As he assessed the end of the Cold War, the reunification of a powerful Germany, and the poor performance of the French military in the Gulf War, Mr. Chirac concluded that past French policy aimed at easing the United States out of Europe would be counterproductive.”

Another key reason why Chirac became motivated to increase ties with NATO concerned state finances. His plan for military reform included the abandonment of conscription and a “professional” army based on the British model of an All-Volunteer-Force. Meanwhile, key European governments like France began a “peace dividend” era, focusing on cuts in defense spending. Europeans at the time also braced for the economic impact of the Euro as the common currency, which was to be adopted by 1999. As a result, NATO became an attractive and cheap security provider for Europeans. The NATO Combined Joint Task Forces (CJTF) in Bosnia was perceived in Paris as an efficient force that did not rely on independent equipment, and, as such, adhered less to the Gaullist philosophy of “self-sufficiency” in all weapons.
With these motivations in hand, Chirac began moving toward reintegration. During the December 1995 Ministerial of the North Atlantic Council (NAC), the French Foreign Minister outlined a path ahead for French-NATO relations. The French Defense Minister would regularly attend the NAC, permanently rejoin the Military Command, participate in the NATO Defense College and the Oberammergau Center, and begin a process of elevating relations with SHAPE. However, this change did not include sending its commanders to attend meetings with the Supreme Allied Commander of Europe (SACEUR), the Supreme Allied Commander of the Atlantic (SACLANT), the Commander-in-Chief Allied Command Channel (CINCHAN), and the Canada United States Regional Planning Group (CUSRPG). As a result, French actions resembled a rapprochement, rather than a reentry, into NATO’s Integrated Command.

And this rapprochement was motivated by economic gain and the strategic consequences of European defense building having been stalled yet again in the EU Treaty.

During his first state visit to the United States in February 1996, Jacques Chirac addressed Congress, where he praised American involvement alongside the Europeans in Bosnia. He also declared that NATO “had to adapt itself to a universe that is no longer that in which it was born,” while calling for a stronger European pillar within NATO to better balance the organization away from the United States.

During his speech, Jacques Chirac proposed the European pillar of NATO to be a part of NATO that could be deployed without U.S. resources, including logistics, transmissions, and aerial surveillance.

Chirac and the AFSOUTH Roadblock

However, the following year would be marked by the Allied Forces South Europe (AFSOUTH) crisis, which was critical to the French given their leading role in the war in Yugoslavia and the French fleet’s logistical importance to the operations in the subregion. The French demand to place a European as head of the fleet trumped prospects for a compromise to adapt the structure of the Alliance by boosting the Combined Joint Task Force (CJTF). While the French did not move claims to AFSOUTH, they simultaneously sought to use the CJTF to further Europeanize NATO, reduce the weight of SHAPE, and increase the influence of the military committee. As demonstrated in Bosnia, the lack of clarity behind NATO’s political guidance in non-Article 5 operations necessitated reform of the Command Structure according to SACLENT Shaheen in February 1996. As a result, the efforts related to defining a role for CJTF subsequently transformed into establishing a clearer relationship between political decision-making and deploying troops abroad – which the French were in favor of, according to former senior member of policy planning Gilles Andreani. The effort to establish CJTF as a political instrument rather than just a military instrument was part of broader efforts aimed at Europeanizing the Alliance, a strategy deemed unpopular inside the Pentagon.

In a New York Times Op-Ed in 1996, new French foreign minister Herve de Charette announced that France would be committed to participating in a new NATO in which Europeans would assume burden sharing, and would open to admitting new Eastern European countries while framing a new relationship with Russia. This commitment,
however, was conditional on attributing two regional commands to European officers. According to another speech, this change particularly concerned the southern command due to the vital importance of the Mediterranean Sea, as seen by Paris. The view of a “Europeanized” NATO was also publicly supported by the French Defense Minister Charles Millon in a piece for NATO Review—suggesting a whole-of-government effort to create public support for European defense via NATO at the time. France was not alone on this front: these demands, in fact, were supported by Italy, Germany, and Spain.

In June 1996, the U.K. and the United States agreed on NATO reform, where they “approved the concept of combined joint task forces both to allow non-NATO countries to participate in NATO missions and to permit the use of U.S.-owned NATO assets in WEU-led operations in which U.S. forces would not participate.”

The consequent NATO Berlin Summit’s communique stated that:

“As an essential element of the development of this European Security and Defence Identity we will prepare, with the involvement of NATO and the WEU, for WEU-led operations including planning and exercising of command elements and forces. It would be based on the identification of the types of separable but not separate capabilities that would be required to command and conduct WEU-led operations and an elaboration of appropriate multinational European command arrangements within NATO, consistent with and taking full advantage of the CJTF concept, are able to prepare, support, command, and conduct WEU-led operations. This implies double-hatting appropriate personnel within the NATO command structure to perform these functions.”

At the end of the Berlin Summit, the French foreign minister declared: “France is satisfied. Today it is a new Alliance that is emerging, and in this new Alliance, France is ready to take its full place.”

Despite the optimism, the United States opposed the text and continued to express its opposition to the implementation of the proposals written and signed at the Summit. Nor did the Berlin Summit see an agreement on SACEUR’s Deputy role. This was vital considering that it generated debates around giving major responsibilities to European generals for NATO reform. On the 28th of August 1996, just a month after the Berlin Summit, Chirac sent an official letter to U.S. President Bill Clinton to manage tensions. Chirac requested AFSOUTH to be ceded to a European on grounds that it would create the conditions required for France to enter “a new NATO,” despite the failure of the Berlin Summit. The Foreign Ministry and the Elysée had prioritized the political objective to “Europeanize” NATO to the point that, according to the press, French diplomats regularly interjected their interlocutors by saying “You mean the reformed NATO.”

French military leadership, however, thought first and foremost of national interest and pinpointed the AFSOUTH as a way to leverage key national interests.

In September 1996, President Chirac wrongly assumed the United States was
ready to concede to French aspirations, which was contradicted by a rebuttal letter from President Clinton that same month. However, in this same letter, the American President did propose to create more responsibilities for a European Deputy SACEUR. In essence, President Clinton refused the French proposal out of fear that Congressional opposition would spill over and derail his larger goal of NATO enlargement in Eastern Europe. German Chancellor Helmut Kohl would later support Chirac in a letter directed at President Clinton while noting that allowing for French and Spanish reintegration into the Alliance should be seen as just as much of a priority by the United States as Eastern European accessions. In November of the same year, Chirac met with the NSC members Anthony Lake and Alexander Vershbow to discuss NATO reform and repeated his demands for AFSOUTH. Yet, the NSC representatives maintained their position, despite the French threat to leave negotiations for NATO reintegration altogether. “For us to give up AFSOUTH would be perceived as the beginning of strategic retreat from the new NATO.”

In response to the roadblock on NATO reform, the French began to formulate less ambitious proposals, including a redrawing of the southern region and “a functional division of labor broadly on the lines of the ACE/AFCENT/ARRC (Allied Command Europe/Allied Forces Central Europe/Allied Command Europe Rapid Reaction Corps).” Allies, however, agreed on the need to downsize the number of regional commanders – in support of their own spending cuts as part of the “peace dividends” era. The idea of Europeans maintaining European regional commands was criticized in Washington as an unviable option to maintain an integrated command, and instead, designing common political goals across the Alliance was seen as more sustainable. For instance, as Henry Kissinger put it, “In the end, it is not the integrated command that generates unity, but a sense of shared political and security interests.”

American skepticism of French intentions on NATO reform was further explained by the actions of French diplomacy in the 1990s. According to a senior American diplomat, French representatives had developed a reputation for aggressively advancing their nation’s influence in diplomatic circles, while at the same time assuming greater leadership roles as a consequence of the French-German relationship. Therefore, French advances in the name of Europe were mostly perceived in American diplomatic circles as hypocritical attempts to boost the French role in NATO once they rejoined the integrated command.

During the Brussels Summit of December 1996, the U.S. proposal for a partnership council to replace the Partnership for Peace (PfP) was refused by the French as a way to apply pressure on the AFSOUTH issue after two refusals by President Clinton. Consequently, during a January 1997 summit between NSA member Samuel Berger and his French counterpart Jean-David Levitte, Berger reported that there was “no real movement in the French approach to their reintegration in NATO,” and that their proposal for a solution was still about AFSOUTH, regardless of guarantees of any comfortable leadership position in the new NATO. In an ultimate effort to save reintegration before the Madrid Summit, Jean-David Levitte met
with U.S. Joint Chiefs of Staff Chairman Shalikashvili to discuss a compromise. Levitte proposed that the Europeans would be in charge of “operational” command while the Americans would be in charge of “functional” command. Shalikashvili refused the final offer and proposed to discuss the matter again in six years – in 2003, a year before France rejoined SHAPE. 49

A compromise was finally struck: a European commander in waiting would take charge of the AFSOUTH in six years, until which it would be equal in authority to the current U.S. commander. Despite the reported agreement between Washington and Paris, the deal was soon canceled over political developments in France: Chirac had dissolved parliament to win a majority before the looming recession could impact the election the following year; he lost the election and was forced to accept a socialist government, which staunchly opposed a return to NATO’s integrated command. 50

In June 1997, a month before the Madrid Summit, European defense would see another significant upset during the Amsterdam Summit. Although the French government had been pushing for a deepening of the WEU, the U.K. blocked efforts to link the WEU to the EU because of their traditional policy of supporting NATO as the primary security institution in Europe. 51 As a result, the ambitions of making the WEU a bridge for the EU into NATO died through the U.K. veto. The end of the French effort toward reintegration was thus confirmed by the defense minister in December 1997, where he underlined that France would remain a country in the Alliance but out of the integrated command because of disagreements over shared responsibility of command. 52

Despite this upset, in a speech in 1998, President Chirac articulated his desire to see more integration in the EU on defense issues - while remaining in NATO. 53 Yet, following the Madrid Summit, no further efforts were made to push France back into the Alliance’s integrated command until after the Iraq War. The failure of the Madrid Summit, therefore, became one of the main reasons for President Chirac’s shift after 1997 to seek European defense through the EU rather than through NATO. The Saint-Malo agreement between France and the U.K. further resolved many security issues between both countries, paving the way for a new Amsterdam Treaty without British opposition and adding optimism to the idea that European Defense through the EU could still be possible.

Chirac After 9/11 and The Iraq War

The accession of several Eastern European countries into NATO after the Cold War further impacted the view of Jacques Chirac and other French politicians towards European defense. Paris saw NATO as an institution that would continue to exist despite the end of the Cold War and would continue to shape European defense long after. 54 With the September 11, 2001, Attacks and the 2003 Invasion of Iraq by U.S.-led forces, considerations of what NATO should become resurfaced.

Relations between Paris and Washington had reached their lowest point ever when President Chirac vetoed Turkey’s assistance to the war in Iraq through NATO; called for a “multipolar world”; and proposed creating
an independent European defense institution with Belgium, the Netherlands, and Luxembourg. In Washington, these developments had convinced senior officials in the George W. Bush administration that France was intent on sabotaging NATO and constraining American interests—coincidently at a time when U.S. foreign policy was considerably more hawkish. The war in Iraq created a stark division within Europe as well. France and Germany were alone in their opposition, which led to the perception from Eastern European states that the U.S. was a more reliable “guardian of democracy.” The split over an Iraq Invasion was the first time the EU and NATO did not have a common view, thereby complicating the choice of Eastern Europeans over which stance to take. Unlike the de facto Franco-German leadership of the EU, NATO’s official position on Iraq was a brief statement at the Prague Summit in November 2002, pledging “full support for the implementation of UN Security Council Resolution 1441” and calling “on Iraq to comply fully and immediately with [...] all relevant UN Security Council resolutions.”

Despite diplomatic and political tensions, military relations between France and the United States remained good, notably during their cooperation in the Balkans and other peacekeeping missions. However, punishment by isolation carried out during the Spring of 2003 led to 18 alliance members circumventing France’s veto and giving NATO protection to Turkey through NATO’s Defense Planning Committee, in which France had still no participation. This action was unprecedented in the Alliance’s history.

Consequently, President Chirac allowed massive French participation in both the Allied Command Transformation (ACT) and NATO Response Force (NRF), both newly created in the same year. Furthermore, for the first time since 1966, France tried to appoint two generals to NATO’s International Military Committee but was blocked by U.S. Defense Secretary Rumsfeld. After a meeting between the new French Defense Minister Michele Alliot Marie and Secretary Rumsfeld on January 2004, as well as threats by the Chief of the Armed Forces General Bentégeat to block French dues to NATO, Secretary Rumsfeld finally lifted his opposition.

The two officers appointed to SHAPE eased most of the tensions between the two countries in the months and years that immediately followed. However, they both differed greatly on what the future of the Alliance would be. Disagreements regarded enlargement to Ukraine and Georgia, the expansion of the ISAF’s role in Afghanistan, and the political role of NATO. On the latter, France feared the Americans wanted to transform NATO into an “alliance of democracies” that would compete with the UN. The failure of the vote on the referendum on the Constitution for Europe, in addition to suffering from a stroke, would paralyze President Chirac for the rest of his term until Sarkozy’s win in 2007.

Sarkozy: Return to the Fold

The geopolitical environment of 2008 incorporated many new elements. Nicolas Sarkozy and his “Americo-phile” foreign policy won the Presidency in 2007, a disappointing record of both the Common Security and Defence Policy (CSDP) and
that of the “new” NATO began to emerge, and finally the increasing weight of France in troop contributions in NATO missions – specifically the Kosovo Force (KFOR). President Sarkozy had explicitly announced very early in his term that he would seek to “normalize relations” with the US in 2007 during a conference with French Ambassadors. Sarkozy made clear he had two conditions for normalization: renewing and investing in CSDP and an explicit recognition by Washington of its role, and second, a rebalancing of responsibilities between the United States and Europe inside of the Alliance. In short, advancing European defense, but this time by pursuing it outside of NATO reform.

For Sarkozy, the issue of NATO reintegration was much more about image than other political reasonings, especially because the steepest steps to reintegration were already overcome. But now the greater struggle was about timing: should U.S. recognition of the CSDP come before or after French reintegration? In 2008, after Sarkozy had made a declaration reuttering his precondition for European defense, President Bush recognized that building a strong NATO Alliance also required a strong European defense capacity. President Sarkozy said after the Summit that “this opens the door to renewed ties with NATO.”

President Sarkozy also evoked during an event in 2009 that the attempts made by Mitterrand and Chirac resulted in some progress but “France had not taken its seat in the western family.” When questioned about his precondition on European defense, he bluntly responded: “It’s done.” Given that his administration had chosen to relaunch the CSDP after reintegration, Sarkozy had commented that if efforts didn’t follow, it would only “be defense on paper.” Importantly, the redistribution of roles proposed by the United States, if France reentered, were not the regional ones in Europe, and were given directly to French officers rather than to Europeans in general.

On the 17th of March, the French parliament voted in favor of reintegration with a majority of votes coming from the right, while most of the opposition came from the socialists - who denounced an “alignment” with the United States. The process of “reintegration” began during the Summit on April 2009 at Strasbourg – in a ceremony commemorating the 60th anniversary of the Alliance. As a result of the redistribution of roles, the Allied Command Transformation became French, with a French officer becoming the Commanding officer of NATO’s Response Force (NRF). France, on the other hand, agreed to send 900 officers to SHAPE and to sit on the Defense Planning Committee.

Hollande and Macron: The Return to the EU

The campaign and consequent Presidency of François Hollande, the first left-wing President since Mitterrand, openly questioned the purpose of French reintegration. There were three reasons for his skepticism: “reintegration was considered a symbol of alignment with the United States, there were doubts about the future of the CSDP, (…) and finally, the other fear was linked to the cost-effectiveness of the measure.” Despite an “Atlanticist” wing of the Socialist Party had welcomed Sarkozy’s decision to fully re integrate in 2009, the party held a
convention on the issue ahead of the elections and concluded on a compromise to conduct a study if victorious to determine whether it should continue to be fully reintegrated or reverse its position. Then-candidate Hollande specified that he would demand an enhanced presence within the NATO military structure, as well as significant progress with regard to European defense as conditions for remaining part of the NATO military command.  

Fulfilling his campaign promise, the newly-elected President Hollande then commissioned the former Foreign Minister Hubert Védrine to examine France’s reintegration into NATO’s command structure. Védrine recommended the Elysée to remain in NATO because of the Alliance’s influence over Europe. However, the “Védrine report” also pressed the government to remain “vigilant” by pressing French national interests as well as pushing for European defense.

Following Hollande’s presidency, the Macron-Trump relationship would shape the transatlantic community until 2020. Despite several successful meetings during 2017, tensions began to rise at the first NATO summit of the year, where President Trump called for Allies to devote 2% of their GDP to defense.

Shortly after, in June 2017, Germany and France, under the EU banner, announced forming the European Defense Fund (EDF), where they would pursue joint work on drones, military transports, and combined efforts to stabilize the African Sahel region while remaining in close collaboration with NATO. This effort culminated in a “triangle” on defense issues at the EU level: the Permanent Structured Cooperation (PESCO) formed in late 2017, along with Common Annual Review on Defense (CARD) and the EDF.

In September 2017, in a speech at Sorbonne University, often considered to be the cornerstone of President Macron’s foreign policy doctrine, Macron announced his intent to continue to push for EU integration on defense as an autonomous, but complementary, initiative to NATO.

In direct reaction to these French efforts, a Pentagon official criticized the EU’s CSDP for pulling forces away from NATO. Transatlantic tensions increased further throughout 2018, when the Trump administration irritated some European allies over pulling from the Joint Comprehensive Plan of Action (JCPOA) agreement, criticism over defense spending in NATO, and energy deals with Russia. As the EU began to ramp up the EDF and its other defense initiatives, the United States also voiced opposition to restrictions that prevented non-EU countries from participating in new EU projects. In an official letter signed in early 2019, the U.S. government said that limitations under consideration on the involvement of non-EU members amounted to “poison pills” and hinted at reprisals if European integration in defense was built without consulting the United States.

After facing hostility from Washington, President Macron showed a more aggressive position. During the centenary of the WWI Armistice in November 2018, where Russian President Putin was the only head of state invited, he warned that Europeans could not be protected without a “true, European army.” Whereas many contend that Macron had been responsible for a
seismic shift in Franco-U.S. relations, he did not fundamentally change anything in the relationship other than rhetoric. As seen through the Mitterrand, Chirac, Sarkozy, and Hollande Presidencies, French perception of European defense-building has remained particularly linear since the end of the Cold War. The current President has followed the same initiatives as his predecessors and follows a reactive approach to roadblocks formed by the broader geopolitical environment. When initiatives within the EU encounter a roadblock or geopolitical shifts, French Presidents resorted to building European defense through NATO, and vice versa.

The strategy of shifting international venues to better pursue national goals is called “Forum Shopping,” and the forum’s specific characteristics affect the actor’s goals. Past attempts to build Strategic Autonomy via NATO necessarily incorporated the United States because of its weight in the Alliance. But current attempts via the EU are now distinctively excluding it because of the forum’s nature as a European, and not transatlantic, organization. Strategic Autonomy never reappeared or changed, but its formulation into policy changed because the EU has been a better avenue for defense-building than NATO in the eyes of Paris.

For the first time in 14 years, France took over the Presidency of the Council of the EU in January 2022. Despite the war in Ukraine taking precedence over long thought-over plans, France still brokered a record 130 agreements on a broad series of engagements, notably in common defense and supporting an EU defense white paper — “the Strategic Compass.” France also put into practice one of the mechanisms of the Compass: the European Peace Facility, an off-budget fund worth approximately €5 billion for the period 2021-2027, financed through contributions from EU member states; the facility was used extensively at the beginning of the war to finance munitions for Ukraine. Thus, France successfully led the EU into becoming a net exporter of security in the region for the first time in its history. Instead of representing a new roadblock, the war in Ukraine proved to become an accelerator for European autonomy.

At the same time, however, the Madrid Summit and the War in Ukraine have convinced many Eastern European states to call for additional U.S. presence on Europe’s Eastern flank. As the EU had been ramping up a security role under the French presidency, NATO had also expanded its enhanced Forward Presence to four additional countries and increased the size of its four multinational battlegroups already present. While the EU is becoming more confident in its vision of security, Europe, as a whole, is still an importer of U.S. defense. Coincidentally, President Macron has spearheaded a new initiative since 2022: the European Political Community. It is unclear what role this initiative accomplishes, other than being an alternative “means of structuring the continent.” According to the “forum-shopping” practice demonstrated by French leadership across this paper, this could be the next forum through which French leadership tries to build autonomy, if it grows frustrated by the EU as an avenue of common defense-building.
Conclusion

The French reintegration into NATO’s integrated command appears to be a natural movement that followed key events and concessions in post-Cold War Europe. Unlike Spain and other European countries that had also reintegrated into the Alliance during the same period, the French experience is unique because of its simultaneous attempt to reshape European security. France has never left its foreign policy doctrine, but against multiple roadblocks, Paris adapted its position over NATO and EU defense integration: when one avenue did not work, it resorted to the other. Today, after a perceived failure to move NATO toward a European pillar, French efforts have resorted to the EU to build a ‘Europe of Defense.’ With the Russia-Ukraine war, NATO has returned to the forefront of European defense, and according to this study, France will likely seek to find compromises that maintain NATO’s role as the first security actor on the continent in exchange for further measures that would allow, in the long term, a ‘civil divorce’ from the United States in defense.

About the Author: Paul Cormarie holds a Master of Science in Foreign Service from Georgetown University. He is a former Researcher at NATO Parliamentary Assembly and junior researcher at the Atlantic Council. A portion of this research was undertaken under the guidance of Dr. Sara Moller, Professor at Georgetown University.
1 Record of conversation between Mikhail Gorbachev and Francois Mitterrand (excerpts). [Link]
3 Frédéric Bozo, p. 425.
4 Frédéric Bozo, p. 424.
5 Frédéric Bozo, p. 426.
6 Frédéric Bozo, p. 428.
7 Statement on the establishment of a Franco-German European Army Corps (La Rochelle, 22 May 1992). [Link]
11 Anand Menon, p. 60.
14 Jacques Lanxade, p. 343
15 François Huguenin, ”Jacques Chirac était-il gaulliste?” Le Figaro, 4 October 2019. [Link]
17 Pichovszky Domonkos, p. 40.
21 Ronald Tiersky, p. 3.
22 Ronald Tiersky, p. 3.
24 Pichovszky Domonkos, p. 41.
26 JFC Naples, “Brief Overview,” [Link]
28 Douglas Berenson, “As France Explores Full Reentry to NATO . . .: FRANCE SEES UTILITY IN GEN. SHEEHAN’S NATO REFORM PROPOSAL.” Inside the Pentagon 12, no. 6 (1996): p.16. [Link]
29 Douglas Berenson. p. 16.
31 NATO, « Intervention du Ministre des Affaires étrangères de la France, Hervé de Charette, devant le Conseil Atlantique. » [Link]


37 Pichovszky Domonkos, p. 43


40 Michael Brenner and Guillaume Parmentier, p. 56.

41 Gilles Delafon and Thomas Sanoton, p. 214.

42 Gilles Delafon and Thomas Sanoton, p. 217.

43 Declassified documents, p. 193.

44 Michael Brenner and Guillaume Parmentier, p. 55.


46 Gilles Delafon and Thomas Sanoton, p. 196.

47 Gilles Delafon and Thomas Sanoton, p. 257.

48 Declassified Documents. p. 208

49 Gilles Delafon and Thomas Sanoton, p. 268

50 Gilles Delafon and Thomas Sanoton, p. 285.


56 Tassos E Fakiolas, p. 300.


58 Henri Vernet and Thomas Cantaloube, p. 319.


60 Henri Vernet and Thomas Cantaloube, p. 319.

61 Henri Vernet and Thomas Cantaloube, p. 321.


64 Dominique David, p. 437.

65 Dominique David, p. 440.


https://www.lemonde.fr/international/article/2009/03/12/otan-les-justifications-de-m-sarkozy_1166809_3210.html.

69 Laurent Zecchini, Le Monde.


73 Jean-Pierre Maulny, p. 10.


76 “Germany, France drafting details of defense fund: German minister,” Reuters, June 10, 2017. 


78 “Sorbonne speech of Emmanuel Macron - Full text / English version” Ouest France, September 26, 2017. 


83 Hannah Murphy, Aynsley Kellow, “Forum Shopping in Global Governance: Understanding States, Business and NGOs in Multiple Arenas” Global Policy, Volume 4, Issue 2, May 2013, pp. 139-149.


86 “NATO’s military presence in the east of the Alliance,” NATO, December 21, 2022. 

Reframing Proxy War Thinking: Temporal Advantage, Strategic Flexibility, and Attrition

Amos C. Fox

Despite superb efforts by a determined group of post-9/11 theorists, proxy war theory has stagnated in recent years. Most modern proxy war theory centers on the idea that actors use proxy strategies because those strategies allow the actor to maintain and launder their involvement behind a number of fronts, including intermediary forces and client governments. Further, by operating behind a façade, the proxy-dependent actor’s strategy generates varying degrees of reasonable refutability regarding their involvement in the conflict. These two basic tenets of contemporary proxy war theory can be reduced to the ideas of indirect control and involvement, and plausible deniability.

Proxy wars in the post-9/11 period, however, have shown that the two core tenets of indirectness and plausible deniability no longer carry proxy strategy’s water. The Russo-Ukrainian War provides an excellent case study to refute indirectness and plausible deniability’s primacy in proxy war. Temporal advantage and strategic flexibility rise to supersede indirectness and plausible deniability as the two core tenets that drive why actors elect to conduct proxy strategies. Resultantly, proxy war theory should begin to use temporal advantage and strategic flexibility as its basic lens for making sense of modern, and future, proxy wars.

Lastly, proxy wars are inherently deadly and destructive, despite the real, and perceived, benefits afforded to the principal actor in principal-proxy dyads. The Russo-Ukrainian War, for instance, clearly illustrates that proxy wars are just as many wars of attrition, as they are anything else. Therefore, to appropriately account for the reality of proxy wars, and not remain anchored to feel-good conceptual myths, theorists must update their theories and analytical lens to account for the proxy war’s association with wars of attrition.

Introduction

Russia began a long-running proxy war in Ukraine in the spring of 2014. The proxy war began with Russian forces, and Russian proxies, annexing Crimea, and large swaths of land in Ukraine’s Donets River Basin (Donbas) region. Cynically attuned to the international community’s values and norms, Russia’s proxy strategy relied on both a political movement and a military action. Following several evolutions, Russia’s two political proxies in the Donbas became known as Donetsk People’s Republic (DPR) and Luhansk People’s Republic (LPR), while the ostensible Republic of Crimea reflects Russian proxy governance. Similarly, Russia’s military proxies in the Donbas congealed into entities collectively referred to as the Donetsk People’s Army (DPA) and Luhansk People’s Army (LPA). As integrated elements of the Russian Army, the DPA and LPA are also often referred to as the 1st Army Corps and 2nd Army Corps, respectively.1 Russia used its Donbas proxies to undercut Kyiv’s power, control, and influence in eastern Ukraine, while weaponizing governance, citizenship, and visas to Russify the region.

Militarily, the Kremlin used the DPA and LPA to make territorial gains in the Donbas between the spring of 2014 and early 2015, capturing significant portions of sovereign Ukrainian land.2 Moscow put a Ukrainian face on its aggressive behavior in the Donbas by using the DPR and LPR to parry
the international community’s concerns regarding Russian meddling. Once overt hostilities ended in February 2015 with the Minsk II Agreement, the Donbas proxies insulated Moscow from outright international condemnation by providing a veneer of political legitimacy, regardless of the Kremlin’s duplicity in Ukrainian domestic politics. Russia relied on the DPR and the LPR to govern its annexed regions in Donetsk and Luhansk and it used the DPA and LPA to maintain a credible deterrent along the contact line with the Ukrainian armed forces.3

Akin to an untreated ailment, the Donbas proxies – a Kremlin Trojan Horse, of sorts – served as a guise and policy time-bomb for Moscow, while presenting Kyiv with a dubious conundrum. Left alone, the Donbas proxies might provide Moscow with the excuse it needed to initiate a full-scale invasion at an undetermined future date, which hindsight has proven correct.4 Yet, if attacked militarily, the Kremlin might break from the obfuscated use of force vis-à-vis its proxies, and instead counter with a full-scale invasion with conventional forces, which happened through late 2014 and early 2015. Given the drubbing Ukraine’s professional and volunteer forces took during the 2014-2015 Donbas campaign, inviting another conventional Russian assault appeared as a non-starter for Ukrainian policymakers.

The Kremlin’s Trojan Horse policy came full circle in February 2022 when Russian military forces invaded Ukraine yet another time in the two states’ long, intertwined history. The Donbas proxies formed the crux of Russia’s justification for invading Ukraine. To be sure, Russian President Vladimir Putin repeatedly cited the need to defend the DPR and LPR against Ukrainian oppression as the impetus for armed conflict, resulting in Russia’s instigation of the largest European armed conflict since World War II.5

The outcome of Russia’s Ukrainian policy gambit remains to be seen. A detailed examination of Russia’s proxy strategy can nevertheless illuminate many blind spots in existing proxy war literature, conflict studies, and international relations theory. Russia’s reliance on the DPR and LPR is relatively straightforward as it relates to political proxies – Moscow used intelligence and special forces operations to identify sympathetic individuals in eastern Ukraine, and then use those individuals to manufacture de facto regional governments in Donetsk and Luhansk. Russia’s reliance on the DPR and LPR (i.e., the political entities) aligns neatly with most of proxy war theory.

Russia’s use of proxy military forces nonetheless highlights a lacuna in proxy war theory. Most proxy war theory values indirectness and plausible deniability as keystone concepts. In most proxy war theory, indirectness is often referred to as Actor A’s use of an intermediary (Actor B) to combat a third-party (Actor C).6 Moreover, proxy war theory suggests that in using Actor B, Actor A attempts to avoid direct, or physical, confrontation with Actor C.7 Plausible deniability, on the other hand, is the theory that Actor A can purposely generate doubt in external actors, and the international community, regarding their involvement in an armed conflict against Actor C.

A cursory examination of the Russo-Ukrainian War (2014 – present), however, discredits indirectness and plausible deniability’s recognized importance in proxy wars and proxy war theory. Contrary to contemporary theory, the Russo-Ukrainian War, and Russia’s use of proxies throughout the conflict, point to the fact that temporal
advantage and strategic flexibility are of strategic importance in proxy war.

In this paper, I examine how Russia’s proxy war strategy confounds contemporary proxy war theory by not adhering to the tenets of indirectness and plausible deniability. I analyze this problem by briefly examining a few of the leading voices in proxy war theory to emphasize the significance they place on indirectness and plausible deniability. As part of that review, I introduce definitions for time and strategic flexibility that are pertinent to proxy war theory. Next, I analyze the Russo-Ukrainian War to identify the importance, or lack thereof, that Moscow placed on indirectness, plausible deniability, time, and strategic flexibility. I then close with what these findings might portend for constructing a more robust and practical theory of proxy war.

I make four basic arguments in this paper. First, most proxy war theory incorrectly categorizes proxy conflict as the indirect activity of one actor, through another actor, against a third party. As my case study will demonstrate, however, indirectness was of passing importance to Russia’s proxy strategy in Ukraine. Moreover, indirectness is an incompatible term vis-à-vis strategic competition and armed conflict. By virtue of any strategic actor employing a military strategy in which they are in any way complicit, they are directly involved or directly engaged in the conflict. What many theorists mean when using the term indirect, or indirectness, is obfuscation, or an actor’s attempt to keep their involvement hidden.

Second, the Russo-Ukrainian War illustrates that plausible deniability is useful insofar as it provides a temporal strategic advantage. Obfuscation in proxy war matters only during the initial phase of a proxy war, or when the purveyor of a proxy strategy is attempting to surprise its opponent or trigger a fait accompli. Once an actor’s ability to stay a step ahead of its adversary, or the international community, has dissipated, plausible deniability is no longer an important or useful concept in proxy war theory. Thus, instead of speaking about the alleged importance of plausible deniability, proxy theory must emphasize the importance of temporal advantage to the purpose of a proxy war strategy.

Third, strategic flexibility, along with temporal advantage, is the second true reason proxy wars dominate strategic competition today. Proxy strategies provide strategic flexibility by offloading the political and domestic risks and costs an actor incurs when using their own armed forces in armed conflict. Further, when an actor uses proxies in conjunction with their own forces, it increases their strategic flexibility by creating a more robust and resilient force, which allows it to endure more readily the rigors of combat while continuing toward its political-military objectives.

Fourth, untreated external proxies (political and/or military) operating within another state’s sovereign borders create a security conundrum for both parties. The external proxy’s principal will likely use the proxy to create chaos within its opponent’s borders. Over time, the principal will likely use its opponent’s treatment of the proxy – regardless of reality – to justify hostile action at a later date. However, if the opponent attempts to excise the external proxy, that might invite hostile action from the principal.

Finally, proxy wars increase what can be characterized as attrition in armed conflict. Attrition increases in proxy war because the principal actor in a principal-proxy dyad is fundamentally disinterested, or less
interested, in the survival of its proxy than it would otherwise be if its own forces were at work. Conversely, the opposing actor is interested in its own survival and will attempt to eliminate the proxy no differently than it would the principal’s own forces.

Moreover, proxies, such as the DPA, tend to be the lead forces in increasingly hazardous combat situations, and as a result, they suffer high casualty rates. David Axe, a correspondent with Forbes, reports that the DPA began the February 2022 campaign with roughly 20,000 soldiers organized into six brigades. By November 2022, the DPA incurred over 19,000 casualties – wounded and killed in action. As Axe flatly states, the DPA are used like cannon fodder.

I conclude this paper by making several recommendations to improve the ontology of proxy war. First, proxy war theory should begin to cashier its reliance on the ideas of indirectness and plausible deniability as its core tenets, and instead exchange those ideas with temporal advantage and strategic flexibility. Next, proxy theory should clearly communicate its relationship to attrition so that policymakers, academics, and practitioners are not caught off guard by the high casualties and collateral damage associated with employing proxy strategies. Lastly, proxy war theorists must begin to push their arguments from behind the dark confines of academia and engage on a broader front, including government policy and military circles, so that their important work can make a meaningful impact on the true praxis of war and warfare.

Research Methods and Limitations

My focal point – examining the efficacy of indirectness and plausible deniability in proxy war – necessitates using theory testing as this paper’s theory-building research objective. To that end, indirectness, plausible deniability, temporal advantage, and strategic flexibility are the independent variables I use to examine contemporary proxy war theory. The Russo-Ukrainian War serves as the case study I use to examine my independent variables and to support my new theoretical contributions to the general theories of proxy war and conflict studies.

The challenge in writing about the Russo-Ukrainian War is that much of the policy and strategy documents, plans, and other essential primary source information is still behind the secretive confines of each involved actor’s respective information classification protocols and networks. As a result, I have defaulted toward secondary source information as the main documentation to support this work’s case study. Acknowledging that bias is inherently built into my secondary source material, I do my best to parse that bias and focus on the verifiable information therein.

Press releases and other official government information are also handled with caution. This is because many of these documents serve dual purposes – not only are they to inform the public, but they are also often a tool of messaging strategies in which an actor attempts to advance its preferred narrative. As a result, I carefully evaluate the official government press releases before using them as a reference within this paper, regardless of the actor. In situations where the reliability of a document cannot be confirmed, it is discarded as a source.

Additionally, some primary source material within this paper is known to be patently inaccurate, or false. For example, I use Vladimir Putin’s speech and treatise on Russo-Ukrainian history as a source within this work, despite the known factual inaccuracies therein. I consciously do this, however, because the falsity of Putin’s
message is an important feature in the narrative that Putin and his attendants used to build their case for war with Ukraine. Therefore, it is a critical lie that helped pave the path from the \textit{fait accompli} seizure of Crimea in the spring of 2014 to today.

Lastly, I use only one case study to support the argument that temporal advantage and strategic flexibility are the true benefits of proxy strategies. In most analytical studies, one data point might skew the data. However, in this situation, the Russo-Ukrainian War has been running in real-time for over eight years. Therefore, in my judgment, the conflict’s duration is a sufficient substitute for multiple data points and thus legitimizes the research findings herein.

**Case Study**

Russia’s use of proxies against Ukraine reflects Moscow’s understanding of the rules-based international order. Instead of Russia playing by the international community’s accepted rules, however, Russia simply used those rules as a handrail to provide a façade of legitimacy to illegitimate action. This case study validates the claim that Russia took advantage of the rules-based international order along three primary lines. First, the Kremlin weaponized social factors, such as ethnicity, language, religion, and a loose interpretation of history to both legitimize its actions, but to also cultivate and manufacture proxy movements in Crimea and the Donbas.\(^{10}\)

Second, Russia’s proxy strategy was not purely military. Aware that condemnation awaits those who operate outside the bounds of legitimacy, Russia used both political and military proxies in Crimea and the Donbas to maintain the appearance of political legitimacy.\(^{11}\) The combination of political and military proxies allowed Russia to point to the political proxy as the representative for repressed Russian minorities in both regions. In turn, this provided Vladimir Putin with the political cover, regardless of its dubiousness, to justify the action of its Ukrainian proxy strategy.

**Crimean and Donbas Campaign (Spring 2014 – Spring 2015)**

In February 2014, Russian President Dmitry Medvedev and Prime Minister Vladimir Putin authorized the annexation of Crimea.\(^{12}\) The use of force relied on the covert insertion of troops, proxies, and political figures ahead of time so that once the operation was given the green light, the results were a forgone conclusion.\(^{13}\) As will be seen more clearly in the Donbas, Russia manufactured an insurgency to help obfuscate Russian involvement and lend a degree of legitimacy to its action.

On 27 February, Russia moved forward with its plan to annex Crimea. A mix of unmarked Russian forces and proxy forces jumped into action, seizing the peninsula largely through \textit{fait accompli}.\(^{14}\) Aware of the international communities’ sensitivities to political legitimacy, Russia quickly installed a proxy government in Crimea.\(^{15}\) Russia’s proxy handlers on the peninsula relied on transactional ties to govern the relationship with its contractual proxies, including the notorious Wagner Group. Further, Russian proxy handlers balanced transactional ties and kinship ties to govern the relationship with its political and military proxies on the peninsula.\(^{16}\) Based on the close ethnic link between Russians and the population of Crimea, the bond between the Kremlin and their Crimean proxy government was relatively strong and resulted in low agency costs for Moscow’s cultural and contractual proxies.\(^{17}\)
Following the success of the Crimean operation, Putin sanctioned covert operations in eastern Ukraine. A smattering of contractors, intelligence officers, and special forces led the effort to manufacture an insurgency in eastern Ukraine. A manufactured insurgency differs from a genuine insurgency in a few meaningful ways. First, a manufactured insurgency is a planned and curated response to an outside actor’s political and strategic goals, whereas traditional insurgencies are often an organic response to deteriorating internal political, domestic, or economic conditions. Second, manufactured insurgencies are a front for the intervening actor, but the insurgency is meant to appear as though it is the result of genuine internal strife. Language can help make this point apparent. The Donbas proxies were commonly referred to as ‘separatists,’ especially in the early post-2014 period, despite being a manufactured insurgency, composed of a variety of proxies, as part of the Kremlin’s strategy relating to Ukraine.

Moving from policy to the applied level, Russia hinged its strategy on disaffected locals, Russian volunteers, and contractual proxies to cobble together the political leadership and military forces required to bring its manufactured insurgency to life. In the Donbas, the manufactured insurgency materialized as a small, but formidable, proxy army in Donetsk and Luhansk. In many cases, Russian army officers and reliable Wagner contractors commanded Moscow’s Donbas proxy force, while local commanders possessed only nominal control.

The establishment of the DPR and the LPR came quickly on the heels of the DPA and LPA’s formation. Through April and May of 2014, the DPA and LPA infested Donetsk and Luhansk oblasts, carving out significant holdings in each locale. In Donetsk, Russia’s proxy land grab stretched as far west as Slovyansk, and Lysychansk in Luhansk oblast. In total (i.e., the Crimean annexation plus the Donbas annexation), the Kremlin’s proxy strategy acquired nearly ten percent of Ukrainian territory.

Analyst Mark Galeotti notes that the Russian border town of Rostov-on-Don served as the springboard for Russia’s proxy strategy in the Donbas. Rostov served as the primary logistics node supplying the DPA, the LPA, and their Russian handlers. Rostov also served as a recruit and training depot where volunteers and contract fighters were screened and trained before being sent to the front to link up with their respective units.

In May 2014, the DPR and LPR declared their independence from Ukraine and the government in Kyiv. Shortly thereafter Kyiv’s forces stepped up their effort to both retake the lost territory, defeat Moscow’s proxy forces, and evict the Kremlin’s proxy governments. Notably, Ukrainian forces took the Donetsk Airport in late May and were chipping away at retaking Slovyansk, Kramatorsk, and other towns through the late spring and early summer. By early July 2014, Ukraine had regained its strategic balance and was methodically pushing Russia’s proxies toward the Russia-Ukraine border.

Recognizing that the balance was shifting in favor of Kyiv, the Kremlin authorized overt participation by Russian land forces in the conflict. Russia moved upwards of 40,000 soldiers close to its border with Ukraine. The 11 July Russian rocket strike on Ukrainian forces at Zelenopillya was the first significant indication that the Kremlin’s proxy strategy had shifted from one of obfuscation to one of speed and flexibility. By August 2014, Russian land forces were unabashedly participating alongside the...
DPA and LPA in combat. Unanimous Russian victories at the battles at Ilovaisk and Luhansk Airport are the most well-known examples of Russian army participation during the summer of 2014 because of the visibility surrounding the movement of the hundreds of vehicles and thousands of soldiers that deployed from Rostov-on-Don to the Donbas. Russian land forces, in turn, helped stave off proxy defeat and to preserve proxy territorial holdings in the Donbas. The summertime defeats at Luhansk Airport and Ilovaisk brought Kyiv to the negotiating table and resulted in the ill-fated, and short-lived, Minsk ceasefire agreement.

Nevertheless, before the ink had dried on the Minsk Protocol, Moscow doubled down on its Donbas strategy and initiated the brutal battle for Donetsk Airport in September, commonly referred to as the Second Battle of Donetsk Airport. After four months of combat, which destroyed the airport, Ukrainian forces withdrew from the area and set up defensive positions along the city’s outskirts. Following the victory at Donetsk Airport, Russia shut the door on the Donbas campaign with a decisive military victory at Debaltseve, which paved the wave for the Minsk II agreement.

Throughout the Donbas campaign’s fall and winter legs, the Kremlin used Russian army officers to command many of its Donbas proxy units due to the proxy’s high agency costs. To be sure, the Russian military found the Donbas proxies unreliable, which generated high agency costs for the Russian armed forces. Further, the Donbas proxies’ high agency costs forced the Russian military to place regular army units alongside the DPA and LPA because the Kremlin lacked confidence in the DPA and LPA’s ability to conduct unilateral or complex military operations. In most cases, a shadow Russian command structure paralleled, and guided, the proxy’s unit commanders. Russian balanced legitimate and coerced power to manage its relationship with the DPR/DPA and the LPR/LPA during this period by tightly coupling its own forces, as well as combat and political advisors, with its proxy.

The Kremlin’s relationship with the Wagner Group, however, was distinct from that with the Donbas proxies. The Kremlin used the Wagner Group solely for military operations. The Wagner Group, unlike the Donbas or Crimean proxies, played no part in governing Russia’s confiscated territory in Ukraine. As a result, Moscow’s relationship with the Wagner Group was contractual, which relied on a transactional bond and reward power to regulate the relationship. Because of the contractual nature of the relationship between the Wagner Group and the Kremlin, the bond between the two was strong, and Wagner’s embrace of risk was (and remains) very high. Considering those factors, Wagner came with low agency costs which provided the group with a significant degree of autonomy. Because of the Wagner Group’s low agency costs, high morale, and financial incentives, it was capable of unilateral operations of increasing difficulty.

Post-Donbas Campaign (March 2015 – July 2020)

In the intervening years, Moscow relied on its governmental proxy, the DPR and LPR, and their cultural symmetries with the Donbas’ residents, to create a politically and domestically advantageous situation for future political and military action against Ukraine. During this period the Kremlin weaponized citizenship by providing passports and visas to the residents of Donetsk, Luhansk, Crimea, and other areas in eastern and southern Ukraine. This move sought to solidify its political position.
in the Donbas and manufacture a sympathetic domestic audience in occupied and non-occupied areas. Moreover, hindsight proves that this move was used to create additional pockets of ‘Russians’ in Ukraine which Moscow would use as a pretense to justify its February 2022 invasion.

From Minsk II through February 2022, sporadic conflict between the Donbas proxies and Ukraine’s armed forces flared up, but generally garnered scant interest outside of the region. Russia continued to use the Donbas proxies as a thorn in Kyiv’s side, but also as a strategic Trojan Horse. Using the inordinate number of Minsk ceasefire violations levied by the DPA and the LPA as an indicator of utility, the DPR and LPR provided Moscow with a useful vector to manipulate Kyiv and the domestic situation across the region.\(^4\) The DPA and LPA also provided the muscle to enforce the DPR and LPR’s dictates, as well as deter any Ukrainian plans to retake the Donbas. Similarly, Crimea’s Russian proxy government kept the Kremlin’s policy at the fore of operations, while traditional Russian military forces garrisoned the peninsula.

Russia likely used this time to normalize Crimea’s annexation and its occupation of the Donbas with the international community. As a sign that the normalization effort was working, technology giant Apple changed Crimea’s designation from Ukrainian to Russian on all its map applications in April 2019.\(^4\) This explicit acknowledgment of Moscow’s hegemony over the peninsula lasted until shortly after Putin’s February 2022 invasion, after which time Apple quietly changed Crimea’s affiliation back to Ukraine.\(^4\)

The Trilateral Contact Group (TCG) meeting in Minsk in July 2020 was another clear sign that the international community was gradually accepting Ukraine’s territorial losses and the primacy of Russian proxy governance in the region. During the meeting, the participants agreed on several key points. First, the DPR and LPR would remain the de facto governments in the occupied region – an unmitigated strategic win for Russia’s proxy strategy.\(^4\) Second, the existing contact line would remain the demarcation line between sovereign Ukrainian territory and the occupied regions.\(^4\) Third, Kyiv agreed to not attack across the demarcation line and attempt to retake its disputed territory.\(^4\) In effect, the TCG’s July 2020 agreement all but ceded the occupied portions of the Donbas to the DPR and LPR, and in reality, to Russia’s sphere of influence – another colossal strategic victory for Moscow’s proxy strategy in Ukraine.

**Renewed Hostilities (July 2020 – Present)**

Not satisfied with its gains from the July 2020 TCG agreement, the Kremlin used the DPR and LPR as a pretense for further escalation. In late 2021 and early 2022, the Kremlin began to amplify Putin’s Ukraine narrative.\(^4\) Putin and his acolytes routinely stated that Ukraine was not a separate or independent state, but instead a lost region of Russia. Moreover, Putin and his retinue forcefully asserted Kyiv was discriminating against Ukraine’s ethnic Russian population.\(^4\) Consequently, Putin argued that as the hegemon of the Slavic world and protector of ethnic Russians and Russophones, Russia was obligated to intervene on behalf of the persecuted Russians in Ukraine.\(^4\) Keenly aware of the international community’s aversion to interstate armed conflict, Putin cited the United Nations charter in his formal declaration of ‘war’ (though he stylized it as a *special military operation*) to legitimize his call for war.\(^4\) Putin specifically referenced the charter’s Chapter VII, Article
51, which alludes to a state’s right to conduct self-defense measures necessary to maintain peace and security.

Putin’s 23 February declaration of war opened the belly of Moscow’s Trojan Horse, bringing the Kremlin’s proxy strategy in Ukraine full circle. The DPR and LPR were the fulcrum to Moscow’s central argument for war and the DPA and LPA proved the anchors upon which Russia’s operational plan pivoted. Nonetheless, the Donbas, replete with a host of Russian political and military proxies, also served as the springboard for Moscow’s failed February 2022 assaults on Kyiv and Kharkiv. Meanwhile, Russia’s successful thrust during the spring and summer of 2022 to create a land bridge to Crimea ran directly through the Donbas and linked its proxy governments in the region with those of Crimea. In effect, the DPR and LPR formed a chain link of proxy governments and proxy military forces from the Russian-Ukrainian border in Luhansk to Crimea, running all along the Azovian coast.

The DPA and LPA, for their part, played central roles in the combat that allowed Russia to create its land bridge to Crimea. The DPA was front and center in the struggle to capture Mariupol, serving as one of Russia’s front-line organizations throughout the battle. In the wake of the siege of Mariupol, for instance, the DPR’s president Denis Pushilin toured the front and provided combat awards to members of the DPA, including the infamous Somali Battalion.

After capturing the preponderance of Kherson and Zaporizhzhia oblasts, placing the legality of the issue aside for the moment, Russia’s annexation of Donetsk, Luhansk, Kherson, and Zaporizhzhia constitutes the unprovoked seizure of approximately 90,000 square kilometers, or 15% of sovereign Ukrainian territory, all spurred by a proxy strategy which maximized the potential of coerced proxies, contractual proxies, and proxy governments.

In the ensuing weeks, the United Nations parried Russia’s nefarious use of international law to legitimate its action in Ukraine. On 12 October 2022, the UN countered, stating that Russia’s reliance on Chapter VII, Article X of the UN’s charter mischaracterized the charter’s purpose and intent, and the body urged its members to not vote in support of Russia’s illegal annexation of Ukrainian territory, which it deemed felonious. The measure passed by a vote of 135 members in favor of the measure, five voting against it, and 35 abstentions.

In the war’s ensuing period, Russia used its Donbas proxies as a blunt force tool as part of its attrition strategy in Ukraine, serving as a bite-and-hold force in Donetsk and Luhansk, and a battering ram in places such as Mariupol and Bakhmut. On the governance front, the Kremlin continues to rely on the DPR and LPR as fronts to govern its annexed territory to provide a veneer of international legitimacy to its unofficial occupation of the Donbas.

Wagner, Moscow’s contractual proxy, on the other hand, has been used more broadly than the DPA or LPA. Russian operational commanders have used Wagner to both supervise and support the DPA and LPA in the Donbas. Beyond working alongside the DPA and LPA, Wagner proxies have played a significant role in augmenting Russian
land forces by spearheading combat in Kyiv, Kharkiv, Mariupol, and the Donbas.  

Further, Russia began the war in February 2022 with approximately 190,000 soldiers, oriented along three major attack axes. The DPA and LPA provided Russia with an additional 10,000 soldiers and a ready-made command structure. Wagner’s numbers are far less easy to discern. Nonetheless, reports indicated that Wagner started the conflict with approximately 1,000 fighters. By December 2022, reports indicate that Wagner has upwards of 20,000 fighters working alongside the Russian army. If those numbers are correct, Wagner now makes up roughly twenty percent of Russia’s fighting force in Ukraine.

Russia eclipsed 125,000 soldiers killed in action in January 2023. It is unclear if those numbers reflect losses across the Russian military, including the DPA and LPA, Wagner Group, and lesser groups. Nevertheless, the Russian military does not report Wagner losses to its public the way that it does with service members. As a result, contractual proxies, such as Wagner, provide a mass, attritable army that can plunge into destructive combat without fear of domestic and political unrest at home.

To be sure, in the wake of stalwart Ukrainian defense against Russia’s Wagner, in turn, provides the Kremlin with a proxy actor that provides it both time and strategic flexibility to circumvent the slow, bi-annual mobilization and training of conscripts, and the strife associated with losing uniformed service members. The Kremlin’s use of Wagner throughout the Russo-Ukrainian War, from 2014 to the present, provides an interesting wrinkle to proxy war that is often overlooked in the literature – plausible deniability and indirect involvement take a back seat to time sensitivity and strategic flexibility.

Finding: Indirectness in Proxy War

Indirectness engagement and indirect control reside at the heart of contemporary proxy war theory. For instance, Andrew Mumford, one of proxy war theory’s prominent contributors, predicates his definition of proxy war on the idea of indirectness. Mumford states that “Proxy wars are defined…as the indirect engagement in a conflict by third parties wishing to influence its strategic outcome.” Additionally, he contends that proxy wars are the byproduct of armed conflict in which the principal actor does not directly commit its own military force. Mumford also summons the ghost of B.H. Liddell Hart by creating a theoretical link between the purported salience of indirectness within proxy wars to Liddell Hart’s idea of the indirect approach. Mumford contends that proxy warfare is one of the leading forms of indirect warfighting within a “New era of the indirect strategic approach.”

Indirect engagement and indirect control are also central in Eli Berman and David Lake’s perspective on proxy war. Berman and Lake assert that state actors engage in proxy wars through the indirect control of non-state, and sometimes state, actors. Further, proxy wars provide state actors with the means to indirectly control political violence on behalf of their strategic interests and policy aims.

Frank Hoffman, a heavyweight in the study of armed conflict, and Andrew Orner use similar language discussing proxy war, describing the use of proxies as an indirect approach to fighting wars. Hoffman and Orner continue, stating that proxy wars will continue to be significant features of strategic competition in the future because of the supposed value of avoiding direct action by supporting intermediaries on the ground.
Candice Rondeaux and David Sterman, from the U.S. think tank *New America*, also base their portrayal of proxy war on indirectness. Rondeaux and Sterman state that “Proxy warfare is best defined as the direct or indirect sponsorship of third-party conventional or irregular forces that lie outside of the constitutional order of states engaged in armed conflict.”

What each of these theorists fails to account for, and the Russo-Ukrainian War makes clear, is that obfuscated participation in a conflict is useful so long as it provides advantages of time. Temporal advantages include the positive impacts an actor generates from strategic operations that occur ahead of an opponent, and/or the international community’s, ability to recognize, process, and respond to the first-mover use of proxy forces.

When the advantage of time passed, the purveyor of the proxy strategy is faced with a decision – carry on, despite their involvement being known; or discontinue their involvement. The Russo-Ukrainian War case study illustrates that once the temporal advantage of a proxy strategy dissipates, and obfuscation evaporates, carrying on with a proxy strategy still yields benefits because of the strategic flexibility provided by operating through intermediaries.

Moreover, the phrases ‘indirect involvement’, ‘indirect engagement’, and ‘indirect sponsorship’ do not clarify proxy war dynamics. These terms may, in fact, confuse policymakers, academics, and practitioners looking to understand and employ proxy strategies. By virtue of being involved at all, strategic actors are directly engaged in a proxy war. Indirectness, when the term is used, actually alludes to varying degrees of obfuscated participation. This distinction might seem metatheoretical and pedantic, but the nuance is important.

**Finding: Plausible Deniability in Proxy War**

In addition to indirectness, plausible deniability is a core concept in contemporary proxy war theory. Mumford asserts that modern proxy wars reflect a principal actor’s desire for plausible deniability and lower political and domestic risk.

Strategic actors also use proxy strategies to create time asymmetry between them and their adversaries. Preeminent military theorist Robert Leonhard opines that:

> “Time… is the first and primary dimension that commanders and leaders have had to struggle with since the dawn of history… Time pervades all decision-making in war… Time comes before, follows after, and orders the sequence and tempo of military operations.”

Time asymmetry allows an aggressor to move toward its strategic objectives at a faster clip than their opponent. Time is so important in war and warfare that Napoleon Bonaparte famously stated: “I may lose ground, but I shall never lose a minute… ground we may recover, time never.”

Think of time asymmetry as the aggressor operating at full speed, while the opponent is reactionary and operating at half-speed, at least in the preliminary period of a proxy war, because of the cognitive and decision advantages afforded to the aggressor by obfuscated involvement. The Crimean campaign and the Donbas campaign’s spring offensive also illustrate how time asymmetry allows the aggressor to gain strategic objectives at a faster rate than their opponent.
Modern proxy war theorists, however, seem stuck wrestling with first-order questions and dealing with topical issues which stand in the way of developing a tangible understanding of proxy wars. This development is important because a richer understanding of proxy war can prevent policymakers, academics, and practitioners from being oblivious to the range of options proxies provide an actor with hostile intentions buried within their foreign policy. The vast numbers of people, institutions, and governments issuing *mea culpas* for failing to connect Russia’s Crimean and Donbas campaigns with larger foreign policy goals in the wake of Russia’s reinvasion of Ukraine in February 2022 reinforces this point.76

**Conclusion**

Theory testing and analyzing proxy war concepts through the lens of the Russo-Ukrainian War through the lens of proxy war provides many insights that are both overlooked in proxy war studies. For starters, the conflict demonstrates that indirect involvement, indirect control, and plausible deniability played almost no part in Russia’s use of proxies throughout the conflict. Instead, the war demonstrates that a proxy strategy’s usefulness correlates to the dependency between a strategic actor and three elements – obfuscation, time, and opportunity. Strategic actors use obfuscated political and military activities to skirt the rules-based international system’s norms and rules by temporarily generating plausible deniability to provide the time required to maximize situational opportunities. Russia’s reliance on political and military proxies during the Crimea campaign and the 2014-2015 Donbas campaign’s spring offensive are exemplars of this idea.77 It therefore follows that indirect involvement, indirect control, and plausible deniability are not core tenets of proxy war strategy, but are shibboleths of a bygone era. The following sections provide a set of critical ideas, as a result of this paper’s theory testing, to supplant the shibboleths of indirectness and plausible deniability.

**Time and the Art of Proxy War**

Time is at the heart of proxy strategy. A proxy strategy works best when it provides its user the time it needs to a) gain a significant strategic foothold on its policy objectives, b) gain that foothold, not necessarily the entirety of its policy objectives, a step quicker than international recognition, condemnation, and counteraction, and c) do so while minimizing political and domestic risk. Put another way, strategic actors use proxy wars to manage external and internal risk. An aggressor manages external risk (i.e., international recognition, condemnation, and counteraction) by moving first and maintaining obfuscation to the degree and time needed to gain a foothold. An aggressor manages internal risk (i.e., lack of political and domestic support for aggressive foreign policy aims) by offloading the preponderance of close combat to surrogate forces.

Indirectness, on the other hand, does not account for the centrality of time to proxy war strategies. Further, indirectness describes a method, and mistakenly conflates the notion of not operating along the frontlines of combat and hidden involvement as the primary reasons for utilizing a proxy strategy. Russia’s war strategy, however, turns the idea of the importance of indirectness on its head.

Russia’s use of proxies in Crimea was a time-based decision. The use of ‘little green men’ – thinly disguised Russian soldiers –
and Wagner Group proxies focused on taking Crimea by force quicker than Kyiv, or the international community, could respond. Moscow’s use of proxies had little to do with hiding the Kremlin’s involvement and more to do with temporarily confusing Kyiv and the international community. Russia’s reliance on confusion was to generate a strategic situation that allowed its military forces—traditional and proxy—to operate at a quicker pace than Ukraine and the international community’s collective ability to recognize what was going on, who was doing it, what response options were available, and acting in response to the aggression.

Moreover, Russia’s initial incursion into the Donbas in 2014 demonstrates that indirect involvement is not a helpful idea for understanding proxy war. Russian proxy forces took control of nearly all of the Donetsk and Luhansk oblasts during April–May 2014. Through May and July 2014, Ukrainian forces had momentum and reclaimed about one-third of Kyiv’s stolen territory. Alarmed by the deteriorating situation regarding its proxy strategy in the Donbas, Moscow authorized the use of conventional Russian land forces to staunch the bleeding. In July 2022, Russian artillery forces delivered a devastating blow against Ukrainian forces at Zelenopillya. In August, regularly Russian forces were fully committed to combat at Luhansk Airport and Ilovaisk. Shortly thereafter, Russian land forces were committed, multiple times, to the long and hard-fought battle for Donetsk Airport. Lastly, Russian land forces helped win a decisive victory at Debaltseve in February 2015. In each of the situations, the involvement of Russian land forces was not indirect, nor obscured. In almost real-time, reports were coming out in open-source reporting that thousands of Russian soldiers and hundreds of tanks, infantry fighting vehicles, and artillery were pouring across the border from Rostov oblast and into eastern Ukraine. Russia’s reliance on proxy forces in the Donbas during 2014-2015 provided Putin with the strategic flexibility he needed to conduct a blistering campaign of territorial conquest and martial brutality, without causing much alarm in the Russian countryside or the halls of the Kremlin. That strategic flexibility provided Putin with the political currency he needed to double down in Ukraine in 2022.

**Strategic Flexibility and Attrition**

The Russo-Ukrainian War, from its inception in the spring of 2014 to today, clearly demonstrates that proxy wars are, and will continue to be, wars of attrition. Proxy force employment fuels wars of attrition for two primary reasons. First, an actor’s use of a proxy is often a gambit to offset the high degrees of political and domestic risk that accompany an aggressive foreign policy. A proxy strategy allows the actor to offload the preponderance of death and dying to a surrogate, thereby decreasing much of the political and domestic risk inherently found in war.

Second, when comparing proxies with a strategic actor’s own forces, the proxy is an expendable force, whereas the principal’s force is a more valuable commodity. From the battlefield, this tends to result in the principal actor fighting at extended range, often with long-range fires and armed drones, while feeding the proxy force into the pulverizing reality of close combat. From a practical standpoint, the tactical employment of proxies fuels high combat casualties, high civilian casualties, significant collateral damage on civilian infrastructure, high numbers of internally displaced people, and depopulation of focal points of combat. Russia’s casualty count—reportedly over 120,000 as of January 2023—and the evisceration of Ukrainian cities,
such as Pisky, Sievierodonetsk, Mariupol, and Bahkmut, make this point self-evident.\textsuperscript{81}

Third, the argument can be made that irregular force proxies care less about international humanitarian law (IHL) than uniformed soldiers, and therefore proxy strategies generate a more wanton and indiscriminate use of force. It logically follows then that civilian casualties, collateral damage, and violations of IHL are higher in proxy-laden wars.

Fourth, due to their ad hoc and temporal character, proxy forces are frequently less trained, and less proficient in the use of force, than regular soldiers. Professional militaries, especially those that do not rely on conscription, value training because it creates proficient forces. Professional militaries value proficiency because it increases their effectiveness and efficiency in warfighting. In turn, effective and efficient warfighting contributes to precise combat operations, less battlefield destruction, less collateral damage, and makes war less deadly, at least theoretically, for all its participants. In contrast, undertrained and untrained proxy force inefficiency fuels high casualties amongst the proxy force, high collateral damage, and high civilian casualties.\textsuperscript{82}

The 2014-2015 Donbas campaign’s blistering battles of Luhansk Airport, Ilovaïsk, Donetsk Airport, and Debaltseve pulverized eastern Ukraine, killed upwards of 14,000 Ukrainians, displaced thousands, and reduced frontline towns like Pisky to double-digit populations.\textsuperscript{83} The Donbas campaign bears witness to the brutality that accompanies outsourcing combat to surrogate forces. Further, the butchery of 2022’s battles of Bucha, Kharkiv, and Mariupol, in which both the Donbas proxies and the Wagner Group played significant roles, reinforces the findings from 2014-2015.

Lastly, opportunistic third parties can change the character of proxy wars. The United States, for instance, pragmatically enlisted Ukraine as a transactional proxy closely following Russia’s February 2022 reinvasion.\textsuperscript{84} Seeking to maximize its strategic flexibility within the conflict the United States opted for the technology diffusion model of proxy war to combat Russia.\textsuperscript{85} Keeping in line with the importance of minimizing political and domestic risk by using another actor’s force to do the fighting on one’s behalf, the United States provided Ukraine with $6.3 billion in security assistance in fiscal year 2022 and $13 billion in military assistance since August 2021.\textsuperscript{86} Collectively, those packages have included thirty-eight HIMARS (High Mobility Artillery Rocket Systems), forty-five T-72B tanks, 1,600 Stinger Anti-Air Missiles, 8,500 Javelin Anti-Armor Systems, 142 155-millimeter howitzers, thirty-six 106-millimeter howitzers, and a vast array of other warfighting materiel.

To conclude, proxy theory must maintain pace with the realities of armed conflict. The Russo-Ukrainian War illustrates that indirect engagement, indirect control, and plausible deniability are outmoded principles of proxy war theory. Proxy wars are useful insofar as they provide a strategic actor with temporal advantages and strategic flexibility. Moving forward, proxy war theorists must account for these factors above all else in their theories to maintain pace with the reality of proxy war.

Proxy wars might make armed conflict appear as a less dangerous option for a principal actor. However, proxy war theorists and practitioners alike must move beyond this cynical, and fundamentally
mistaken, notion. If the Russo-Ukrainian War demonstrates one thing above all else, it is that proxy war is a savage game of butchery, and routinely falls well within the bounds of wars of attrition.

About the Author: Amos Fox is a PhD candidate at the University of Reading. Further, Amos is an associate editor at the Wavell Room, and he is on the staff at the Irregular Warfare Initiative.
Ibid.


3 Franklin Holcomb, “The Kremlin’s Irregular Army: Ukrainian Separatist Order of Battle,” Institute for the Study of War (2017), 7-23

4 Fox, “The Donbas in Flames.”

5 Michael Hirsh, “Why Putin’s war is the West’s biggest test Since World War II,” Foreign Policy 24 (2022).


9 Ibid.


15 Galeotti, Putin’s Wars, 173.


17 Ibid.


19 Galeotti, Putin’s Wars, 180.


21 Fox, “The Donbas in Flames,” 4-6.


23 Galeotti, Putin’s Wars, 183


26 Galeotti, Putin’s Wars, 184

27 Trent Telenko, “Ammo, Railroads, Tyres, and Logistics are Driving Military Strategy in Ukraine,” on Peter Roberts (host), This Means War (podcast), July 28, 2022. Available at: https://podcasts.apple.com/gb/podcast/ammo-railroads-tyres-and-logistics-are-driving/id1629454648?i=1000571443399


Holcomb, “The Kremlin’s Irregular Army,” 8-12.


Ibid.


Ibid.

Ibid.

61 Galeotti, Putin’s Wars.
69 James Risen and Ken Klippenstein, “The CIA Thought Putin Would Quickly Conquer Ukraine. Why Did They Get It So Wrong?” The Intercept, 5 October 5, 2022. Available at: https://theintercept.com/2022/10/05/russia-ukraine-putin-cia/
78 Fox, “‘Cyborgs at Little Stalingrad,’” pp. 5-16.
82 Kono Carragee, “Historical Failures in the Use of Proxy Forces and Implications for Future Conflicts,” Johns Hopkins University, 2020, 34-38
84 Amos Fox, “Ukraine and Proxy War: Improving Ontological Shortcomings in Military Thinking.” 3-5.
85 Ibid.
When an Intelligence Agency Loses its Mind: An Analysis of the KGB’s Operation RYAN and What it Means for Russia’s War against Ukraine

Patrick Hutson

As an authoritarian leader, Vladimir Putin likely influences Russian intelligence to produce conclusions confirming his own views and prejudices—which likely influenced his decision to launch the brutal invasion of Ukraine in 2022. But this process of forcing information to fit conclusions creates the intelligence failure of confirmation bias. This failure is potentially disastrous. Disturbingly, there are parallels between the Russian war against Ukraine and a previous Soviet intelligence failure that followed this exact pattern: Operation RYAN. During Operation RYAN, the paranoid Yuri Andropov ordered Soviet intelligence to confirm his suspicions that the U.S. was planning a surprise nuclear preemptive strike on the Soviet Union. Forced to comply, the KGB ultimately produced bizarre intelligence confirming Andropov’s fears while having no connection to reality. The resulting feedback loop increased Soviet paranoia so much that its leaders nearly caused a nuclear war themselves during NATO’s Able Archer 83 exercise. Operation RYAN poses disturbing questions that can be applied today as Putin refuses to back down. Since Putin has a habit of eliminating everyone who disagrees with him, what if he's created a fantasy world where everything he believes is true? And if so, how much worse could things get?

Introduction

If someone wanted to determine what Vladimir Putin is thinking about at any given moment, one way would be to learn what his intelligence agencies are telling him. The whole point of an intelligence agency is to inform leaders about factual information necessary for their work. Rather, that’s the way an intelligence agency should work. Under an oppressive regime like Vladimir Putin’s, the truth is often the first thing to be repressed. There is no guarantee that intelligence agencies under an authoritarian regime actually produce factual content. The paranoia of a leader can easily cause intelligence agencies to become massive echo chambers, instead producing intelligence that confirms their world view while creating catastrophic intelligence failures. An organization that is supposed to provide factual updates provides fiction—effectively losing its mind.

This article makes the argument that, under a totalitarian system, a leader can inadvertently coerce an intelligence agency into reporting information that confirms their own beliefs, and, if left unchecked, such a negative feedback loop can lead to catastrophic consequences. I will support this argument by examining a historical intelligence failure under the Soviet Union, Operation RYAN, through the lens of Russia’s pre-Ukrainian invasion intelligence failure. This paper will research and analyze the origins of Operation RYAN—a Russian acronym of “Raketno Yadernoye Napadenie,” meaning “nuclear missile attack.” I will then discuss the operation’s immediate failure and its consequences during NATO’s 1983 Able Archer exercise. This discussion will be followed by an retrospective analysis of why Operation RYAN failed from an intelligence perspective. After that, I will discuss the possibility that such a failure is occurring again today in Putin’s Russia and how such a failure may have provided the motivation for his invasion of Ukraine. By understanding the intelligence failure created by Russian intelligence under one paranoid former KGB officer decades ago, it is possible to understand how a similar
failure could be happening today under similar circumstances.

**Literature Review and Methodology**

After Russia’s cruel and disastrous invasion of Ukraine, Western intelligence agencies, media, and academia now have some of the information necessary to determine why the invasion occurred. Many articles have explained that the war was influenced by factors such as Russian imperial ambition, President Vladimir Putin’s megalomania, or the prejudicial hatred that many Russians feel toward Ukrainians. But there is another factor that explains both the invasion and Russia’s subsequent poor performance. According to open source reporting in outlets such as *The Washington Post*, Russia’s FSB (Federal Security Service; the domestic security agency that also handles intelligence regarding former Soviet states) predicted that an invasion would be an easy victory. FSB officers claimed that Ukrainian leaders would flee the country and that invading Russians would be joyously welcomed by a gracious Ukrainian population. These FSB officers—most likely influenced by Russia’s totalitarian system—created an echo chamber that reinforced Vladimir Putin’s pre-existing view that a war against Ukraine would be easy for him to win. Like the FSB’s conclusions, Putin’s internal thought process is not easily determined. However, further open source reporting in Western outlets shows that Putin is isolated and distrustful of advice that conflicts with what he believes. The result was a war likely launched because Putin’s authoritarian leadership influenced his intelligence organizations to give him information that was biased in favor of what he wanted to hear.

To understand the relationship between Operation RYAN and the ongoing invasion, I analyzed the origins and failures of Andropov’s Operation RYAN to determine what similarities it might share with Putin’s intelligence in the lead up to the invasion of Ukraine. Putin’s intelligence failure can be better understood if it is first viewed through the lens of Operation RYAN. Both Andropov and Putin were former leaders of Soviet and Russian intelligence, both were heads of an authoritarian state that punished opposing thought, and both launched operations that appeared to be total disasters and threaten their own country’s security.

I utilized both primary and secondary sources that chronicled and assessed Operation RYAN. I used primary documents directly from the KGB and made available through George Washington University’s National Security Archive. I also utilized recent works such as Ben MacIntyre’s excellent *The Spy and the Traitor*—a biography of KGB officer Oleg Gordievsky who served through the height of Operation RYAN. Unfortunately, I faced a number of limitations during this research. I was limited to only English language accounts. Another major limitation is the overall availability of information about the inner workings of both Soviet and Russian intelligence. While Operation RYAN occurred long enough ago for some sources to be available, the closed and secretive structure of Russian politics makes it difficult to find information beyond personal accounts. Meanwhile, information about current intelligence practices in Putin’s Russia was even harder to find. Much of the information I was able to find comes from Western speculation of its workings and Putin’s mindset. Next, I will discuss the origins and failures of Operation RYAN.
A Suspicious Andropov Creates Operation RYAN

In May of 1981, senior Soviet intelligence officers attended a Moscow conference hosted by General Secretary Leonid Brezhnev and KGB Chairman Yuri Andropov. Andropov had an important announcement: Ronald Reagan “was actively preparing for nuclear war.” Why Andropov believed this is unclear, although Marc Ambinder’s account of Operation RYAN discusses Andropov’s unsteady mental state, highlighting his notorious paranoia and that Andropov often “saw conspiracies everywhere.” Immediately after the conference, Andropov created Operation RYAN, which quickly became the biggest peacetime Soviet intelligence operation ever launched. Andropov’s orders were the push. And like a row of dominos beginning to fall one after another, his orders had complicated and immense consequences. In November of that year, instructions were telegrammed to KGB stations—or rezidenturas—around the world. Officers were ordered to “uncover plans…and to organize continual watch…for indications of a decision being taken to use nuclear weapons against the USSR.” Soviet leadership provided a list of these indications for their officers. By 1983, the KGB “had compiled seven binders” full of observable nuclear war indicators that would be sent back to Moscow so Soviet leaders could prepare for the expected strike. Once Andropov became General Secretary after Brezhnev’s death, Operation RYAN increased in importance. In 1982 “all rezidenturas were instructed to make RYAN a top priority,” even if it meant “a de-emphasis on regular foreign intelligence work.”

As Operation RYAN escalated, its scale required the KGB to make institutional changes. Nate Jones extensively discusses organizational changes necessary to launch Operation RYAN in Able Archer 83: The Secret History of the NATO Exercise That Almost Triggered Nuclear War—a useful source that describes the logistics of the operation. This was the first cooperation between the KGB and GRU (Soviet military intelligence) in Soviet history. Neither intelligence agency had the resources necessary to implement an operation of this magnitude alone. The KGB itself was also restructured. The entire operation was overseen by a new division in the KGB’s First Chief Directorate, ironically named the “Institute for Intelligence Problems.” Meanwhile, “three hundred positions were created” to “report on and monitor” incoming intelligence while fifty new officers were hired for coordination. The KGB even developed a “primitive computer system” to process the immense quantities of intelligence coming in from rezidenturas. Not even the full power of the Soviet Union was enough. Observing signs of a sudden attack also required the Soviet Union’s allies. “The primary mission of Warsaw Pact intelligence agencies” was gradually changed to compiling intelligence that indicated an attack was imminent. The East German Stasi, as well as Bulgarian and Czechoslovakian intelligence agencies, became major providers of intelligence for the operation. The Warsaw Pact’s entire intelligence apparatus gradually transformed to support RYAN and prove that the United States was preparing for a preemptive nuclear strike. In doing so, Operation RYAN arguably became one of the most important events in the history of Soviet intelligence.

An Inherently Flawed Operation Becomes a Failure

Despite the colossal investment, Operation RYAN was a catastrophic intelligence failure. Its entire premise was wrong: the
United States had no intention of attacking the Soviet Union by surprise. While Reagan aggressively opposed communism, he also feared nuclear war. In 1981 he publicly stated that, in such a conflict, “all mankind would lose.” Reagan simply did not want a nuclear war—let alone a preemptive one. KGB officers regardless began collecting intelligence on their assigned indicators, confirming something that didn’t exist. At first, many of these indicators represented expected ways a foreign intelligence service would try to target the United States and determine military plans. Officers were ordered to surveil “key nuclear decision makers”, monitor “nuclear depots [and] military installations”, and recruit sources in “government, military, intelligence, and civil-defense organizations.” However, these indicators became increasingly bizarre and detached from reality, revealing that the Soviet leaders knew shockingly little about the United States. Churches and banks were listed as key nuclear decision-makers that would decide when to strike. Collectors were even ordered to report the price of blood in blood banks, because Soviet leadership believed they operated as literal capitalist businesses. These obviously wrong indicators gave no indication of the United States’ war plans, and KGB officers assigned abroad knew it. Despite their reservations, the officers complied and began producing the evidence they were required to find. Whether or not they shared Andropov’s beliefs about an incoming nuclear attack was irrelevant. By following skewed collection requirements, the KGB began producing skewed intelligence.

What followed next was what KGB officer and spy for MI6 Oleg Gordievsky described as “a vicious spiral of intelligence gathering and evaluation, with foreign stations feeling obliged to report alarming information even if they did not believe it.” As a result, Soviet leadership became convinced their imaginary impending nuclear attack was real and ordered the KGB to collect further proof. Operation RYAN transformed the KGB and Warsaw Pact intelligence apparatus into a giant feedback loop. Paranoid leaders ordered officers to prove their fears, officers obeyed, leaders became more paranoid, and officers received more orders to collect more proof. The already massive operation continued to expand through the eighties as the KGB devoted more resources towards discovering something that did not exist. Even mundane information, such as the construction of a highway, was presented as proof of a coming attack, convincing leadership they were right all along. It was as if the KGB changed their mission to prove the conspiracy theory their leader believed was true. One could argue that the point of an intelligence organization is to collect the intelligence that policymakers request. After all, a policymaker may know things intelligence leaders do not or they could want to point their intelligence organization towards something they feel is overlooked. However, even if a policymaker requests to collect intelligence about something nonexistent, that organization must be free to respond and deliver facts the policymaker may not want to hear. If intelligence is conducted in a totalitarian climate, intelligence organizations will fear their policymakers and deliver the information they want to hear. By forcing the KGB to prove the imaginary attack was real, the Soviet Union led its entire intelligence apparatus into an intelligence failure. Thanks to Andropov’s paranoid orders, the KGB lost its mind.

**Why Operation RYAN Failed**

The reason Operation RYAN failed is obvious now: its premise was incorrect from the start. The entire operation followed the
“garbage in, garbage out” principle. If ordered to collect intelligence that proves a nonsensical idea, an organization will provide nonsensical intelligence to ‘prove’ that idea. But more importantly, why were these orders made and why were they implemented? Soviet leader Yuri Andropov shares a significant portion of the blame. His belief in a surprise nuclear attack “probably had more to do with [his] personal experience than rational geopolitical analysis.”

Andropov served as the Soviet Ambassador to Hungary during the 1956 Hungarian Revolution, where Soviet power over the satellite state was threatened by mass protests. His participation in brutal suppression as Chairman of the KGB likely also led to his assumption that all leaders—such as Reagan—would pursue the same “extreme measures” against their opponents that he would, such as launching preemptive wars. All of Andropov’s actions were influenced by his personal experiences and beliefs. If new information did not conform to his view of reality, it was not true. As a result, he ordered the KGB to be an extension of his own thought process. They were to find him intelligence that confirmed his beliefs, institutionalizing his confirmation bias.

The totalitarian culture of the Soviet Union and the KGB are also to blame for RYAN’s failures. The orders to provide proof of a nuclear attack came directly from Andropov, the most powerful person in the Soviet Union. KGB officials had no choice but to comply. The Union’s authoritarian society created a culture of “terrified conformity,” eliminating any ability to stand up to leaders. The authoritarian Soviet regime enforced groupthink, creating a climate where “obedience was more powerful than common sense” in the intelligence workforce. If any KGB officers questioned why they were required to collect ridiculous intelligence about blood prices, highway construction, or nuclear armed clergy it would have meant the end of their career, if not their life. There was nothing KGB officers could do to stop this political interference; it was a standard part of their work. The KGB was not a neutral, information-providing intelligence agency: its primary role was to be the “servant of the [communist] party.” Politicizing intelligence was the very goal of the KGB. By design, all of Soviet intelligence was inherently influenced and distorted by the Communist Party’s totalitarian politics.

A Vicious Cycle Nearly Leads to Catastrophe

As Operation RYAN expanded further, its vicious cycle of false intelligence began to have dangerous repercussions. Because the operation gathered more intelligence that confirmed their fears, Soviet leaders became further convinced that war was imminent. The détente period of calmer diplomatic relations between the Soviet Union and the United States ended as the Soviet leaders intensified their rhetoric. In 1982, the Soviet Union proclaimed that it would end any “attempt to achieve mutual cooperation with the United States.” Peace was now off the table. In 1983, Andropov—now General Secretary—warned that the United States was “moving toward the dangerous ‘red line’” of nuclear war. This sudden escalation was met with confusion in the United States. Reagan administration officials struggled to determine whether these statements represented genuine fear or some sort of calculated ploy. Concerned by Soviet overreaction and fearing conflict, the United States began preparing NATO military exercises in response to Soviet rhetoric. Operation RYAN was inadvertently escalating the Cold War, pushing it towards real conflict.
In 1983, NATO began Able Archer—its annual simulated military exercise. The exercise was based on a theoretical invasion of Yugoslavia, Finland, and NATO allies such as Germany, Norway, and Greece by an obvious stand in for the Soviet Union called “Orange.” The practice invasion was designed to quickly escalate to simulated chemical and nuclear warfare. Its scale was immense. Able Archer 83 involved 40,000 troops from the United States, 19,000 of whom deployed from the United States to Europe. B-52 bombers deployed in an Able Archer exercise for the first time and NATO leaders including President Reagan, British Prime Minister Margaret Thatcher, and German Chancellor Helmut Kohl planned to participate. Soviet intelligence immediately noticed these preparations. Observation of NATO planes appeared to show they were armed with nuclear weapons and signal traffic indicated that the United States moved to DEFCON 1—an unprecedented event. While the United States and NATO prepared to fight an imaginary nuclear war, their exercise looked similar to a real one.

Able Archer 83 appeared to be exactly what the Soviets were waiting for. Believing the long-feared attack was finally upon them, the Soviets panicked—nearly turning simulated nuclear war into an actual one. The Soviet Union immediately prepared a response “unparalleled in scale” to previous war mobilizations. All three elements of the Soviet nuclear triad activated: planes were equipped with nuclear weapons, submarines were deployed to defensive positions, and “around seventy SS-20 missiles targeted on Western Europe were placed on heightened alert.” Disturbingly, there is even speculation that the Soviet Union’s ICBM silos were activated and prepared to launch. Operation RYAN’s effects cascaded to create what historians now believe was “the moment of maximum danger” in the late Cold War, one which could have very easily escalated to nuclear conflict. Operation RYAN nearly caused the nuclear war it was created to predict. Thankfully, the Soviet Union decided not to pursue a preemptive strike. Able Archer 83’s quiet conclusion made it obvious that the exercise was not the surprise nuclear attack Operation RYAN was looking for. In February of 1984, three months after Able Archer 83’s conclusion, Yuri Andropov died in power. Both his death and the nuclear war scare marked a turning point for Operation RYAN. The operation’s mission and goals changed in the following years. Instead of gathering evidence to prove a nuclear war was coming, officers collected intelligence on general Soviet security concerns such as NATO missile specifications and the United States’ activities in the Middle East and Central America. Operation RYAN declined in importance each year until it ended in 1989. Russian intelligence seemed to have learned from its mistakes and Operation RYAN’s influence was dramatically reduced.

**Why Operation RYAN Matters Today**

History never repeats itself exactly the same way, but the parallels between Operation RYAN and Russian intelligence failures in Ukraine are too prevalent to ignore. Today, another former KGB officer is the head of state in Russia. And similarly to Andropov, Putin previously served as Director of FSB—one of the KGB’s successor organizations. Both Andropov and Putin are clear examples of paranoid, conspiracy-obsessed thinkers. During his speech to KGB leadership, Andropov claimed that every war was actually caused by imperialists who wanted to stop “the forward march of history.” This claim
effectively blamed the West for any military action the Soviet Union would take and framed the Soviet Union as simply defending itself from Western aggression. Putin has similarly concocted a conspiratorial narrative of his own to justify his attack on Ukraine. Based on his previous statements, Putin’s worldview is based on the idea that Russia is threatened by a bizarre collection of forces including NATO, neo-Nazis, multiple gender identities, and “cancel” culture all coming from the West. Just like Andropov before him, Putin likely believes he is threatened by secretive, sinister forces.

A leader falling to delusional, paranoid thinking has ramifications for the rest of the administration. Andropov believed he was facing an existential threat, removing anyone who disagreed—something Putin is likely doing again even more ruthlessly. During his time as KGB Chairman—and later as General Secretary—Andropov was notorious for his suppression of dissidents. The KGB “became a more sophisticated and effective means of enforcing near absolute political control” while leaders of independent groups such as an anti-nuclear citizen’s council were sentenced to imprisonment in psychiatric hospitals. Putin is likely waging a similar campaign of suppression and taking it even further. A suspiciously high number of oligarchs have recently died under mysterious circumstances, such as falling out of windows. While it is not definitively known, a common factor is that many of the deceased were critics of Putin’s invasion of Ukraine—leading many to suspect they were murdered under Putin’s orders. Andropov was content to imprison those who questioned him: Putin wants them dead. And by silencing his critics, Putin is almost certainly creating the exact same dysfunctional climate of fear that led to Operation RYAN’s production of fake intelligence.

Russian intelligence is likely falling into another intelligence failure because little in the institution has changed since the Cold War. There have been no meaningful reforms to Russian intelligence since the collapse of the Soviet Union. According to Julie Anderson’s “The Checklist Takeover of the Russian State,” Russian intelligence has spent the years since the collapse of the Soviet Union taking over all aspects of the Russian state. In 2000 Putin described the beginning of his presidency as a “a successful [intelligence] penetration operation,” and immediately began empowering his former colleagues in Russian intelligence. In her article, Julie Anderson extensively documents how Russian intelligence gradually seized all elements of power within the Russian state including the judiciary, the economy, the media, and even local governments. The same authoritarian culture that led to the KGB’s failure during Operation RYAN never went away: the FSB simply picked up where the KGB left off. In fact, Russian intelligence has only become more powerful since the collapse of the Soviet Union. Anderson describes Russia as an “FSB state,” where Russian intelligence and the government are one and the same. Nothing has changed in Russia since Operation RYAN, and, if anything, it has only gotten worse. Russian intelligence is more powerful than it was in the 1980s and it has sustained the exact same environment that led to the failure of Operation RYAN in the first place.

It does not strain the imagination to consider that Putin has inadvertently recreated his own version of Operation RYAN, giving himself a constant feed of biased or skewed information from officers too frightened to tell the truth. Today, Putin has publicly
declared that he does not believe Ukraine is a real country. Before his invasion, he could have ordered his officers to find intelligence proving Ukraine has a weak national identity or that a Ukrainian response to a Russian invasion would be a positive one. Given his preference for conspiracy theories, he might have also ordered them to prove that Ukrainian leaders were secretly a part of some imagined western, Nazi, LGBT cabal. Soviet intelligence officers complied with Andropov’s bizarre request for intelligence about fluctuating blood prices and nuclear clergy, and this time would be no different. Whether or not Putin directly ordered Russian intelligence to prove his beliefs, or if Russia’s and the FSB’s culture of compliance indirectly led to it, Russian intelligence likely came to a conclusion that Putin was already sympathetic to. Disagreeing with Putin is a death sentence. And if the penalty for disagreement with Putin is death, no one is going to disagree. All it took were orders from Andropov and the KGB lost its mind, becoming an intelligence agency that did not produce intelligence. Nothing has changed in Putin’s Russia. Russian intelligence exists to make him happy—and if that means delivering nothing but falsehoods about Ukraine, it’s what it will do.

Both Operation RYAN and Putin’s likely delusions about Ukraine are part of a much larger trend of conspiracy-minded thinking in Russian politics and security. This paranoia has almost certainly infected Russian institutions and is proclaimed by leaders at various levels of government. During protests in Russia in 2017, head of the National Guard Viktor Zolotov claimed that protesters were caused by “brainwashing through Western media.” Meanwhile, former head of the analytical department of the SVR Leonid Reshetnikov—someone supposedly responsible for factual and well-reasoned analysis—claimed that a secret, Illuminati-like “world government” is actually responsible for orchestrating world affairs. Instead of producing well-researched critical analysis, Russian national security scholarly journals frequently reference familiar conspiracy theories and tropes about George Soros and the “new world order.” These publications and statements go beyond propaganda and misinformation. This evidence indicates that Russia’s national security apparatus is incapable of providing actual analysis and can only see the world through the lens of simplistic conspiracy theories. This thinking has dangerous and obvious repercussions. Geopolitical crises are viewed not as “spontaneous events but…plans for [Russian] ‘regime change’ (smena vlasti) developed in Washington.”

No one has learned anything since Operation RYAN. Thanks to generations of paranoid leaders and a perverse culture of fear, Russian intelligence as an organization is no longer willing or able to distinguish truth from fiction. Perhaps the average Russian intelligence officer knows this is paranoid and that there are no merits to these claims, but at a certain point the beliefs of an individual no longer matter. If they are being repeated by leaders and mandated through official channels, eventually these beliefs become policy. The introduction of these paranoid theories into the intelligence process contaminates the political outcome. Operation RYAN showed this mindset can cause a nuclear war, and the invasion of Ukraine shows it can lead to horrific war crimes. How much more damage will this thinking cause?

Operation RYAN’s consequences are largely theoretical, as the nuclear war never came. But Putin’s intelligence failures have left devastation across Ukraine and could become even worse. Because Putin thought that a war would be easy to win, thousands...
of civilians have been killed and millions have been displaced in Ukraine. But Ukraine still resists Russia, and the war is dragging on. If Russian intelligence solely exists to confirm Putin’s beliefs, does he even know that he’s losing? Putin’s intelligence has lied to him before, and he’s created an environment that encourages them to continue delivering the lies he wants to hear, so what lie will they tell him next? Imagine for a moment that Putin demands confirmation that NATO will soon launch a preemptive nuclear attack on Russia, just like Andropov did. Russian intelligence is not likely to tell Putin that such a belief is deluded and has no chance of occurring. Putin’s underlings might even believe him. In the right authoritarian political climate, an intelligence service and its leader can lose their minds. And when a nuclear-armed authoritarian leader and his intelligence service lose their minds, the lives of millions are threatened. Operation RYAN ended as an intelligence failure. Today, Putin’s Russia shows us that things can also end much worse.

**About the Author:** Patrick Hutson is a graduate student in Georgetown University’s Security Studies Program (SSP) with a concentration in Intelligence. He received a bachelor's degree in Government and Politics with minors in History and Global Terrorism Studies from the University of Maryland, College Park.


3 Miller, Greg, and Catherine Belton.


9 Ambinder, 60.

10 Macintyre, 143.


12 Macintyre, 143.


14 Macintyre, 143.

15 Ambinder, 61.

16 Ambinder, 60.

17 Jones, 13.

18 Schaefer, Bernd, et. al. “Forecasting Nuclear War.”

19 Jones, 13

20 Jones, 22.


22 Macintyre, 145.

23 Ambinder, 120.

24 Macintyre, 146.

25 Ambinder, 120.

26 Macintyre, 146.

27 Ambinder, 120.

28 Macintyre, 143

29 Britannica, “Yury Andropov.”

30 Macintyre, 143.

31 Macintyre, 138.
32 Macintyre, 146
34 Jones, 16.
35 Jones, 23.
36 Jones, 23.
38 Macintyre, 178
39 Ambinder, 62
Daalder, Ivo. “Column”
41 Jones, 37.
42 Macintyre, 179.
43 Jones, 37.
46 Macintyre, 179.
47 Jones, 53.
52 Godfrey, Elaine.
58 Kragh, Martin, et. al. “Conspiracy Theories in Russian Security Thinking.”
59 Kragh, Martin, et. al. “Conspiracy Theories in Russian Security Thinking.”
60 Kragh, Martin, et. al. “Conspiracy Theories in Russian Security Thinking.”
Turning Grievance and Vulnerability Into Violence: Lessons on Identity-Based Conflict from the United States and Overseas

Jordyn Iger

High-profile global incidents of identity-based violence in recent years, such as the 2022 Buffalo, New York supermarket shooting and the Tigray conflict in Ethiopia, have demonstrated that hate can be a powerful mobilizing tool in any community perceiving its own marginalization for any reason. However, research on the linkages between emotional vulnerability, prejudice, and violence is still relatively new; further, sharing of information and best practices between practitioners working in the United States and abroad on similar issues is rare. Through literature reviews and interviews with practitioners working in the United States and Myanmar, I examine how identity-based grievance can serve to meet human needs for belonging and self-definition, and I outline histories of identity-based grievance in the United States and Myanmar. I analyze salient differences in efforts to combat identity-based grievance and violence between domestic American efforts and international aid programs; and provide policy and practice recommendations based on practitioner experiences. This research demonstrates that while the United States and Myanmar have different histories regarding identity-based grievance and violence, their experiences share a few notable patterns. Further, the main differences between interventions in the United States and elsewhere tackling prejudice and identity-based violence are largely related to the weight of the First Amendment of the U.S. Constitution, which treats hateful speech as a right. Finally, in the experience of U.S. and global practitioners, the most effective approaches involve reducing the salience of the cleavage at issue, addressing conditions (e.g. material) underpinning identity-based grievance, and leveraging current political conditions and the incentives of key stakeholders in conflict.

Introduction

In the wake of World War II, after Holocaust survivors were liberated from concentration camps, academic attention to questions of prejudice and hatred began in earnest to try to make sense of German state-sponsored persecution. Interest in identity as a mobilizing tool for violence intensified further after the breakup of the former Yugoslavia, the Rwandan genocide, and the failure of the Oslo Accords in Israel and Palestine. These events resulted in a growing body of research on how prejudice and hate manifest in humans, as well as scores of interventions to bridge divides and improve social cohesion aimed at reducing the likelihood of new or reemerging conflict. More recently, the events in Washington, D.C. on January 6, 2021, made clear that the United States is not immune to threats posed by the spread of hateful rhetoric and beliefs which have been present for centuries.

Broadly, two schools of thought dominate discussions around identity-based conflict. One suggests that the distinction between “self” and “other” drives ethnic violence at the group level.1 Another holds that identity-based cleavages are political tools superimposed onto local conflicts that operate based on individual grievances that are not directly related to identity.2 These are not mutually exclusive, as hate can be a powerful mobilizing tool in communities perceiving their own marginalization for any reason. A literature review and interviews with researchers and practitioners about identity-based conflicts in various contexts reveal that these two schools of thought work in tandem to facilitate identity-based violence.
Instances of identity-based violence show that states and individuals participate in a process to translate grievance to hate to violence. Economic and social growth in most countries is uneven. This uneven growth often occurs along identity cleavages, particularly if the distribution of public services and other state resources is unequal. As a result, citizens experience unequal economic and social outcomes and lose trust in the state. The state feels threatened and extremist groups feel empowered by this public loss of trust; both groups then introduce and impose hateful “othering” narratives to explain the unequal outcomes. Citizens who feel emotionally vulnerable due to these unequal outcomes and loss of trust subscribe to those narratives and may then respond positively to calls for violence against outgroup members.

This process of interpreting grievances to lead to violence involves many steps. Evidence and experience from both domestic and overseas programs show that all steps of the process must be addressed at once to effectively stem this violence. Yet significant differences in language, framing, and approach are evident in studies of and interventions to address identity-based cleavages in the United States and elsewhere. As a result, guiding philosophies driving intervention design and impact analysis can be very different. This divergence can be leveraged as an asset, as varied perspectives can better inform program design and implementation. This paper aims to bring together lessons and best practices from practitioners working on domestic and international issues, who often lack incentives to share information in part due to organization structures and in part because of “American exceptionalism,” a belief that the United States is immune to certain problems that occur abroad.³

This paper will first examine literature in psychology, sociology, and peacebuilding to analyze the roots of prejudice and hate, then apply that framework to two case studies that are rarely directly compared: white supremacy in the United States and Burmese Buddhist supremacy in Myanmar. Finally, the paper will review trends in interventions addressing identity-based hate and violence in the United States and internationally and provide policy recommendations based on the experiences of practitioners and researchers across contexts.

Psychosocial and Relational Roots of Prejudice and Hate

Understanding a group’s path from prejudice and hate to violence must begin with an examination of the role of “ingroup” formation in addressing human needs, the role of “outgroup” formation in generating a sense that these needs are threatened, and how that feeling of vulnerability generates prejudice strong enough to respond positively to in-group encouragement to commit acts of violence.

Definitions of human psychological needs exhibit clear patterns. Although Maslow does not directly explain how meeting these needs influences conflict, the four most immediate needs on the hierarchy include, in order, basic survival, safety, belonging, and personal fulfillment.⁴ A few decades later, Burton identified four main human needs: a sense of security and identity; a consistent response from the environment in order to learn; recognition and valued relationships (i.e. bonding); and some control over their environments to ensure their needs are fulfilled.⁵ In the early 2000s, Deci and Ryan outlined self-determination theory, whose three psychological needs (autonomy, competence, and relatedness) underpin internalized motivation.⁶
These needs are partly met through the formation of social groups. Humans do not associate with social groups for merely pragmatic benefits of improved chances of survival; they also derive an emotional sense of “belonging” from such groups (i.e., a sense of security and identity as well as recognition and valued relationships, from Burton’s list of needs). This comprises social identity, which, as defined by Tajfel and Turner in the 1970s, are “those aspects of an individual’s self-image [that] derive from the social categories to which he perceives himself as belonging.”

Prejudice, stereotype formation, and discrimination “meet” human needs when a major component of these needs is superiority to other groups. The feeling that one’s own group is superior to others requires little more than “in-group” and “out-group” designation. This was demonstrated by a 1954 study known as the “Robber’s Cave” experiment. In this study, boys as similar as reasonably possible in family background and age were randomly assigned to two groups. They were situated at a campsite, isolated within their respective groups, in order to build a sense of identity and cultural cohesion in their “in-group.” The groups were then brought in contact with each other in a series of competitions. The separation of the boys into competing teams was enough to generate hostility so strong it obscured the boys’ perceptions of reality. There were no intrinsic differences between the two groups, yet they easily created in-group identities and out-group hostilities.

As evidenced by the Robber’s Cave experiment, perceptions of ingroups and outgroups are influenced by emotions that obscure rational thinking and induce vulnerability to calls to violence. When intergroup competition tests group perceptions, people react emotionally. The basis of intergroup emotions theory is that “people can and do feel emotions on behalf of the[ir] group;” a stronger sense of group identification correlates with stronger emotions in response to group threats. In a 2019 case study, researchers outlined four emotions involved in the development of prejudice that leads to violence: resentment, fear, rage, and hatred. While this study asserts that resentment is the strongest of these, intergroup emotions theory qualifies that when a person perceives their ingroup as stronger than the outgroup, intergroup competition sparks anger. In the opposite situation, intergroup competition generates anxiety. These emotional reactions inhibit the ability to rationally process new information that might contradict the source of the anger or anxiety.

Strong negative emotions that underlie violence are not easily disarmed through counterfactual information. Earlier research found that “anxiety was found to lead to more thoughtful processing of information, while anger led to more reliance on easily available cues such as social identities.” However, anxious citizens look for information, particularly threatening information, in a biased way. Researchers Grant-Halvorson and Rock (2015) identify a series of cognitive biases that happen outside of conscious awareness yet influence how people process new information. The biases most relevant to information about in-groups and out-groups can be classified as similarity (i.e., favoring ingroups), safety (i.e., avoiding loss is more emotionally powerful than securing a gain), and expediency (i.e., making decisions using the minimum amount of cognitive effort required). In stressful situations, the effects of bias on decision-making are amplified, and the ability to think critically about new information suffers.
Taking actions that lead to violence (for example, starting and spreading rumors about the out-group), or even merely consuming and accepting without questioning such actions, activates intergroup and individual emotions in various ways. The resulting motivations for either taking or supporting inflammatory actions range from: (i) heightening personal status among others by conveying the impression of uniquely having access to important information; (ii) desiring to inform others of impending danger; (iii) sharing anxieties about potential threats and seeking reassurance; and (iv) expressing hostility, fear, or wish, either directly or indirectly through projection. These motivations align with the three types of internalized motivations as defined by social determination theory: (i) identified regulation, or the conscious valuing of a behavior even if it is not enjoyable (for example, to heighten personal status among a group valued as important); (ii) integrated regulation, or valuing a behavior because it is important, an expression of the self, and coherent with other personal goals and values; and (iii) intrinsic motivation, or the mere pleasure (i.e., emotional reward) of a certain behavior. The two lists combined convey how motivation can be externally motivated to elevate social status or internally motivated by the alignment of negative emotions and values. Critically, when the opportunity arises to express these negative emotions as insulting or derogatory sentiments towards the outgroup, self-esteem that has been damaged due to (perceptions of) ingroup loss becomes restored.

Understanding the role of emotion in intergroup tension helps to explain why state support can be a powerful catalyst that aids majorities in resorting to violence. Political leaders can take advantage of group emotions to gain support and power. In 2019, Ethan Busby et al. found that subjects asked to frame failures of governance in terms of “dispositional blame,” or the attribution of behavior or circumstances to someone’s inherent characteristics, are much more likely to express populist attitudes and to support populist candidates. Social, political, and religious leaders exploit this for their own gains, employing violent rhetoric that evokes these negative emotions to great effect even among people who are not typically aggressive.

While not directly related to identity-based prejudice and violence, Emilie Hafner-Burton (2013) hints at the possible interests leaders might have in stoking such violence by explaining four “benefits” to perpetrators of human rights abuses: a sense of superiority or satisfying a sense of idealism, intelligence, revenge, and money. These align loosely with the human needs outlined earlier: a desire for superiority or idealism corresponds with a “sense of security and identity” as people derive their sense of identity from perceived moral superiority to “out-groups”; intelligence corresponds to a need for “consistent response from the environment” as people develop their understanding of society from lived experience; and revenge corresponds to a need for “control over the environment to ensure needs are fulfilled” as the desire for retaliation often stems from a perceived need to assert control and dominance over those who have committed harms.

Thus, states, as entities led by humans, can be motivated by something resembling emotions just as individuals can. Further, states and other authorities are uniquely positioned to not only promulgate hateful narratives, but also transform these narratives into social norms to which people feel pressure or enthusiasm to conform.
Linkages Between Prejudice and Political Violence

The cases of recent domestic violence in the United States and Myanmar demonstrate the process by which individuals translate grievance into violence. The majority group’s feelings of vulnerability and threat caused by anti-minority group narratives are not merely harmful at an interpersonal level but can have profound societal, political, and even institutional implications that persist through generations.

United States

The history of white supremacy in the United States, from the first slave ships to the January 6, 2021 attack on the U.S. Capitol, highlights that the entire definition of “whiteness” was built on superiority over “blackness” for economic reasons. Scholars of white supremacy in the United States generally indicate its main point of origin as European colonialism’s effort to divide economic and social classes along racial lines between colonial subjects (i.e., Africans) and Europeans. Europeans were influenced by medieval North African scholars who associated blackness with “filth, evilness, ugliness, and sin,” and passed these stereotypes and prejudices to Spanish, Portuguese, and other European colonialists. At the same time, European cultural consensus grew around the idea that white Christians were not enslaveable (anymore), and the sight of Black slaves in Spain and Portugal became a role model for the British to follow. Further, 17th century white settlers in North America attempting to enslave Native Americans were not successful because contact with foreign diseases meant that the latter simply could not “survive in sufficient numbers to become a reliable source of labor.” Therefore, although it was theoretically economically costly to transport people from one continent to another for enslavement, Black Africans became the victims of slavery in the early 1600s. This victimization was reinforced by the state where the first directly racially discriminatory law was passed in 1623, in British Bermuda.

In the United States, the growing need to justify the existence of slavery, even as the transatlantic slave trade was outlawed in Britain in 1807 and existing slaves in the British Caribbean were granted their freedom in 1833, further entrenched racist rhetoric and views against Black people. This need became particularly salient during the 1850s. During that time, as new states were added, debates grew over the balance of “slave” and “free” states in the union, the primary line of political division at the time. Proponents of slavery cited numerous theories and pseudo-science as justification. One popular theory, called monogenesis, drew from Christianity: namely, that God had turned Cain black as punishment for killing his brother Abel, and all Black people were thus descended from Cain. The opposite theory, polygenesis, argued that blacks and whites belonged to different species. Polygenesis gained more traction in and after the 1850s, influenced not only by “scientific” research but also even by race scientists’ emotions of “pity” upon encountering Black servants.

The guiding assumption underpinning these theories and the racist rhetoric that invoked them was that Black people were predestined to perpetual servitude to white people. Meanwhile, race “scientists” and racist doctors diagnosed “diseases” with names such as “Drapetomania” (given to a Black person who sought to escape from servitude) and “Rascality” (“misbehavior” of enslaved Black people). These “diagnoses” further reinforced both that servitude was a “natural” state and that
deviation from that state was to be considered deviance. Southern news publications drew from that conclusion to justify enslaving individuals, and declared that American founding principles, such as those enshrined in the Declaration of Independence,

...involve the assumption that the negro is the white man, only a little different in external appearance and education...Ethnology and anatomy, history and daily observation, all contradict the idea...34

The Civil War, and Reconstruction after it, facilitated the first identity-related loss of trust in state institutions by upending the previous social order. Around the end of the Civil War, global economic trends demanded increases in agricultural production of the primary crop cultivated by enslaved people: cotton.35 White Southerners thus reeled from not only the humiliation of clear defeat, but also the loss of a significant labor force at precisely the wrong time.36 One of the first acts of Reconstruction was the advancement of U.S. troops into Southern territory, seizing plantation land in the process.37 This was viewed as Northern encroachment and was counterproductive to building trust in the Union and curbing North-South and racial tensions. With this in mind, President Andrew Johnson reversed General William T. Sherman’s plans to redistribute plantation lands to free Black Americans because he was wary of upsetting both economic and race relations in such a direct way.38 Soon after, sharecropping, a system that involved labor as a means of repaying increasingly burdensome debts, arose to replace slavery as a means of maintaining plantation owners’ livelihoods with forced labor.39 However, as sharecropping contributed to growing economic inequality between landowners and the landless, indentured servitude ensnared poor whites as well.40 These white Americans, threatened by the idea of sharing a socioeconomic status with Black sharecroppers, emphasized racial differences to distinguish themselves from Black sharecroppers. Thus, “the symbolic good of whiteness offered these laborers a ground for full membership within the republic, no matter their diminished material status.”41

This sense of white vulnerability emboldened anti-Black extremist groups. As the last American forces withdrew from the South after Reconstruction, pro-South authors, starting with Edward Pollard, widely disseminated the idea of the Lost Cause.42 The Lost Cause held that the Civil War was not motivated primarily by slavery concerns but rather by concerns for the Southern “way of life,” and further that slavery was a benevolent force, rather than a societal evil.43 White supremacist groups such as the KKK, which was founded in 1865, terrorized Black freedmen in hopes of squashing their demands for equal rights and opportunities.44 White Southern backlash against freedom for former slaves was supported by lukewarm attitudes in the North toward granting Black Americans full, equal rights. These attitudes persisted even as Northerners, whose economy did not rely on slave labor, tended to support the end of slavery.45

Influenced by the racist science of a decade earlier, both Southerners and Northerners agitated over the “Negro Problem,” or what free Black Americans would do if they were no longer bound by slavery. At best, the “Negro Problem” held, they might be idle and struggle to navigate the world. At worst, they posed a societal threat, particularly to white women.46 Newspapers continued to write about the Black American as “a savage
or a slave” and references to Black Americans’ supposed inferiority remained as explicit as before the war.47

As Charles Darwin’s theory of evolution gained currency, so did “Social Darwinism,” or the idea that the wealthy and powerful were biologically superior, and Black individuals belonged at the bottom of the social order because of the inferior traits they developed from evolution.48 Even writers who viewed slavery as a problem for poor white Americans described Black Americans as something other than human.49 Critically, these groups had the support of Southern state governments, representatives in Congress, and the Supreme Court, which, respectively, enacted and then upheld discriminatory “Jim Crow” legislation that overtly segregated blacks and whites in all public spaces and sought to covertly strip Black Americans of various rights and protections, such as the right to vote.50 This legislation shows that the state can enforce individuals’ desire to strip other citizens of their rights, which is a result of their misinformed vulnerability-driven biases.

Mass media, which often draws on existing attitudes, further spread and normalized narratives of violence as an acceptable response to perceived racial threat. In 1905, author Thomas Dixon published his second novel *The Clansman: A Historical Romance of the Ku Klux Klan*, at a time when the Klan itself had lost most of its momentum.51 The novel “characterizes Reconstruction as an unmitigated disaster” and draws upon white supremacist themes such as tropes of the innocent white woman victimized by a brutish, savage Black man.52 The popularity of Dixon’s novel inspired D.W. Griffith’s infamous film *The Birth of a Nation* a decade later.53 The film is credited with igniting the rebirth of the Ku Klux Klan and facilitating the height of its power.54 That the movie simultaneously pioneered modern film techniques and story construction, thereby remaining in the minds of later directors who drew from it, likely expanded the reach of the film’s supremacist rhetoric and themes.55

Over the last century, white Americans have felt threatened by increased political enfranchisement of minority groups and have responded accordingly. In the 1960s, increased rights for Black Americans were successfully institutionalized as a result of the Civil Rights Movement and subsequent Supreme Court rulings that integrated public spaces and codified legal and civil protections.56 Not coincidentally, the Ku Klux Klan experienced a revival, gaining renewed notoriety for bombings, lynchings, and other acts of violence against not only Black Americans, but also other Americans considered to be non-white.57 As well, at least 101 Southern Congressmembers signed a “Southern Manifesto” encouraging opposition to the integration of schools.58 Throughout the 1960s and beyond, perceived progress with racial equality never fully succeeded at diluting white entitlement. More recently, the 2008 election of President Barack Obama led to renewed white anger as white Americans perceived a loss of political dominance. The Obama years witnessed the creation of the Deferred Action for Childhood Arrivals program and the rise of the Black Lives Matter movement, which aimed to protect Hispanic Americans and advocate for Black Americans, respectively.59 These developments have led television personalities such as Laura Ingraham to claim that America is undergoing “changes that none of us ever voted for and most of us don’t like,” namely, that America “has become estranged or alien” to white Americans.60
This perceived loss among some White Americans found a mouthpiece in President Donald Trump, whose rise to political prominence was accompanied by a rise in violent hate crimes against nonwhite Americans and by executive orders sanctioning outright discrimination. President Trump successfully drew upon white Americans’ sentiments that a white Christian background, rather than mere subscription to American civic and patriotic values, makes an American. Further, as evidenced by the discourse around remedying historical injustices towards other groups, for President Trump and his followers, “the historical primacy of whiteness (reflected by a historically majority status) is in a state of terminal decline.”

Famously, the Charlottesville rally in 2017 invoked the rhetoric of loss with the chant “You will not replace us.” Another version of this chant is “Jews will not replace us,” which was also periodically used in Charlottesville and explicitly connects the rally to the white supremacist and anti-Semitic views underpinning the Holocaust. Two years later, Federal Bureau of Investigation (FBI) Director Christopher Wray told the Senate Judiciary Committee that the majority of domestic terrorism investigations conducted by the FBI were connected to white extremist ideologies. This remark eerily foreshadows the white extremist motivations behind the Capitol attack on January 6, 2021.

Although Trump is no longer in office, the identity-based grievances his administration amplified linger. Benjamin Gellman reported that the most significant determinant of participation in the January 6 attack is white identity in a region where the percentage of white residents is in decline. Gellman explicitly notes that socioeconomic grievance alone was not enough to attract people to violence against the state, as the January 6 participants included highly educated professionals. As demographic trends erode white numerical majorities, these white Americans believe that with their majority, their economic and political power will dwindle as well. The perceived shield whiteness provides from other types of deprivation (such as poverty) becomes an entitlement that white citizens believe they are “owed,” and will resort to violence to “reclaim” if it is lost.

**Myanmar**

The history of in-groups and out-groups in Myanmar points to the ways that even groups long habituated to peaceful coexistence can, after experiencing economic and demographic shifts, turn towards emotional responses that facilitate hate and violence. Prior to British colonial rule, kingdoms in what is today’s Myanmar, particularly those on the western frontier such as the Arakan, maintained peaceful relationships with their diverse constituents. When settlers from the Ayeyarwady Valley (who would become today’s Rakhine ethnic group) mass migrated into the Arakan in the 11th century, Arakan governors even encouraged these settlers to intermarry with the descendants of earlier Persian and Indian traders. Further east, although King Anawratha essentially imposed Buddhism as the state religion and the monarchy as its official keeper, Muslims were present in his administration and those that followed. In later centuries, mosques were built alongside Buddhist temples in the Rakhine capital of Mrauk U. The Kaman, a primarily Muslim ethnic group, descended from a unit of archers that defended the court there in the 17th century.

Ethnicity in Burma was fluid and difficult to categorize until the British imported their system of racial classification. The British process of recording ethnicities calcified not just identities, but allegiances, along racial
Until British colonial annexation upended Burmese society, “[t]hese rulers knew the value in keeping the country’s various religious communities on side, and in doing so they cultivated a degree of communal harmony.”73

Buddhist supremacy in Myanmar originated with an upending of the existing social and economic order. Upon annexation in 1886, British Burma, a territory heretofore populated mostly by ethnic Bamar Buddhists, was administered as a province of British India.74 Encouraged by the British perception that Indians were hard workers and loyal to the regime, Indians flowed throughout the British-administrated territory, taking jobs in every sector of society wherever they went.75 As a result, Burma rapidly diversified, such that Indians outnumbered indigenous Buddhist Bamar in Yangon, the area’s largest city, within twenty years.76 By the 1930s, commentators observed that Yangon had “developed the feel of an Indian city.”77 As the Great Depression ravaged heavily indebted rice farmers, more than half of the arable land in the Ayeyarwady Delta was controlled by mostly-Indian “nonresident landlords.” Further, local Buddhist women who married Muslim men often converted and raised their children as Muslims.78 As in the United States, swift yet thorough societal and economic changes soon engendered resentment among Myanmar’s indigenous population.

For many Buddhist Burmese, the issues of independence from Britain and reversing the unfavorable demographic changes were the same. Like in the United States, perceived threats empowered extremist groups that promoted violence. Two of the first known Burmese independence/Buddhist supremacy groups were the Young Men’s Buddhist Association (YMBA), founded in 1906 with the goal of uniting various Buddhist groups to “assert a cultural identity distinct from the Western culture of the colonisers;” and the We Burman Association, founded in 1930 as racial and independence-related tensions escalated during that decade.79 A popular anti-colonial rallying cry during the 1920s and 1930s, “Amyo, Batha, Thathana,” declared the need to “protect the race and the nation, the majority language and religion, and the Sasana” (i.e. teachings of the Buddha). This phrase was printed and distributed prolifically by We Burman.80 Also around this time, the term “taingyintha,” which today is loosely translated as “national races,” became a means of distinguishing indigenous Burmese from foreigners in British Burma such as Chinese, Indians, and British.81

The British interpreted Bamar grievance against this situation as subversive resistance to British rule. In response, they further exacerbated the grievance by giving preferential treatment to minority groups regarding representation in the colonial administration and the military.82 During World War II, while the Burma Independence Army elected initially to side with the Japanese, minorities such as the Rohingya were recruited to fight with the British.83 On the eve of independence, the British even promised rights to secession for minority groups including the Muslim Rohingya and the Christian Karen and Kachin.84 After independence, however, these secession promises were summarily discarded by the first independent Burmese government under Prime Minister U Nu, leading to the first small-scale ethnic armed rebellions.85

As in the United States, policy solutions intended to promote resolution ultimately backfired as they attempted to work around, rather than directly address, underlying grievances. The new government, although not completely interested in promoting
multiculturalism, prioritized national unity. Most minority groups, including the Rohingya, were treated comparably with other ethnic minorities. In the wake of World War II, taingvintha, here translated as “indigenous races,” was used in the draft constitution of Burma to advocate for cultural rights for minorities. Taingvintha was also used to advocate for national unity: U Nu appealed to taingvintha to accept and discharge a duty of responsibility to the state.

The first National Registration Cards (i.e. citizenship papers) printed in 1952 did not ask for or mention religion or ethnicity, meaning “all who could prove a family presence in the country going back two generations or who had lived in the country for eight years prior to independence were granted citizenship, regardless of their group identity.”

However, “national unity” could not resolve the original pre-independence grievance that indigenous Burmese-ness was under threat. Almost immediately after Ne Win led a military coup in 1962 that replaced U Nu’s government with martial law, foreign influences in Myanmar were expelled from the country. Those expelled included major Western foundations, education institutions, and libraries run by the U.K., the United States, India, and Russia. The military, also known as the Tatmadaw, initiated mass deportations of Indians and Pakistanis living in Burma. Political dialogue with minorities stopped as xenophobia took root. Roughly ten years after the coup, Ne Win pushed for a new constitution for Burma that defined 135 recognized ethnic groups as indigenous (notably excluding the Rohingya). In 1988, the Burmese legislature passed a law recognizing three tiers of citizenship (full, associate, and naturalized). The law defined “full citizenship” as the ability to demonstrate clear documentation of living in Burma since before 1823, a signal that implied residency before British rule over Burma began. The choice of date, coupled with the obvious intention of excluding Rohingya Muslims, clearly demonstrates that resentment over demographic changes in Myanmar before independence from Britain was never resolved.

The end of military rule in 2010 proved a critical moment in the determination of who to include in the new democracy. More specifically, although the Tatmadaw retained significant control of the new government according to the 2008 constitution, it felt vulnerable to the direct loss of a share of its political control. The Tatmadaw’s concerns were not unwarranted: Americans who spent significant time in Myanmar found that resentment for the military was stronger after almost fifty years of hardship, and anti-minority sentiment was much more indoctrinated from the top-down than natively internalized.

As in the United States, mass media played a role in spreading and normalizing violent narratives. Extremist Buddhist groups experienced a revival around 2011 and gained support through social media. Promoting, or at least permitting, these extremist groups served the interests of the military by providing rhetorical ammunition to help direct popular grievance away from the military and towards Myanmar’s Muslims. The 969 Movement grew out of the 1988 political uprising and has had considerable influence over the NLD, the party that came to oppose the military and lead the country. The movement would instruct Buddhist shopkeepers to post stickers with the numbers 969 on their storefronts and encourage Buddhists to boycott Muslim shops. Its leader, U Wirathu, was known for giving internationally published interviews claiming that Muslims were “breeding so fast” and “stealing…raping” Burmese
women.\textsuperscript{96} When 969 fell out of prominence due to government restrictions, MaBaTha, the acronym for “Organization for Protection of Race, Religion, and Sasana,” arose in 2010 to replace it and became even more powerful than 969.\textsuperscript{97} MaBaTha’s key claim to prominence was its advocacy for a series of Race and Religion Laws grounded in anti-Muslim narratives that were ultimately passed in 2015.\textsuperscript{98} MaBaTha has also gained influence over religious education, which it uses to advance a more anti-Muslim version of Buddhism. Although the liberalization of telecommunications and the plummeting price of a SIM card around that time were expected to improve the exchange of information and therefore reduce prejudice, bias, and manipulability, instead they amplified an alliance of 969 and MaBaTha that orchestrated violence in Rakhine in 2012-13.\textsuperscript{99} However, the Myanmar population lacked adequate time to develop the internet literacy skills needed to filter through the information overload before both the military and MaBaTha began wielding the internet, and Facebook specifically, to spread misinformation and promote their agendas.\textsuperscript{100}

This violence was, unfortunately, supported by the rise of the Arakan Rohingya Salvation Army (ARSA) in October 2016. Although ARSA is just one of many ethnicity-based armed groups battling the military at any given time, its appearance in the headlines provided the perfect scapegoat around which the military could rally support for anti-Rohingya armed action. State-run media published various claims with unclear veracity about ARSA, such as ARSA alliances with foreign terrorist networks and jihadist groups, and that ARSA fighters posed as refugees as they recruited new fighters in Bangladesh. According to military chief Min Aung Hlaing, “the broader ‘Bengali issue’...had become ‘a national cause and we need to be united in establishing the truth.’” His comments were supported by remarks made by Aung San Suu Kyi and her party, reflecting the sentiment that Muslims were receiving outsized support from the international community.\textsuperscript{101}

Once attitudes of emotional vulnerability and grievance lead to violence, these attitudes are extremely difficult to dislodge. Although the government formally banned MaBaTha in 2017, its influence persists. Its views continue to be represented by the military online through partnerships with Burmese influencers and propaganda hidden among seemingly harmless entertainment pages. Facebook’s capabilities proved instrumental in fomenting popular acceptance of military actions against the Rohingya in 2017.\textsuperscript{102} In particular, Facebook helped the military gain support among Buddhists in Rakhine State who already subscribed to narratives that suggested Rohingya would eventually take away and occupy all of the land there. The Rohingya have fallen out of the headlines since the February 2021 military coup; in its wake, young urban Burmese Buddhists have begun vocalizing their growing awareness of minority group struggles and even training with minority-led forces to fight militarily with a new unit of the National Unity Government in exile called the People’s Defense Force.\textsuperscript{103} Despite a military-imposed ban on access to Facebook, Facebook identified scores of fake accounts, pages, and groups with ties to Tatmadaw-linked users, after receiving reports from civil society as recently as July 2021.\textsuperscript{104} Meanwhile, the National Unity Government in exile has publicly pledged to include the Rohingya in conversations about a new constitution and work towards birthright citizenship for Myanmar.\textsuperscript{105} However, skepticism lingers among minority groups about whether these pledges would come to fruition, as there had not been a previous
commitment to an inclusive federal government.¹⁰⁶

A long history of imperialist economic and societal divisions facilitated the takeover of anti-minority narratives and the permeation of identity-based violence in a society that once valued coexistence. Further developments may demonstrate whether and how a century of exposure to anti-“other” narratives can be overcome by the superordinate goal of defeating a common enemy.

Similarities and Implications

On the surface, the United States and Myanmar have very different histories leading to white supremacy and Buddhist Burmese supremacy, respectively. White European settlers in the Americas had no previous contact or intermingling with the African-descended peoples they would exploit, while Burmese Buddhists had centuries-long records of peaceful and even state-encouraged intermingling with Muslims. Further, differences in the permitted levels of free expression in both contexts had implications for how centralized identity-based hateful narratives would become. In part because of these two factors, the Burmese state post-independence played a much larger role in defining anti-Muslim narratives than the American state did in defining anti-Black narratives. The Burmese state created and sowed anti-minority attitudes in areas where they did not substantively exist previously, while American political figures merely amplified already virulent anti-minority rhetoric.

Despite these differences in history, however, majority group supremacy has worked psychologically and rhetorically in many of the same ways. First, both contexts were tainted by the British “obsession with racial classification” that distinguished between white and Black Americans and between Buddhist Burmese and others in Myanmar and cemented judgments of superiority and inferiority of groups.¹⁰⁷ In both the United States and Myanmar, a majority group’s sense of economic and political loss, vulnerability, and threat led to the development and promotion of negative, dehumanizing stereotypes that facilitate violence against supposedly threatening minority groups. Further, such violence is often organized through extremist groups that take advantage of popular media to sow hysteria and escalate tensions. After a period of conflict, both countries experienced a reconciliation period that was supposed to help the different groups work toward participatory democracy. In practice, however, these reconciliation periods were far too short to address the historical depth of conflict and the emotions underpinning it. Finally, by framing the genuine challenges of majority groups in a dispositional way that draws from these stereotypes throughout history, political actors stoke extant identity-based resentments for political gain.¹⁰⁸ In both contexts, these political narratives produce, and appeal to, a similar spectrum of viewpoints: people who already hold extreme views and people who are aware of the extreme views that exist. The path of extreme hate toward generalized acceptance transpired in different ways. In Myanmar, hateful narratives imposed by the state and extremist groups became mainstream among the people. In the United States, hateful narratives became mainstream from the grassroots to the political elite and back again. But for both white Americans and Burmese Buddhists, “Emotions…serve as resources (similar to weapons and money) for conflict entrepreneurs, who use them against enemies.”¹⁰⁹
Domestic and International Approaches to Combat Identity-Based Hate and Violence

While all countries, including the United States, face challenges of emotions and identity-based conflict, key differences arise in how the United States approaches solutions in domestic and international contexts. Relevant lessons can be derived from each. The next section will now analyze patterns in United States’ domestic and international approaches to combating identity-based hate and violence as a basis for its recommendations on best practices.

Key Trends in Approaches

A review of both domestic and international programs reveals key trends in how various programs seek to minimize identity-based hate. First, it is important to note the United States’ unique approach to speech protection. In the United States, hateful expressions are treated as protected speech under the First Amendment; hateful expression is not subject to legal consequences. A lack of legal consequences has facilitated a culture in which those who express hateful speech perpetually assert their “right” to make such remarks free of cultural or societal consequences. Unfortunately, when hate speech is not regarded as a problem to be corrected, expressions of hate are not treated as a warning sign for potential violence until actual violence occurs. The distinction of protected speech is less salient for organizations working overseas, which are not bound by U.S. legal culture and therefore can more easily link hateful rhetoric to anticipated violence. For these organizations, counter-extremism programming may feed into community-based social cohesion goals.

This difference has facilitated the development of a few distinct approaches, from addressing individual-level intervention, broad-scale bias reduction, direct engagement with political leaders and public figures, and others. Some organizations, such as Life After Hate and Parents for Peace, aim to rehabilitate members of violent extremist groups or prevent the vulnerable from joining one. These programs may take the view that hate is a public health problem akin to addiction. This view follows two general trends in social science research: first, that community-level violence can be treated as an “epidemic” similar to contagious disease; and second, that social cohesion can be a dependent variable for other health outcomes.

Other programs work to engage ordinary citizens with different identities in conversation to reduce ordinary prejudice. In service of this, some American organizations, such as Facing History and Ourselves, develop materials targeting individual educators or institutions. They may choose to work with universities precisely because their student bodies represent a “cross section of society” where students from diverse backgrounds and communities interact. These types of programs tend to promote social cohesion for its own sake, using the language of countering “a divided America” or of training the next generation in civic skills. Organizations such as Soliya, which hosts virtual dialogues between university students in the United States and overseas, may adopt this lens as part of their strategy for combating selection bias and inviting a wider audience to participate, as people who feel threatened by the rhetoric used to frame a conversation will not come to the table. International organizations, particularly multilateral bodies, have been able to
coordinate directly with local government entities to influence, for example, locality-wide educational curricula. Common to initiatives funded by the United States Agency for International Development (USAID) and other international donors, an overarching project in a country or community may contain a diverse mixture of program components such as dialogues, bridging activities (e.g., soccer matches between youth from different groups), trauma awareness-raising and healing, and coordination to address issues such as resource management disputes that drive identity-based grievance. These types of interventions often directly instrumentalize social cohesion programming as a means of preventing or recovering from identity-based violence. Finally, some organizations, such as Canopy of Northwest Arkansas and the Centre for Information Technology and Development (CITAD) in Nigeria, seek to directly engage lawmakers, policymakers, and other leaders and persuade them to leverage their role in social cohesion through policy, public awareness, or both.

Despite the broad-based strategies to combat violence, programs operating in both the United States and overseas shared that they face common challenges, including: (i) political will and buy-in from influential authorities, (ii) the rapid spread of misinformation and disinformation online, (iii) community pessimism or disengagement due to previous intervention failures, (iv) selection bias in program participants; and (v) the need for conflict sensitivity in intervention design.

**Recommendations**

Given the science behind vulnerability to emotional responses that beget hate and violence, differences in framing and approach, and the broad experience of practitioners and researchers, below is a non-exhaustive list of recommendations for policymakers and practitioners addressing identity-based conflict.

**Find ways to make identity less salient as a mobilization tool.** The salience of identity cleavages in daily life influences how easily identity can be weaponized by states for political ends. Reducing the salience of identity could impact how quickly and emotionally people react to changes related to their perceived identity status. This reduction could include appealing to a different identity that supersedes the identity of cleavage and that allows multiple forms of identity to exist. For example, Canopy of Northwest Arkansas, a refugee services organization, drew community support for its initial founding from common values as Christians in the face of hateful anti-immigrant rhetoric. This type of appeal is supported by research showing that in- and out-groups are malleable and context-dependent. As such, emphasizing commonalities over differences can be an effective way of countering hateful narratives. One particularly effective approach is highlighting common “superordinate” goals that are unrelated to the identity cleavage. Reducing the influence of identity may also involve rehabilitation from trauma: organizations working to support people seeking to leave extremism, including white supremacy, “go around [ideology] to go beyond it” by seeking information about events in extremists’ lives, such as trauma, that make them emotionally vulnerable to the influence of hate. Similarly, organizations working to rebuild social cohesion after violent conflict and prevent it from reoccurring incorporate trauma counseling and rehabilitation services into their programs.
Address hate and underlying conditions that facilitate vulnerability to hate simultaneously. As emotions underpinning outgroup prejudice and hate are typically reactive, isolating the prejudice and hate as a target for intervention is inadequate as is isolating the material grievance without addressing societal-level emotional responses. A USAID-funded Mercy Corps evaluative research report finds that in contexts where natural resource management is a conflict driver, building trust between groups in conflict most effectively reduces support for extremist violence when paired with other interventions that substantively address resource distribution. Multiple practitioners with experience in Myanmar shared their assessment that international programs’ inability to address the core challenges of citizenship and power imbalances undermines their social cohesion efforts. Similarly, a main critique of Search for Common Ground’s evaluation report on social cohesion programming in Tanzania is that the choice to avoid direct mention of the contentious issues to secure government buy-in came at the cost of project efficacy.

In the United States, projects encouraging racial reconciliation often fail to integrate the underlying grievance and emotion behind hateful narratives or provide opportunities to develop substantive community efforts to that end, and efforts to temper the social acceptability of hateful rhetoric and views, while experiencing modest success at a classroom level, falter under the weight of the First Amendment. However, tailored interventions to explore both hateful narratives and lifetime experiences and trauma behind a person’s motivation to join an extremist group have helped to lead extremists to disengage. Similarly, Mercy Corps finds that dialogue programs between Arab and Kurdish municipal leaders in Iraq have built sufficient trust that changing the inequitable distribution of municipal service provision was not only possible but correlated with reductions in local violence; and that interventions that only focus on facilitated interactions or public awareness are not successful if they do not address underlying material inequities or other grievances.

Take advantage of prevailing political conditions and interests as well as local community consensus. Experience with USAID programs has shown that, while the goal is to end bloodshed expeditiously, interventions that are not mindful of political conditions will fail because division almost always serves a political interest. Through its work in countering hate speech, PeaceTech Lab has observed that interventions to educate communities about harmful and hateful rhetoric are most effective in the time immediately following elections. PeaceTech Lab’s local partners have reported that after the urgency of mobilizing for an election subsides, the public is calmer and more receptive to questions about the nature of their actions during the election season and even the prospect that they might have been lied to. For Canopy, the refugee resettlement organization in Arkansas, two local conditions collided fortuitously. First, Governor Asa Hutchinson was term-limited and could not legally campaign for reelection as Governor of Arkansas. Given this, voicing opinions that ran counter to dominant political party narratives would not incur direct political risks. Second, as the 2015 Syrian refugee crisis dominated headlines, an outpouring of sympathy emerged across northwest Arkansas as it did elsewhere in the United States. These conditions supported Canopy’s advocacy to persuade Governor Hutchinson to continue to resettle refugees in Arkansas despite President Trump’s September 2019
executive order permitting state and local governments to end new refugee resettlement in their localities. Ultimately, Canopy was a participant in successfully convincing Governor Hutchinson to continue resettling refugees in northwest Arkansas. According to Emily Linn, Canopy’s former executive director, he chose to publicly defend that decision independently of any of Canopy’s demands. According to George Washington University researcher Jon Lewis, the range of solutions available to policymakers looking to combat the hate behind white supremacist violence has narrowed considerably over the last four or five years given the extent to which expressions of identity-based hate have been normalized. However, legal or institutional tools are still available to address extremist violence—for example, establishing a domestic terrorism unit in the Department of Justice.

**Identify key stakeholders influencing identity-based narratives and leverage incentives for action.** Related to the previous recommendation, it is important to work with states and other authorities and build their trust, appealing to both their emotions (as entities directed by humans) and their more tangible incentives. PeaceTech Lab has found that shaming organizations and individuals from a position outside the relevant context and without a clear plan or approach to combat the hateful narratives they spread does not work. CITAD, a partner of PeaceTech Lab, has been successful at publicly calling out politicians and other figures for using hateful language and tracking how these figures stop their use of such language afterward. While outside organizations often shout their condemnation into the void, CITAD attributes its success to the reputation it has developed within Nigeria in tracking and combating hate speech over the last six years. Similarly, key to Canopy’s advocacy efforts were relationships built at each level of authority: community leaders, then local government officials, and finally Congressmembers. Demonstrated support from each level of leadership provided Canopy credibility when appealing to the next level.

**Conclusion**

While the field of political psychology and exploration of vulnerability and prejudice as policy questions is still relatively new, this paper demonstrates that the research on human needs and the role of emotions in conflict is already aligned with the experience of practitioners in both the United States and overseas. Best practices from programs in the United States combating identity-based hate that leads to white supremacy-motivated violence can inform programs related to counter-extremism, conflict prevention, and peacebuilding abroad, and vice versa.

Leaning into sharing learnings between U.S. and overseas programs raises questions about the limits of the universality of human psychology given local cultural contexts. While some interventions or ideas may be received differently by audiences with different cultural frames of reference, some best practices from each context could be beneficial to consider universally. U.S. domestic practitioners can learn from global programs’ non-securitized (i.e., avoiding clear focus on minimizing security threats such as terrorism) approach to countering hate-related violence, as overly securitized approaches have been shown to be less effective overseas. Further, U.S. programs could, as overseas programs do, clarify goals of reducing violence through social cohesion programs, particularly as hate crimes rise in the wake of COVID-19 and possibly as a means of sidestepping concerns about
running afoul of the First Amendment.\textsuperscript{146} Meanwhile, practitioners working overseas can improve the framing of project initiatives and project work with family members of individuals engaged in identity-based violence to more precisely attract the target group to participate, and learn from U.S. programs’ framing of violence as a public health issue, which may improve likelihood of support for initiatives from host governments.

The systematic integration of behavioral psychology into interventions and the sharing of learnings between international practitioners and American practitioners is long overdue. As hateful speech and misinformation spread across rapidly developing technology platforms, those seeking to fight against prejudice and hateful rhetoric must make full use of every available tool, experience, and lesson.

\textbf{About the Author:} Jordyn Iger received her Masters in Global Human Development from Georgetown in May 2022. Her research was funded and supported by the Institute for the Study of Diplomacy’s Fellows in Diplomacy program. She is currently a Senior Project Management Associate at Chemonics and is a Returned Peace Corps Volunteer.
3 Interview with William Mark Habeeb, Adjunct Professor - Master of Science in Foreign Service, Georgetown University, November 16, 2021.
9 Mason, p. 2.
10 Mason, p. 2.
11 Mason, p. 23.
13 Mason, Uncivil Agreement: How Politics Became Our Identity, p. 86.
14 Mason, Uncivil Agreement: How Politics Became Our Identity, p. 86.
19 Amiot et al., “Bringing Together Humanistic and Intergroup Perspectives to Build a Model of Internalisation of Normative Social Harmdoing.”
20 Mason, Uncivil Agreement: How Politics Became Our Identity, p. 84.
25 Morgan, p. 93.
27 Morgan, p. 93.
32 Gates, Jr., pp. 57-59
34 Gates, Jr., p. 65
35 Gates, Jr., p. 16
36 Gates, Jr., p. 16
39 “Sharecropping, Black Land Acquisition, and White Supremacy (1868-1900) | Duke Sanford World Food Policy Center.”
46 Gates, Jr., pp 89 & 141
47 Gates, Jr., p. 67
48 Gates, Jr., p. 68-69
49 Gates, Jr., p. 67
53 Brook, “The Birth of a Nation.”
54 Brook, “The Birth of a Nation.”
60 Feola, “‘You Will Not Replace Us.’”
62 Habeeb, interview.
63 Feola, “‘You Will Not Replace Us.’”
64 Feola.

Mukherjee, “Race Relations, Nationalism and the Humanitarian Rohingya Crisis in Contemporary Myanmar.”

Cheesman, “How in Myanmar ‘National Races’ Came to Surpass Citizenship and Exclude Rohingya.”

Wade, Myanmar’s Enemy Within: Buddhist Violence and the Making of a Muslim Other, p. 56.

Wade, Myanmar’s Enemy Within: Buddhist Violence and the Making of a Muslim Other, p. 90.

Mukherjee, “Race Relations, Nationalism and the Humanitarian Rohingya Crisis in Contemporary Myanmar.”

Akins, “The Two Faces of Democratization in Myanmar.”

Interview with Allie Smith, Former Employee, CARE Inc., January 27, 2022; Interview with Steve Ross, Senior Advisor and Program Director, Richardson Center for Global Engagement, February 24, 2022.

Mukherjee, “Race Relations, Nationalism and the Humanitarian Rohingya Crisis in Contemporary Myanmar.”

Mukherjee, “Race Relations, Nationalism and the Humanitarian Rohingya Crisis in Contemporary Myanmar.”

The numbers 969 stand for the nine attributes of Buddha, six attributes of his teachings, and nine attributes of the Sangha, or the order of monks in the courts of the precolonial monarchies.

Wade, Myanmar’s Enemy Within: Buddhist Violence and the Making of a Muslim Other, p. 197.

Akins, “The Two Faces of Democratization in Myanmar.”

Wade, Myanmar’s Enemy Within: Buddhist Violence and the Making of a Muslim Other; Mukherjee, “Race Relations, Nationalism and the Humanitarian Rohingya Crisis in Contemporary Myanmar.”


Wade, *Myanmar’s Enemy Within: Buddhist Violence and the Making of a Muslim Other,* p. 76.

Busby, Gubler, and Hawkins, “Framing and Blame Attribution in Populist Rhetoric.”

Mahony, Albrecht, and Sensoy, “The Relationship between Influential Actors’ Language and Violence: A Kenyan Case Study Using Artificial Intelligence.”


Interview with Emma Jouenne, Director of Programs and Research on Extremism, Parents4Peace, March 14, 2022.


Elbeblawi, interview.


Elbeblawi, interview.


Interview with Muaiao Parpieva, Former Employee, Mercy Corps, January 28, 2022; Sarota, “Katika Usalama Tunategemeana and Pamoja! Strengthening Community Resilience in Tanzania.”

Interview with Muaiao Parpieva, Former Employee, Mercy Corps, January 28, 2022; Sarota, “Katika Usalama Tunategemeana and Pamoja! Strengthening Community Resilience in Tanzania.”

Interview with Emily Linn, Former Executive Director, Canopy of Northwest Arkansas, January 11, 2022; Hamza Ibrahim, Project Lead, Countering Hate Speech Project, Center for Information Technology and Development, February 22, 2022.

Interview with Anonymous USG Employee, February 16, 2022.

Interview with Emily Linn, Former Executive Director, Canopy of Northwest Arkansas, January 11, 2022.

Mahony, Albrecht, and Sensoy, “The Relationship between Influential Actors’ Language and Violence: A Kenyan Case Study Using Artificial Intelligence.”


Jouenne, interview.

Interview with Jonathan Lewis, Research Fellow, George Washington University Program on Extremism, February 24, 2022.


Ross, interview; Smith, interview.

Sarota, “Katika Usalama Tunategemeana and Pamoja! Strengthening Community Resilience in Tanzania.”

Nagai-Rothe, “Congressional Conversations on Race.”; Evaluation Department, “How Do We Know It Works? Researching the Impact of Facing History and Ourselves since 1976.”; Lewis, interview.


Parpieva, interview.

Anonymous USG Employee, interview.

Interview with Achol Mach Jok, Specialist, PeaceTech Lab Africa, January 28, 2022.


Mach Jok, interview.

Interview with Hamza Ibrahim, Project Lead, Countering Hate Speech Project, Center for Information Technology and Development, February 22, 2022.

Linn, interview.

Gender Roles and Military Necessity: Women’s Inclusion in the Liberation Tigers of Tamil Eelam

Kara Joyce

The Liberation Tigers of Tamil Eelam (LTTE, or the Tigers) were unique in many ways, one of which being their inclusion of women in combat roles during the Sri Lankan Civil War. This paper applies Jennifer Eggert's 'factors influencing women's inclusion in terrorist organizations' theory to the LTTE. Eggert proposes three factors that influence whether or not women are included in combat: the level of pressure applied by the opposing force, the availability of male manpower, and the degree of support from external actors and media. This paper considers how these three factors change over different eras of the conflict and analyzes how women's participation varies accordingly. Women's inclusion in combat during the civil war did change according to changes in pressure, support, and available manpower, supporting Eggert's thesis and opening up a new pathway of analogous research in the realm of insurgency and counterinsurgency. This research implies that gender roles, sometimes viewed as immutable or natural, can and will change when military necessity demands it. Future avenues of research include understanding the bureaucratic politics within the LTTE to determine how these changes were applied, as well as analyzing gender relations within the LTTE to understand the intersection of the personal and the political.

Introduction

The Sri Lankan Civil War raged for almost thirty years, killing 70,000 and displacing hundreds of thousands of civilians while the Sri Lankan (Sinhalese) army and the Liberation Tigers of Tamil Eelam (LTTE) fought. From 1983 to 2009, the Sinhalese (the ethnic majority in Sri Lanka, making up 75% of the population) and the Tamil (16% of the population) fought for control of ancestral Tamil territory, with four attempts at peace having failed. The conflict ended in 2009 when the Sri Lankan army killed the founder and supreme commander of the LTTE, Velupillai Prabhakaran, and secured a complete military victory. This is a common enough story, but there are some unique aspects to it: one being that one-third of the Tigers were women.

The Tigers included women in all aspects of their organization, from combat to naval expeditions to logistics. Although their total numbers were a military secret, some researchers think that women may have made up as much as one-third of all forces. This article will attempt to fill a gap in the scholarship about the Tigers’ organizational culture by answering the question of why the Tigers included women so readily in their forces despite Tamil cultural norms that would indicate otherwise. Tamil culture is highly patriarchal, with women’s roles being primarily domestic; so why did the Tigers break with tradition and expand the roles that women could play in the organization? This paper argues that the Tigers included women in all aspects of their organization because they faced a manpower deficit and lacked steady external support in the face of a well-funded and powerful adversary. This change temporarily expanded the roles of women in Tamil culture, but did not impact the outcome of the war.

Drawing on Jennifer Philippa Eggert’s work studying gender roles in terrorist organizations and Lee Ann Fujii’s
framework of culture as a script to be written and acted out, this paper will argue that the Tamil Tiger leadership changed the gender script, expanding the permitted roles of women to achieve military and political goals during the conflict. I will compare the roles of women in the organization over time in the context of Eggert’s three most important factors: the level of pressure applied by the opposition, the availability of male manpower, and the level of external support from potential supporters. I will review the relevant literature and describe the research methodology; discuss the case study through the chosen lens; rebut some criticisms of why women were included; provide alternative explanations for this inclusion and posit future avenues of research; and consider the end of the war and what it meant for Tamil women.

**Literature Review**

*Women in Combat*

War and warfighting are largely considered masculine institutions, but women have participated in both for all of human history. Therefore, this research is less focused on why women would choose to participate than on the factors that influence why masculine organizations choose to include women. Darden, Henshaw, and Szekely also ask, “Why do insurgencies recruit women?” in their book *Insurgent Women*, and focus on case studies in Ukraine, Colombia, and the Kurdish insurgencies in Turkey, Syria, and Iraq. This research follows in their footsteps by asking the same question in Sri Lanka.

*The Mutability of Gender Roles*

This research belongs to the poststructuralist tradition because it takes as given that gender is a construct dependent upon context. Structuralism is “a theoretical approach that identifies patterns in social arrangements,” and poststructuralism “holds all meaning to be fluid rather than universal and predictable.” Michel Foucault, one of the key scholars in this field, believed that “the operation of power temporarily stabilizes meanings into a discourse.” In other words, gender is a pattern in social arrangements, but it is fluid, based on context, and can be manipulated. The social power that the LTTE held gave them the ability to define gender in the Tamil cultural text and then change it when necessary. Poststructuralism is notoriously controversial in international affairs, but it is important to this case study so that we may acknowledge the inherent fluidity in gender as a social construct and its importance in conflict.

To delve deeper into the poststructuralist idea of culture as text and power-holders as writers, this research will build on Lee Ann Fujii’s concept of a script as a way to understand violence between ethnicities as defined in her groundbreaking book, *Killing Neighbors: Webs of Violence in Rwanda*. She viewed “state-sponsored ethnicity not as an external force that acts on people, but as a ‘script’ for violence that people act out.” Because “the rules and expectations for normal behavior change, sometimes in radical ways” during periods of violence, power-holders can take advantage of these shifts to create new scripts that serve their political goals. Her work is specifically about the actions that individuals take when called to genocide, but I will use her theoretical framework in this research to frame the options that the Tigers offered to Tamil women: to join the Tigers or to remain a civilian.
Changing Gender Roles in Terrorist Organizations

Jennifer Philippa Eggert’s article, “Women Fighters in the ‘Islamic State’ and Al-Qaeda in Iraq: A Comparative Analysis,” is crucial to this analysis. She concluded the Islamic State (IS) only changed its stance towards women’s inclusion when a worsening security context demanded it. The three variables she identified as influential on whether women were included were the level of pressure by opposing military forces, possession of sufficient male manpower, and how the media covered the conflict along with the degree of support from external actors. While the social context and organizational goals differ between IS, Al-Qaeda, and the LTTE, the similarities between these terrorist organizations fighting a state military make this a useful framework through which to analyze the factors that influence Tamil women’s inclusion. Based on Eggert’s findings, this paper will argue that the Tigers changed the script for Tamil gender roles in response to these three factors.

Methodology

To demonstrate the connection between the external circumstances facing the Tamil Tigers and the gender script they followed, I will discuss three different eras of the civil war: from the 1970s to 1987; from 1987 to 1990; and from 1990 to the end of the war in 2009. In each section, I will analyze the role of women in the organization, the manpower and external support available to the Tigers, and the counterinsurgency capabilities of the Sri Lankan government.

Pre-Conflict Tamil Gender Roles

Women in Tamil culture had and have a “respected but simultaneously ambivalent and somewhat restricted status.” They were celebrated for being “weak, chaste, shut up in the home and ignorant about the world, and shy.” Though schools for girls were established in the 1800s, their education was for the benefit of their future husbands so they would make good wives. Similarly, women gained social standing or respect for their roles as wives or mothers to men. Women who later became LTTE soldiers also said that there were cultural superstitions that prevented them from pursuing certain activities in their daily lives. They were told by elders that “girls should not climb trees, go out alone or ride bicycles.”

In the script of Tamil culture from before the war, women were discouraged from masculine activities and encouraged to focus primarily on domestic and reproductive pursuits. This script was left unchallenged by the Tigers in the first phase of the war. Subsequent sections will discuss how this script was changed over time based on the needs of the Tigers.

Phase One of Women’s Inclusion (1970s-1987): “Nice to Have”

Women were recruited and trained in the LTTE during this era (known as Eelam I), but only in limited numbers because it was not a military necessity. While the Tigers were not going to turn down women who wanted to fight for Tamil sovereignty, the leadership was in a strong enough position so as not to require changing the gender script at this time.

This first phase of the civil war began in earnest in 1983 when the Tigers killed 13 Sri
Lankan soldiers. This violence resulted in 400 to 2,000 Tamils being killed in response. After this original catalyst, the Tigers issued four principal demands: “first, that the Tamils of Sri Lanka be recognized as a distinct nation; second, that the northeast of the country be recognized as their historical homeland; third, that the Tamil population be allowed the right of self-determination and finally, that all Tamils be granted Sri Lankan citizenship.” These terms were unacceptable to the Sri Lankan government, and so the violence ensued.

**State of Manpower of the LTTE**

When Tamil-Sinhalese violence began in the 1970s, there were five main Tamil separatist groups on the island—the Tigers, the Tamil Eelam Liberation Organization, the People’s Liberation Organization of Tamil Eelam, the Eelam People’s Revolutionary Liberation Front, and the Eelam Revolutionary Organization of Students. All of these organizations recruited women in the 1980s for propaganda, medical, and logistical roles. During this time, the LTTE committed fratricide against the other organizations to consolidate control over the available manpower. The LTTE recruited men from the other organizations as they disbanded, making themselves the only group left to join. The broad availability of experienced Tamil men made recruiting untrained women unnecessary.

**Economic Support for the LTTE**

After the Tamil killings in 1983, thousands fled to India. The Indian government and the Tamil diaspora both provided support to the LTTE. They provided “money, sanctuary, and training camps” for the Tigers, with over 20,000 militants trained in Indian camps. This level of support, both from their own people and a foreign government, ensured that the Tigers did not have to change their strategic outlook at this time. This support continued through the end of Eelam I.

**Strength of the Sri Lankan Opposition**

The Tigers also did not originally need to recruit women because of the ineffectiveness of the Sri Lankan government. First, the government did not take the threat of the LTTE seriously, and it underestimated how serious the demand for independence was and how willing the Tamils were to join the organization. Second, government forces were untrained in counterinsurgency tactics, relying instead on ineffective conventional warfighting means. Third, the military’s reliance on collective punishment strategies, such as the forced displacement of Tamils from the capital city, only made LTTE recruitment more effective. Even with only half the available population participating as combatants, the Tigers were able to easily hold off the Sri Lankan soldiers.

**Outcome**

Due to the ineffectiveness of the Sri Lankan army, the ongoing support from diaspora Tamils and the Indian government, and the plentiful manpower available from the Tamil community, the LTTE leadership did not need to consider changing the gender script at this time. Although they did make the organization open to women—the LTTE women’s corps was officially created in 1983—they did not focus on women’s recruitment. In response, only a few women joined in the first phase of the conflict. They were trained in mining, explosives, weapons technology, and other combat-related skills, but their numbers were limited. This would all change during the next period of the conflict: the occupation by the Indian Peacekeeping forces.
Era of Change: 1987-1990

This brief interlude changed the strategic calculus of the Tiger leadership: the loss of Indian support after the Indian withdrawal from Sri Lanka and the deaths of LTTE soldiers without any competitive Tamil militant groups from which to poach new ones created a new need for support and manpower. During this era, the LTTE leadership institutionalized the Women’s Front and created a new script for Tamil women to follow: they could support their people and the fight for their homeland by joining the Tigers as combatants.

Although the Indian government was originally a staunch supporter of the Tigers, this sentiment shifted during the Indian occupation of Sri Lanka. In 1987, the Indian government pressured the Sri Lankan government to sign a peace agreement with the Tigers, which was enforced by Indian troops left on the island as peacekeepers. The Tigers, who had started to run a quasi-state in Tamil-majority lands, refused to cede control of land or institutions to the Indians. The Sri Lankan army resented the foreign interference, and the Indian government’s domestic audience did not support the expenditure of blood and treasure. This phase of the war ended in March 1990 when India elected a new government that opposed the campaign and the peacekeepers were recalled home.

Changes in Manpower

The Indian Peacekeeping forces, which previously had been on the side of the LTTE, lost control of the Tigers and were unable to disarm them. During the occupation, the goals of the Indian forces changed from peacekeeping to counterinsurgency by attrition. This new policy was considered a success solely because the Indian army was able to wear down the LTTE; however, the LTTE countered by sliding between conventional and guerrilla tactics as well as refilling their ranks with women. In 1987, the Women’s Front was given its own training camp by Prabhakaran and was large enough to have its own leadership structure by 1989. The change in allegiance by the Indians and their ability to use conventional warfare, alongside their effective counterinsurgency campaign to kill Tamil soldiers, made recruiting and training women necessary for the survival of the LTTE.

Changes in Economic Support

The end of support and training from the Indian government drastically undercut the resourcing capabilities of the Tamil Tigers. This shift from government-funded external support to relying only on the diaspora for support created a new need for domestic recruitment. This reduction in funding stressed the leadership, causing them to lean on the Tamils in Sri Lanka for more support. One way that they asked for support was for families to give at least one member. Whereas before they only asked for sons, they now asked for anyone.

Changes in Opposition Capabilities

During the period of Indian occupation, it was less the capabilities of the Sri Lankan army than the presence of the Indian army that created the need for more manpower in the LTTE. The LTTE had to prepare to fight its former ally, which was blocking them from achieving their aim of political independence by trying to force disarmament before the LTTE deemed it ready to do so. Relative to the period before, the LTTE was under much more pressure and needed new options to continue fighting.
Outcome

These contextual changes provided an impetus for the LTTE to update the Tamil gender script: they decreed that women can and should fight for their homeland against both the Indian and the Sri Lankan armies, which had violated them and stolen their freedoms. It was an effective strategy: the number of women in the LTTE dramatically increased from the 1980s to 1990 and beyond. This change was visible in the newly published goals of the Women’s Front in 1991: “to secure the right to self-determination of the Tamil Eelam people and establish an independent democratic state of Tamil Eelam; to abolish oppressive caste discrimination and divisions, and semi-feudal customs like dowry; to eliminate all discrimination against Tamil women and all other discrimination, and to secure social, political and economic equality; to ensure that Tamil women control their own lives; and to secure legal protection for women against sexual harassment, rape, and domestic violence.”

One brutal but effective example of the way that LTTE changed the gender script is with regard to sexual assault. The traditional Tamil gender script called for women who had been dishonored through rape to isolate or kill themselves because they could no longer be a virgin or get married. Instead, the LTTE gave women the opportunity to get revenge. The LTTE also changed the script for men: in the 1980s, male leaders spoke out against sexual violence only because the shame of it made Tamil women unfit to marry Tamil men. However, when the Women’s Front was established and institutionalized, things changed: the leadership encouraged LTTE veterans to marry women who were raped to ensure their social inclusion, and told women that there was no shame in being raped because it was not their fault. This dramatic and progressive change may have been partially inspired by the LTTE’s Liberatory principles, but it had strategic military benefits as well.

Phase Two of Women’s Inclusion (1990-2009): “Need to Have”

Through the last four phases of the conflict from 1990-2009 (Eelam II, Eelam III, Ceasefire Agreement, and Eelam IV), the LTTE relied heavily on women to fill its ranks, especially in the Black Tigers (the suicide bombers) and the Sea Tigers (the navy). This reliance persisted through the end of the conflict. Try as they might, the Tigers were unable to overcome the superior funding and manpower of the Sri Lankan army, especially in the face of defection and the destruction of the tsunami in 2004.

With the departure of the Indian army in 1990, the LTTE returned to fighting the Sri Lankan army for the next nineteen years. The two sides took turns holding, losing, and reclaiming territory. It was only after the Sri Lankan army outspent and outmanned the LTTE on the heels of a huge defection and the 2004 tsunami that the conflict came to an end with a Sri Lankan victory. The number of women who participated in the Tigers remained steady at around one-third of the whole organization through 2009.

State of Manpower of the LTTE

Manpower became a serious issue for the Tigers; with the Indians gone and the other Tamil separatist groups decimated, they running out of new soldiers after seven years of conflict. However, they might have been able to hold their land and maintain the stalemate had there not been two disasters in short succession: a massive defection and a tsunami.
In 2004, the first high-level defection broke the ranks of the Tigers. Colonel Karuna, who had commanded the Eastern theatre, defected to the Sri Lankan side along with 3,000-6,000 Tiger cadres. His defection severely strangled the LTTE’s geographic reach and strategic depth while engendering a leadership crisis, and the intelligence that he provided to the Sri Lankan army gave them an advantage in future conflicts. The December 2004 tsunami disproportionately impacted the Tamils, who held much of the land on the western coast. 40,000 people died, including 3,000 Tiger cadres, and one-quarter of the Tiger’s naval fleet was destroyed. These two disasters severely decreased the manpower available to the LTTE, ensuring that they would accept anyone willing to join.

Economic Support for the LTTE

Another disaster outside of the control of the LTTE severely harmed their economic stability. Much of their funding came from expatriate and diaspora communities in other countries, but in 2002 the international community passed various domestic laws that strangled their funding sources. The global response to the 9/11 terrorist attacks was increased support for the International Convention for the Suppression of Financing of Terrorism; as the Tigers were a terrorist organization, the ratification of this convention cut them off from millions of dollars while the Chinese supplied the Sri Lankan army with over $1 billion in aid. This decline in support, both relative and absolute, also ensured that the LTTE would not alienate any of their surviving supporters by reverting to the pre-conflict gender script or rejecting the work that women did for the cause.

Strength of the Sri Lankan Opposition

The nail in the coffin for the LTTE was the increase in pressure that the adversary was able to apply in the last years of the war. By increasing the defense budget by 30% over five years (2005-2009) and increasing the size of the armed forces by almost 350,000 people, the Sri Lankan government was able to use its sheer material preponderance to overwhelm the Tigers.

Outcome

By the end of the war, the LTTE was so desperate for manpower that it had resorted to kidnappings and forced conscription. The last phase of the war, Eelam IV, lasted from the summer of 2006 to May 2009. The Sri Lankan army was able to capture, clear, and hold territory, slowly pushing the LTTE forces back into a small corner of territory where they were outgunned and defeated, with Supreme Commander Vellupillai Prabhakaran killed. In the end, the same factors that drove the LTTE to recruit women in higher numbers were the ones that led to their defeat: the lack of manpower and economic development were the two deciding factors that led to the Sri Lankan victory. Changing the gender script had helped the Tigers increase recruitment in the early days of the war, but it was not enough to counteract the size imbalance between the ethnic minority Tamil and the dominant Sinhalese army. Based on these two snapshots in military capabilities and gender scripts, as well as the LTTE’s willingness to transgress social rules to stave off defeat, it seems highly likely that the shifts in gender roles in the Liberation Tigers of Tamil Eelam throughout the conflict was in response to military necessity.
Feminist Critiques of Women Combatants and the Women’s Response

Despite documented testimony by women fighters cataloged in the body of research, some scholars imply that joining the LTTE was not real empowerment; i.e. just because the LTTE said that women could do these things now did not mean that the women were exercising free choice. One article on women in LTTE concluded that “...governed by a rigid code of conduct, women are reduced to soldiers obeying orders which is not inherently empowering or liberating.”

Miranda Alison, however, calls the binary of victim/agent or liberated/oppressed “unnecessary and unsophisticated.” These women were born into a patriarchal and restrictive society, and while they were made vulnerable by conflict, they made decisions to shape their own lives by providing for their own security. With the exception of those who were kidnapped or conscripted into service, these women authored their own destinies. Arguing that they were disempowered because they joined a male-dominated organization ignores the reality of the male-dominated lives that they had before.

Alternative Explanations for Changes in Gender Roles

Eggert’s model for factors of change in terrorist organizations is a compelling theory. It does not, however, cover all the variables that might impact gender dynamics within a terrorist organization such as the LTTE, nor does this brief case study conclusively prove that military necessity was the only reason that the LTTE changed its gender roles. It would be remiss not to mention some of those here as opportunities for future research.

This research cannot prove causality. While it seems likely that women joined in response to public statements based on the timelines of LTTE organizational change, perhaps it was the existence of women volunteers that inspired the change. A future research question may ask: are military organizations responsive to civil society’s change and demands?

The bureaucratic politics model would also be interesting to apply to this case: which of the men who founded the organization alongside Velupillai Prabhakaran pushed for women’s inclusion? Who fought against it? Did the early women, those who contributed from 1983 onward and founded the Women’s Front, negotiate for an added focus on the recruitment of women? These questions could best be answered by a scholar who speaks Tamil, has access to the community, and could interview survivors. Unfortunately, when Prabhakaran was killed in 2009, so too was the opportunity to ask him about his life’s work. The same questions could be asked about the different branches; the Women’s Front naturally had a preponderance of women, but the Black Tigers and the Sea Tigers were also both known for being significantly female.

What level of influence did those organizations have in decision-making as a whole, and how did women come to dominate them?

Other avenues for future research would include more consistently analyzing speeches and publications from the LTTE for gendered language usage over time and mapping when women’s liberation became part of the official rhetoric. Similarly, analyzing this conflict from the Sri Lankan government’s perspective could provide context for other influencing variables: why did the Sri Lankan army commit so much sexual assault, when the LTTE did not? What did the Sri Lankan government think...
of Tamil women’s participation, and did they ever consider offering more freedoms to both Tamil and Sinhalese women to weaken support for the LTTE? Finally, applying an overtly feminist constructivist lens to this conflict may offer keener insights into gender roles within the Tigers, within Tamil culture, and within Sri Lanka as a whole.

Military Defeat: What the Sri Lankan Victory Meant for Tamil Gender Roles

In 2004, Colonel Thamilini (real name Subramaniam Sivakami), the leader of the Women’s Front, said that though the war provided significant changes to the roles of women, they could not be considered final until the peace negotiations were completed. She thought that women’s new role could only be cemented through a new definition accorded by legislation. She was right—but the peace talks never occurred.

There was no negotiation period for the Tigers wherein they could institutionalize the rights of women because the Sri Lankan government won outright. In the language of Dr. Fujii’s framework, the Sri Lankan government got to write the script. Since Sri Lanka and the Sinhalese had not changed their traditional gender roles during the conflict the way the Tigers had, they had the power to make the female Tigers conform to the Sri Lankan script and traditional gender roles. For example, “after her release from a rehabilitation camp for former Tamil Tiger fighters in 2011, one former member of the women-only -Malathi Brigade was presented by Sri Lankan government officials with a sewing machine.” These women now face the cultural isolation and stigmatization from which the Tigers protected them; it will take a careful and gender-sensitive reintegration process to mitigate this issue. Though this paper has

focused on the gender script of the Tigers, terrorist organizations are not the only ones capable of changing gender roles and dynamics. Non-violent organizations would do well to challenge the dynamics that prevent them from taking advantage of all possible resources—the power and ingenuity of women included.

Concluding Thoughts

This research contributes to the body of literature on the Sri Lankan civil war and the Liberation Tigers of Tamil Eelam by considering the manipulation of gender roles as a military strategy. It also creates space for thinking about post-conflict gender roles in other situations. All eyes are on Ukraine right now, where Ukrainians (including many women) are fighting for their sovereignty against the Russian invasion. Should Russia win this war without resorting to negotiation, it will write the script for Ukraine’s culture afterward. Ukraine has, in recent years, come closer to the West culturally; Russia promotes itself as the last great bastion of conservatism where masculinity is celebrated. What would a Russian victory mean for gender equality and the rights of homosexuals in Ukraine? How would Russian values be exported and enforced in a newly conquered Ukraine? This area of research deserves timely attention for the sake of those who might have to live under those conditions.

Though the Sri Lankan government was able to win its war, other militaries around the world face insurgencies within their own borders without the ability to spend more money or find more soldiers. Understanding the gender dynamics that shape insurgency recruitment and commitment is critical to fighting smarter counterinsurgency campaigns and in analyzing conflict.
About the Author: Kara Joyce (she/her) is a writer and editor focused on the nexus of climate change, gender security, and public policy. She earned a Master of Arts in Security Studies from Georgetown University (‘23), and Bachelor’s degrees in English Literature and International Affairs from the University of Georgia (‘18). Her current independent research project is creating a database of gendered risk factors related to anti-abortion terrorism.


Oliker, Olga. “Fighting While Female: How Gender Dynamics are Shaping the War in Ukraine,” Foreign Affairs, November 21, 2022.

Spies in the Skies: Analyzing the Development of the U.S. Commercial Satellite Industry vis-à-vis the Hermit Kingdom

Sue Kim

This paper explores the past, present, and future of the U.S. commercial satellite imagery industry vis-à-vis the Democratic People’s Republic of Korea (DPRK, North Korea). The first section will examine the government monopoly of overhead imagery before the development of the commercial satellite industry in the 1990s and explain how U.S. military reconnaissance aircraft and satellites were used during the Pueblo Hostage Crisis of 1968. Next, it will address how non-government actors have used and benefitted from using high-resolution commercial satellite imagery to analyze North Korea’s military development and the domestic environment with reference to specific think-tank programs and reports. The third section will assess the dangers and risks of global transparency and “democratized intelligence” that have emerged from non-government actors and adversaries’ use of publicly available high-resolution commercial satellite imagery. Lastly, this paper will close with future opportunities that commercial space companies have in both countering the North Korean threat and offering internet services to average North Korean citizens by examining SpaceX’s Starlink services.

Introduction

In the 1960s, former Secretary of Defense Robert Gates described North Korea as the “toughest intelligence target in the world.”1 From the Korean War to the present day, the U.S. government’s overhead imagery, provided by military reconnaissance aircraft and satellites, has been a critical source of information in observing and analyzing this black box of an intelligence target, which has been invaluable compared to many other intelligence-gathering operations. With the development of reconnaissance planes and intelligence satellites in the twentieth century, the United States—to a limited degree—was able to circumvent the persistent problem of inadequate and inaccurate information about North Korea. With the commercialization of satellites in the 1990s, many non-governmental organizations (NGOs), universities, news outlets, businesses, and even individuals have taken the opportunity to analyze North Korea with publicly available satellite imagery, public access to internet-based mapping platforms, and the widespread availability of web 2.0 cartographic technology. As more actors beyond the U.S. government can observe what is transpiring within the regime’s territory, and as higher resolution imagery offers ever-increasing clarity of the ground below, there is a growing expectation that an individual has the power or “enhanced vision” to resolve geopolitical conflicts.2

However, a shift towards global transparency and “democratized intelligence” made possible through the commercialization of the satellite industry is precipitating North Korea to demonstrate greater hostility towards American reconnaissance activities in the region. Such a reaction only increases the chances of non-military and non-government U.S. aerial and space assets becoming targets of North Korea’s aggressive provocations. However, the regime, to date, lacks advanced space capabilities to extend its asymmetric strategy—such as terrorism, cyberattacks, cryptocurrency heists, and the development
of nuclear weapons—that it enjoys in the terrestrial domain. Therefore, the United States must dedicate more satellite imagery collection and analysis to North Korea’s domestic activities—including both military and civilian developments—and explore possibilities for commercial actors to provide internet access to ordinary citizens with the goal of undermining the Kim Jong-un regime. Analyzing how SpaceX’s Starlink constellation demonstrated the potential of commercial satellites in a conflict zone, this paper will conclude by examining future opportunities for commercial vendors and how cooperation with U.S. treaty allies, such as Japan and South Korea, can deploy information-based operations targeting North Korean civilians.

The U.S. Government’s Monopoly and Use of Overhead Imagery

Overhead imagery to observe the geographical landscape of the Far East and the Korean Peninsula initially emerged during the Korean War when “American and ‘free-world’ forces encountered Soviet and ‘Red Chinese’ Communist forces in what Americans saw as remote lands.” As a result, photo interpreters like David S. Doyle, at the National Photographic Interpretation Center (NPIC) responsible for interpreting images of Soviet territory from KH-4 CORONA satellites—the first series of American strategic reconnaissance satellites operated by the Central Intelligence Agency (CIA) since 1960—had his first aerial view of the earth while serving as a U.S. Air Force pilot during the Korean War. The U.S. investment in satellite technology as early as the late 1950s proved to be crucial in penetrating adversaries’ territories and bolstering U.S. national security in the subsequent decades of the Cold War.

A U-2 spy plane’s confirmation of nuclear missile sites in Cuba is the better-known account of aerial reconnaissance during the Cold War in 1962. However, an A-12 Oxcart’s photographic intelligence of North Korea prevented the Pueblo Hostage Crisis of 1968 from spiraling into a Second Korean War or “another Vietnam.” The A-12 Oxcart—the CIA’s high-speed, high-altitude reconnaissance aircraft and a precursor to the U.S. Air Force’s SR-71 Blackbird—had originally been designed to replace U-2 spy planes operating over Soviet territory. However, the A-12, which is “the fastest and highest flying piloted operational jet aircraft ever built” to date, was first deployed in 1967 to record potential surface-to-surface missile installations in North Vietnam as part of a covert surveillance program named BLACK SHIELD. As the Tet Offensive was nearing in Vietnam, the USS Pueblo (AGER 2) was on a signals intelligence-gathering mission off the coast of Korea to “monitor DPRK radio and electronic transmissions.” However, the USS Pueblo was captured by North Korea on January 23, 1968, for allegedly entering their territorial waters without permission. With authorization for BLACK SHIELD Mission BX-6847, an A-12 made three high-speed passes over Wonsan Bay in North Korea three days after the seizure, and the processed images from Japan were finally delivered to Washington two days later. Despite the high risk of getting shot down by Chinese and Soviet air defense systems, the A-12 was able to initially collect two pieces of information from its flight. First, the overhead imagery locating the USS Pueblo at 11 miles north of Wonsan Harbor confirmed the vessel’s last communication with the National Security Agency whereby it had “been requested to follow into Wonsan” before being intercepted by three guided missile patrol boats (Komar PTG).
Second, there was no discernible damage to the USS Pueblo anchored in Wonsan Bay.\textsuperscript{11}

Image 1: Imagery of the USS Pueblo taken by A-12 reconnaissance aircraft on January 26, 1968, at 10:10 am local GMT. The A-12 was traveling at 1,863 knots ground speed. The imagery from the Type I camera is estimated to be showing one-foot resolution.

Image 2: Controlled Image Base (CIB) imagery of the Wonson area taken in 1995.\textsuperscript{12}

In addition to supporting the initial assessment of the USS Pueblo seizure, the A-12’s overhead imagery answered what Washington craved to know most: the possibility of a full-scale North Korean invasion of South Korea. At this time, the United States government feared another war on the Korean Peninsula, as North Korea’s call for unification intensified with an unusual number of DPRK military guerilla incursions into South Korea not seen since the Korean War and the signal of hostile intent with an unsuccessful attempt to assassinate South Korean President Park three days before the Pueblo seizure.\textsuperscript{13} As the Lyndon B. Johnson administration explored more than a dozen options both diplomatic and military in response, Mission BX-6847 was further tasked to obtain a “comprehensive baseline coverage of most of North Korea’s armed forces and industry as well as large portions of the transportation system” to determine whether North Korea was planning for a ground offensive.\textsuperscript{14} As tensions between the two Koreas worsened, the Pentagon considered multiple retaliatory plans including a naval blockade, preemptive air strikes, and even the use of nuclear weapons.\textsuperscript{15} Fortunately, imagery analysts confirmed that neither a military buildup or the unusual activity that is typically observed before an all-out offensive were detected near the demilitarized zone. However, interpreters cautioned “that a large number of tunnels, caves, and underground facilities complicated the determination of North Korean force levels.”\textsuperscript{16} Satellite imagery analysts today similarly struggle to identify and track military developments that are intentionally hidden under its terrain from foreign observers, thereby making North Korea the “ultimate nightmare of an intelligence target.”\textsuperscript{17}
Just as how A-12 intelligence “often directly influenced LBJ’s decision to commit to air raids during the Vietnam War,” President Johnson was able to rely on imagery intelligence throughout the Pueblo crisis to negotiate with North Korea.\(^{18}\) Although a series of negotiations lasted for almost a year to procure 82 surviving crew members’ release—after 335 days of captivity and torture—the United States was able to resolve the crisis without resorting to military options with the help of A-12’s overhead imagery.\(^{19}\) After all, the United States “would have found it difficult to afford (either politically or financially)” a second military front in Asia at the height of the Vietnam War.\(^{20}\) Yet, as a further testament to North Korea’s heightened sensitivity towards U.S. reconnaissance operations during this period, a DPRK MiG fighter shot down a U.S. Navy intelligence aircraft EC-121 only four months after the USS Pueblo crisis, killing all 31 American crewmen on board.\(^{21}\)

Although the CORONA satellites were also tasked to “enhance coverage of Wonsan” during the hostage crisis, the resolution of its images were lower than those captured by the A-12.\(^{22}\) Although satellite imagery could not offer the high-resolution images that Washington needed during the USS Pueblo Crisis, satellite imagery technology only improved over the next few decades. With such developments, reconnaissance aircraft, which were always vulnerable to shoot-downs from high-speed interceptors, were able to retire soon thereafter.\(^{23}\) During this period of technological development in the Cold War era, states maintained a monopoly on the collection and analysis of satellite imagery intelligence. However, only the geopolitical superpowers, the United States and the Soviet Union, had unique access to high-resolution reconnaissance satellite imagery, primarily for military applications.\(^{24}\) As Timothy Barney describes, “Traditionally, the state (especially the military) informed how we saw the world and at what resolution we would see it, and so much of that could be censored or classified.”\(^{25}\) Because of the sensitivity of such intelligence activities, the existence of the National Reconnaissance Office (NRO)—a member of the U.S. intelligence community responsible for satellite imagery collection and analysis alongside the National Geospatial-Intelligence Agency (NGA)—was classified to the public for over 30 years until 1992. The commercial satellites in the nascent space industry in the 1990s reportedly had resolution and computational capabilities behind those of the U.S. intelligence agencies by approximately 20 to 30 years.\(^{26}\)

**The Current Use of U.S. Commercial Satellite Imagery of North Korea**

The race to launch a new generation of commercial satellites in the United States began in 1994 when the Clinton administration lifted technical restrictions on private companies to develop satellites whose capabilities would eventually rival governments’ “military spies in the sky.”\(^{27}\) The first two American companies to receive licenses from the Commerce Department to operate commercial satellites were Space Imaging Corporation and DigitalGlobe, which have since become incorporated into Maxar Technologies.\(^{28}\) In 1999, Space Imaging Corporation launched the world’s first commercial satellite, built by Lockheed Martin Corporation, Ikonos, which captured 1-meter high-resolution satellite images of the ground below. Ikonos was regarded as “the world’s most powerful civilian spacecraft” at the time, and competition among new vendors to win military contracts to support U.S.
reconnaissance missions blossomed.\textsuperscript{29} As expected, DigitalGlobe launched its Quickbird satellite, which was able to capture even higher resolution images of 61 centimeters, in 2001.\textsuperscript{30}

In 2003, the U.S. Committee for Human Rights in North Korea (HRNK) published a report called \textit{The Hidden Gulag: Exposing North Korea's Prison Camps} with the donation of North Korean satellite imagery taken by Ikonos and Quickbird between 2001 and 2003.\textsuperscript{31} At a time when the second Bush administration was in the early stages of formulating U.S. policy towards North Korea, the “recently obtained high-resolution, commercial satellite images of seven North Korean prisons and prison camps” helped the Committee identify the exact locations of North Korean replicas of Soviet-style labor camps and expose the horrific conditions of at least 36 prison camps.\textsuperscript{32} Corroborated by North Korean defectors’ claims about the North Korean regime’s practice of forced labor, such photo evidence directly refuted the leadership’s continuous denial of the existence of these camps. As a result, this report was able to influence U.S. foreign policy that had previously overlooked the problem of the DPRK’s human rights violations as it prioritized denuclearization above all.\textsuperscript{33} Furthermore, this report was revolutionary because it was the first instance in which an NGO used commercial satellite imagery to record sites of human rights violations in the DPRK.\textsuperscript{34} In the second edition of the report, published in 2012, the author David Hawk was able to use even higher resolution images made available by Google Earth to release the exact coordinates and satellite images of 41 gulags, once again raising awareness about the human rights violations in the DPRK.\textsuperscript{35} In the latest HRNK report published in 2021, which specifically focused on Political Prison Camp No. 14, satellite imagery analysis of agricultural fields, livestock, and construction of facilities taken by Maxar DigitalGlobe and the European Space Agency (ESA)’s Copernicus Sentinel satellites “confirm[ed] an expansion or reorganization of guard forces, a modest increase in the prisoner population, or both” in the past few years.\textsuperscript{36}

The Transitional Justice Working Group (TJWG), a human rights documentation and advocacy NGO based in Seoul, South Korea, is another example that uses modern tools such as Google Earth satellite imagery and Geographic Information System (GIS) technology to map and analyze public execution sites, burial sites, and cremation sites in North Korea. An updated TJWG report published in June 2019 called \textit{Mapping the Fate of the Dead} did not reveal the precise coordinates of these state-sanctioned killing sites, considering the sensitive nature of the data, but showed that the northernmost province of the country, North Hamgyong, which shares a border with the People’s Republic of China (PRC), had the greatest number of reported killing and body disposal sites by province.\textsuperscript{37} In addition to geographical coordinates, TJWG highlighted the power of geospatial analysis, which offered “critical information related to patterns of killing and burial that are often not visible in interviewee testimonies.”\textsuperscript{38} For instance, GIS technology allowed TJWG to view and analyze site coordinates and their relationships to spatial variables such as site elevation and proximity to roads, infrastructure, natural landscape, and, most importantly, political prison camps, correctional labor camps, and disciplinary labor centers.\textsuperscript{39} In 2016, the American Association for the Advancement of Science (AAAS) reported that the most widely used geospatial technology for human rights documentation is high-resolution satellite imagery, followed by medium and low-
resolution imagery, conventional GIS techniques such as cartography and spatial statistics, and thermal and radar imagery.\textsuperscript{40}

Despite NGOs’ laudable efforts to document human rights violations in North Korea with satellite imagery and GIS technology, there are limitations to commercial satellite imagery analysis, which TJWG recognized. For instance, the mapping of geographical coordinates and subsequent geospatial analysis of these locations in the TJWG report were used to substantiate personal accounts provided by a sample population of the escapee community in South Korea.\textsuperscript{41} Because the majority of its participants in the project were originally from the northernmost provinces like North Hamgyong and Ryanggang provinces, it is likely that North Hamgyong province’s dense geographic distribution of body disposal sites, as illustrated in the report, could be a result of the narrow sampling of interviewees.\textsuperscript{42} Furthermore, it is difficult to assess via satellite imagery whether the regime had already tampered with the sites before the images were taken. For similar limitations, satellite imagery and GIS-based evidence are typically only used “in support of conventional evidence such as witness testimonies,” although they were admitted in the 1995 International Criminal Tribunal for the former Yugoslavia for the Srebrenica massacre.\textsuperscript{43} The Bosnian Serb Army had dug up remains from mass gravesites that were formed after the Srebrenica massacre and had reburied them in different locations after the U.S. Secretary of State publicly alluded to the U.S. awareness of those sites.\textsuperscript{44} However, international courts today still lack specific regulations dedicated to the use of geospatial evidence in court.\textsuperscript{45}

Despite multiple challenges—interviewee sampling challenges, resource constraints, and the institutional and legal environment—TJWG began the mapping process hoping that their work today will accelerate the speed of human rights investigations in North Korea, return remains to families, and someday hold the regime accountable for the human rights abuses. Apart from the killing and burial sites, TJWG also mapped locations of Hoeryong City’s local Ministry of People’s Security offices, Ministry of State Security offices, military units, and administrative offices as they are regarded as some of the “location[s] of possible repositories of documents which may contain evidence of human rights abuses.”\textsuperscript{46} Just as how the well-documented Stasi records after German unification exposed dark truths about East Germany’s repression of its people, TJWG hopes that mapping these potential repositories via satellite imagery will be used for future legal proceedings to hold the DPRK authorities accountable.\textsuperscript{47}

With the growing number of satellite imagery analysts working at think tanks in the past decade, the NGA started a new program called the Tearline Project to “extend access to commercial and unclassified geospatial data” to partnered non-profit organizations.\textsuperscript{48} Such an initiative allows public reporting on and analysis of “various strategic, economic, and humanitarian intelligence topics that tend to be under-reported” but are critical to U.S. national security.\textsuperscript{49} As one of its analytic outreach efforts, the NGA is collaborating with the Stimson Center’s 38 North program to offer geospatial intelligence imagery of North Korea taken by different commercial satellite companies including Planet Labs, Airbus Defense & Space, and Maxar Technologies. Some of the recent 38 North publications include reports of the regime’s tideland reclamation efforts to develop polders into arable farmland—in the hopes of combating food shortages—and the
DPRK’s mining activities in the Komdok region, which is reportedly full of mineral resources that could be a source of economic revenue for Pyongyang.⁵⁰

The Beyond Parallel program at the Center for Strategic International Studies is another source of satellite imagery analysis of North Korea’s military and domestic activities. In February 2022, Beyond Parallel published a report that the DPRK’s five-megawatt electrical (5MWe) reactor at the Yongbyon Nuclear Research Center was operational based on a thermal infrared analysis drawn upon by Maxar’s satellite imagery and National Aeronautics and Space Administration LANDSAT 7 and LANDSAT 8 satellites’ thermal infrared imagery.⁵¹ Whereas the operational status of the IRT-DPRK reactor, the Radiochemistry Laboratory, and the centrifuge plant were not confirmed definitively from the satellite imagery, the 5Mw reactor’s ice-free cooling discharge pipe and “a semi-circle of thawed water in an otherwise frozen river” were sufficient to prove that North Korea is still in pursuit of fissile material for its nuclear weapons program.⁵² Even as the 5Mwe reactor—shut down in December 2018 as a result of the second inter-Korean summit in which Kim had agreed to dismantle nuclear facilities at Yongbyon—is the only reactor in the DPRK that reprocesses spent fuel rods for plutonium production, its operational status indicates that Kim “sees a need for additional nuclear material for an expanding nuclear arsenal.”⁵³

When the International Atomic Energy Agency (IAEA) published a report in August 2021 suspecting the resumption of North Korea’s 5MWe operations, both the Beyond Parallel and 38 North programs released satellite imagery and analyses to the public in support of the IAEA’s assessment.⁵⁴

On North Korea’s domestic activities, Beyond Parallel published an assessment of the regime’s food production and crop yields—critical factors of the regime’s domestic stability—by using advanced geospatial and remote sensing technology. By using the ESA’s PROBA-V sensor data and applying it to the Normalized Difference Vegetation Index (NDVI), which indicates the density of green vegetation in each pixel, the report was able to observe North Korea’s “crop growth vigor and the quality and health of crops” from early July to late August of 2021.⁵⁵ At a time when concerns about a potential humanitarian crisis stemming from food insecurity were at their highest due to the COVID-19 pandemic and subsequent border closures, this analysis validated that certain regions were in fact suffering from worsened agricultural conditions. Despite the significant benefits of commercial imagery, however, some drawbacks have been apparent.

**Limitations of Global Transparency & Dangers of “Democratized Intelligence”**

Commercial U.S. space companies today are vital and integral members of the space domain, supporting various missions of the U.S. government and contributing to the rise of global transparency. One of the earliest claims in favor of commercial satellite imagery at the time of the Ikonos launch in 1999 was that it could “provide an independent check on what the government is saying, for example about mass graves and other wartime atrocities in the Balkans.”⁵⁶ Beyond providing secondary verification, commercial satellite imagery and geospatial technology have been proven useful in “extracting information from inaccessible areas” or otherwise closed societies beyond North Korea like in Iraq, Eritrea, the Democratic Republic of the...
Congo (DRC), Chad, and Burma. However, various risks and dangers associated with global transparency, such as lost privacy, have also been raised since the inception of this industry. As commercial satellite technology continues moving forward, we must aim to understand the limitations and dangers of global transparency.

The first limitation to complete, global transparency is the “uneven, selective, and politicized transparency” stemming from the commercial sector’s financial dependence on the U.S. government, which is accentuated by the former’s activity in regions that receive the U.S. military’s attention for national security concerns like North Korea. As Dr. John Klein states, “commercial space” is difficult to define, and the definitions vary depending on each organization’s perspective. In the twenty-first century, improved public and private sector cooperation and the integrated nature of government and commercial assets in the hybrid space domain are making the delineation even harder. The U.S. National Space Policy, published in December 2020, states that “commercial” in the context of space “refers to goods, services, or activities provided by private sector enterprises that bear a reasonable portion of the investment risk and responsibility for the activity, operate in accordance with typical market-based incentives for controlling cost and optimizing return on investment, and have the legal capacity to offer those goods or services to existing or potential non-governmental customers.” Per this definition, commercial space companies that have a government agency as their “anchor customer,” or customer that generates most of their revenue, should be regarded as government contractors instead. For instance, Maxar DigitalGlobe, which is considered a commercial vendor, has the NGA as its anchor customer with the U.S. government agencies comprising over 60 percent of its revenue, based on its financial reports.

To test whether commercial satellite imagery is “a neutral instrument of security governance,” as many NGOs, academics, and policymakers believe, Phillip Olbrich created a heatmap of high-resolution satellite imagery taken of North Korea by DigitalGlobe over a 15-year period from 2002 to 2017. Olbrich’s geospatial distribution of photographed satellite images refutes the claim that commercial satellite imagery offers “uniform satellite-based transparency” or equal coverage of the world below. Instead, Olbrich’s heatmap suggests that there are specific regions that high-resolution satellites were “tasked” or programmed to focus on by DigitalGlobe, most likely due to time and resource constraints. Some of the regions on the heatmap that had the highest concentration of high-resolution images were areas that had military establishments, such as the Yongbyon Nuclear Complex, Punggye-ri Nuclear Test Site Tunnel 5, Tonghae Satellite Launching Ground, and the Sohae satellite launching station in North Korea. As these locations are of high interest to the U.S. Department of Defense (DOD), such a finding indicates that not only is DigitalGlobe reliant on the U.S. government for its revenue but for its mission, too. Thus, this heatmap concludes that commercial satellite imagery available for public consumption is not free from geopolitical interests as is widely believed. The more worrisome consequence is that users of satellite imagery who are oblivious to this reality may use biased datasets to generate new findings, resulting in unforeseen second-order and third-order effects. With the growth of artificial intelligence software and machine learning programs today, for
instance, an individual or group with malintent could intentionally use biased datasets to draw conclusions that will support their dangerous claims. Thus, users of satellite imagery should be mindful that the images may have been captured and processed for political reasons, especially as the NRO and NGA seek to increase public-private cooperation in the future.67

The commercial satellite industry’s trend toward expanding government contract shares, especially in defense and intelligence, is also hindering the growth and neutrality of global transparency. At the time of the Ikonos launch in 1999, Space Imaging expected that its main customers in the short run would primarily be domestic and foreign governments and militaries, like the Pentagon.68 However, the company also anticipated that its non-government customer share would quickly increase once the public saw and understood the power of commercial satellite imagery.69 Until a few years ago, Maxar Technologies had still hoped to “split its sales one third each between commercial, civil, and national security space” and has won NASA contracts in recent years to diversify its business portfolio.70 Although its plans for commercial and civil sectors are intact, the company has since pivoted its strategy to become a more competitive U.S. defense industry contractor and subcontractor, and in producing smaller, low Earth orbit (LEO) constellations to offset a downturn in its commercial geostationary satellite sales. As a result, Maxar won a $3.2 billion ten-year contract from the NRO in August 2022 and a $192 million five-year contract from the NGA for its satellite imagery.71 The corporate leadership’s new assessment that the defense and intelligence industries are more stable in comparison to the commercial communication satellites market suggests that fewer tasking operations in the future will be dedicated to non-military-related regions or targets.

As a result of increased military-centric commercial satellite imagery, NGOs that are currently making important contributions to the public dialogue surrounding the domestic conditions of North Korea and depend on donations of images from commercial vendors will have difficulty analyzing non-military targets. Although satellite images could be made available for public use, they may not be captured sufficiently in either quality or quantity absent companies’ specific tasking requirements for its satellites. In the future, for instance, it would be difficult to track whether North Korea is moving or destroying evidence of burial sites, and consequently fewer images taken of these locations would reduce the chances of satellite imagery ever being accepted in court as evidence of the DPRK’s human rights violations. Today, Washington’s prioritization of military-related satellite imagery collection is expected, especially concerning North Korea, yet such a trend could also suggest Washington’s loss of opportunity to observe softer variables—indicators of the domestic political, economic, and social environment—that may help Washington reach a breakthrough in U.S.-DPRK relations, or even toward denuclearization.

Another danger of global transparency is the misreading and misinterpretation of satellite imagery, whether intentional or not, and how inaccurate reporting thereof could jeopardize diplomatic relations. One of the earliest mistakes of satellite imagery analysis was when a newspaper called European in 1992, not only published a satellite image “upside down and backwards,” but claimed that a military airbase in Algeria was a nuclear research
In 1998, *Newsweek* magazine published satellite imagery of a site in Rajasthan, India citing that the image was taken a week before India’s five nuclear tests. However, the used image was in fact taken over five years ago, and the hole that *Newsweek* identified as one of the explosion sites was later revealed to be an animal holding pen. As commercial satellite imagery has become more widely accessible to non-government customers, the chances of misinterpreting satellite imagery have also increased. Admittedly, interpretation of satellite imagery is challenging for trained experts but is much more challenging for non-experts who lack technical expertise, experience, and an in-depth understanding of the target country. However, the two examples cited above by news outlets in “an attempt to be the first with breaking news” could “easily embitter relations among nations and prevent the resolution of outstanding disputes.”

In dealing with countries like North Korea that will capitalize on its adversaries’ mistakes to justify prohibited activities like banned weapons testing as self-defense, the costs of misinterpreting and misreporting DPRK installations—especially military installations—are much higher. Given the power of visual images and the risks they bear, the context and facts surrounding them must be verified before public dissemination.

A potential conflict of interest between different governments and NGOs is another relevant concern in regard to the global transparency and, more specifically, “democratized intelligence” that have been made possible with the commercialization of the satellite industry. When *The Hidden Gulag* was published, the U.S. government declined to declassify photos of North Korean gulags because it wished not to “reveal precisely what spy satellites are able to detect” as a means to protect its intelligence sources and methods. Former South Korean President Moon Jae-in criminalized domestic NGOs’ flying of propaganda leaflets to North Korea in 2020 to pursue his political agenda and appease Kim. Therefore, it is possible that satellite imagery reports that call out the regime’s human rights violations could be discouraged from publication by the U.S. and South Korean governments to pursue diplomacy with North Korea.

Moreover, the ease of access to high-resolution satellite imagery today could be easily exploited by malign actors such as North Korea that wish to track adversary military movements but lack sufficient reconnaissance satellite technology. In January 2021, Kim Jong-un publicly announced the regime’s ambitions to build new military reconnaissance satellites as one of the “five major goals for the development of national defense capabilities.” Although North Korea successfully launched self-proclaimed weather satellites into orbit in December 2012 and February 2016, experts argue that the two satellites failed to transmit imagery back to North Korea and “the new device displayed in state media appeared too small and crudely designed to support high-resolution imagery.”

Although satellite imaging and data transmission capabilities were improved by December 2022, the latest reconnaissance satellite launch failure on May 31, 2023, reportedly due to second-stage engine failure, indicates North Korea’s lack of appropriate space launch vehicle (SLV) rocket technology to launch military reconnaissance satellites. Even though there is no public information on North Korea’s current use of open-source high-resolution images, their continued launch failures could suggest that the regime will continue to pursue satellite imagery by every

---

*Georgetown Security Studies Review* 105 Volume 11 | Issue 1
means possible until they can successfully launch three to five indigenous spy satellites, which would be sufficient for “a space-based surveillance system that allows it to monitor the Korean Peninsula in near real-time.”  Until then, Google Earth’s satellite coverage of East Asia could be highly resourceful to Kim.

Lastly, totalitarian regimes like North Korea, which oppose widespread accessibility of satellite imagery of their countries, are investing in counterspace weapons to deny infiltration and deceive enemies, thereby creating a more hostile space environment for global transparency. Although the regime has not yet demonstrated kinetic or non-kinetic physical counterspace capabilities, it began jamming operations, a type of electronic warfare (EW) counterspace capability, against U.S. space systems and South Korean forces in 2010.  Analysts suspect that the DPRK military is operating jammers intermittently in peacetime to gain operational experience for future combat.  In addition to North Korea’s use of GPS jamming during the 2010 annual U.S.-Korean military exercise, it has also interfered with civilian GPS navigation and disrupted South Korean air traffic in 2012. Although such jamming operations today are not expected to affect U.S. military GPS signals, commercial satellite communications may become vulnerable to future DPRK downlink and uplink jamming.  Even as military space-based communications and navigation are likely to become targets of DPRK EW and cyber attacks, the U.S. government must devise a response plan for when North Korea improves its counterspace capabilities to target U.S. and allied commercial and/or civilian space assets in a future conflict.

Future Opportunities for Commercial Vendors: Satellite Imagery & Internet Services

Commercial companies operate in an environment where tensions between North Korea and the U.S. do not cease, creating more business opportunities. South Korean Defense Minister Lee Jong-Sup’s visit to the NGA headquarters in Springfield, Virginia on November 2, 2022, following a series of North Korean ballistic missile launches, marked the first official visit of a ROK defense minister to the agency.  The timing of this unprecedented visit evoked a sense of urgency to the accelerating North Korean provocations and was a reminder of the importance of U.S.-ROK intelligence cooperation in the field of satellite imagery to maintain “absolute superiority vis-à-vis North Korea in the field of intelligence, which is a core element of deterrence.”  NGA Director Vice Admiral Frank Whitworth’s statement that “the monitoring of the North [is] a primary mission of the agency,” served as both reassurance to its ally and a warning to its adversary of the U.S. response to continued DPRK missile tests.  Despite the two nations’ reiteration of commitment, North Korea’s vote against the UN General Assembly resolution in October 2022, which called all countries to self-impose a ban on direct-ascent anti-satellite (ASAT) missile tests, suggests that the DPRK regime may develop physical kinetic counterspace capabilities in the future to deny and degrade U.S. surveillance operations targeting North Korea.  Although “North Korea has yet to exhibit the necessary sensing and altitude control capabilities for a direct-ascent ASAT weapon,” and has never publicly announced its plans to develop such capabilities, its improved ballistic missile technology with over 100 missile tests since 2022 may accelerate its development if it decided to
pursue such a path. As scholars note, North Korea, like Iran, “could field a crude direct-ascent ASAT capability in the near-term by adapting a ballistic missile to launch an unguided warhead to detonate in the vicinity of a target satellite.”

Given North Korea’s continued weapons testing and the ongoing Russo-Ukrainian War, U.S. commercial satellite imagery companies are expected to encounter more business opportunities to support the U.S. defense and intelligence industries. To address growing security threats in Asia, the Department of Defense announced in November 2022 that a Space Force unit will be formally established within the U.S. Indo-Pacific Command. More recently in December 2022, the U.S. Space Force launched a component command in U.S. Forces Korea, making the unit the first Space Force command on foreign soil to provide overseas support to a partner nation. Moving forward, it is expected that the Space Force will play a pivotal role in surveilling the Korean peninsula alongside the U.S. intelligence agencies. Thus, despite the aforementioned dangers of commercial satellites focusing on certain regions flagged by the U.S. military, commercial companies will naturally seek more federal contracts with the DOD especially since “the Space Force has the fastest growing budget of any of the service branches.” Furthermore, domestic companies are likely to have a competitive advantage in such a bidding process against non-U.S. entities given the sensitivity of U.S. national security. For instance, one of the reasons why Maxar Technologies has been a competitive bidder for the defense and intelligence contracts since 2020 was that it was able to rebrand itself as a U.S. corporation after having sold one of its Canadian subsidiaries, MDA. However, smaller start-up commercial satellite imagery companies such as Capella Space and HawkEye 360—both capable of tracking enemy troop movements via their own satellites—are less likely to win long-term federal projects against giants like Maxar primarily due to the “Pentagon’s slow-moving, risk-averse military procurement bureaucracy.” However, Pentagon officials have acknowledged this problem and recognized how commercial satellites are more suitable for persistent surveillance given their “more frequent coverage of the world than even the U.S. government can provide.” As public-private cooperation in the satellite industry increases, commercial companies will be able to expand their businesses because satellite imagery and other relevant satellite technology will meet the demands of the Defense Department.

Commercial satellite companies that do not provide imagery will be able to offer communication services to regions or countries that lack the appropriate network infrastructure. The NGA’s satellite imagery, which distinguishes anomalous behavior from a baseline, confirmed the build-up of Russian forces along the Ukrainian border and served as a decisive early indicator of an imminent Russian invasion in February 2022. Despite the NGA’s work, SpaceX’s satellite internet constellation, Starlink, has received greater public attention as a leading space actor in support of Ukraine since the invasion began. With 4,000 of its small satellites in LEO (as of May 2023), Starlink has provided high-speed, low-latency broadband internet services to millions of Ukrainian civilians across and outside the country, and secure communications to Ukrainian forces. It was reported in October 2022 that Starlink’s broadband services helped the Ukrainian military’s drone operations—especially with drone feeds, identification of targets, and correction of fire in real-time—with a Ukrainian
commander commenting that “fighting without Starlink service at the front line is like fighting without a gun.”

However, Starlink has begun taking steps since February 2023 to prevent the Ukrainian government from using Starlink satellite internet service “for offensive purposes,” claiming that Starlink “was never intended to be weaponized.” Despite SpaceX’s restrictions to limit the use of Starlink services by the Ukrainian military, the world has already witnessed the potential and power of a commercial space company’s satellite services in a conflict zone. As military assets in orbit have distinct technical, operational, and environmental requirements that are not going to meet the needs of civilian populations, the market demand for commercial satellite communications will continue to soar.

Beyond satellite imagery, commercial companies will be more active in the wars and conflicts of the future. Starlink’s internet services offered to the Iranian people, with the U.S. government's approval and support, indicate that commercial satellite companies can target regions and countries that lack internet access. In September and October 2022, the Iranian government cut off internet access in certain parts of the country to keep its people in the dark about the expanding protests across the country. Once the Biden administration granted technology companies, including SpaceX, a general license to offer “services that support the communication tools to assist ordinary Iranians in resisting repressive internet censorship and surveillance tools deployed by the Iranian regime,” Starlink satellite receivers were smuggled inside Iran. Reports suggest that approximately 1,000 receivers have been activated in Iran as of January 2023. Despite activists’ efforts to remain informed about the protests against government censorship, there are arguments that Starlink satellites “are not ready for countries with hostile governments.” While it would be remarkable if SpaceX developed “more practical and safer” means to offer information access to those living inside repressive regimes—like Iran and North Korea—a commercial company is not obligated to support a foreign grassroots campaign or coup d'état to fulfill its host country’s foreign policy goals, without a government contract. Corporate investments follow market demand, and corporate leadership is unlikely to invest in costly projects that are not going to evolve into a sustainable and lucrative business model. If commercial companies lack financial or legal support from the host government, business plans to initiate operations in hostile environments could lose support from their shareholders.

With SpaceX’s success in Ukraine and recent efforts in Iran, the viability of a commercial satellite internet provider offering internet access to North Koreans inside the Hermit Kingdom has also been discussed. Fortunately, civilian access to electronic devices within the country is greater than one would imagine. The 38 North Program assessed in November 2022 that “mobile communications has become an integral and indispensable part of North Korean daily life,” with smartphone ownership potentially ranging between 50 and 80 percent of the adult population and mobile phone distribution reaching one per household. Nevertheless, North Korea still bans the use of unauthorized devices, heavily censors the media, and executes those who access foreign media. It was reported in early December 2022 that two teenage boys were killed by a firing squad in front of locals for watching and distributing South Korean movies, and Kim has publicly executed at least seven North Koreans for
circulating banned South Korean content in the past decade.\(^{107}\) Considering these instances of capital punishment, one can only guess the cost of North Korean citizens getting caught smuggling Starlink receivers across the Chinese border. Nonetheless, ordinary North Koreans accessing uncensored content in the future does not seem to be entirely out of the question. Just as how Japan’s Toshiba developed compact, battery-powered receivers and South Korea’s Samsung Electronics and LG Electronics developed mobile phones with built-in satellite receivers in the mid-2000s to receive broadcasts from “a satellite with a massive 12-meter antenna area,” commercial companies today can end the state-imposed information blockade with similar technology. Although the satellite broadcast services were shortly discontinued due to low profits and the opportunity to target North Korea was missed, Martyn Williams, a Senior Fellow with the Stimson Center, argues that “satellite Internet companies such as Starlink and OneWeb” can still bring back a similar opportunity with the development of “technology that can communicate directly with portable devices.”\(^{108}\)

Although the development of technology may seem to be the biggest barrier to providing internet access to North Koreans, U.S. government policy, including its legal response framework and financial support to commercial companies, presents a bigger challenge. Fortunately, SpaceX and T-Mobile announced a partnership to bring satellite-to-cell service in August 2022 and will begin testing their new service this year to “end mobile dead zones” across the Earth.\(^{109}\) With Amazon’s Kuiper Systems also partnering with Verizon, and Apple’s investment in Globalstar for similar satellite-to-cell underway, an information campaign targeting North Korea may not be too far away.\(^{110}\) However, the DPRK leadership is extremely sensitive to any information campaign targeted toward North Korean civilians since the regime’s domestic stability directly rests upon Kim’s ability to control the media and broadcast state propaganda. Thus, the United States can expect the regime’s severe retaliation against any foreign attempt to provide global internet access inside North Korean territory. Just as how Russia has been launching cyberattacks, signal jamming, and testing new EW systems to disrupt Starlink transmissions in Ukraine, North Korea may also use a range of non-kinetic physical, EW, and cyber counterspace weapons to deny and degrade foreign operations, even if the campaign were to be authorized and initiated.\(^{111}\) Whereas Russia has refrained from using kinetic physical counterspace weapons to disable Starlink satellites so far, a more desperate North Korea may be inclined to use its ballistic missile technology to take more drastic measures if deemed necessary. Therefore, a commercial space company’s novel attempt, without relevant U.S. policies in place, to provide internet access to North Koreans might initiate a conflict, putting innocent civilians inside the country in greater danger and leaving American and ROK armed forces vulnerable to unsolicited attacks.

Taking SpaceX’s activities in Ukraine as an example, there are other two potential conflicts of interest that could arise between commercial satellite internet companies and the U.S. government in a future information campaign targeted toward North Korea. First, SpaceX has been providing all the actual internet service—more costly than the physical satellite terminals themselves—in Ukraine for free.\(^{112}\) Despite Elon Musk’s pledge to support Ukraine with Starlink days after the Russian invasion, the significant cost burden without additional U.S.
government funding left Musk to reconsider future internet services to Ukraine. Although SpaceX did not confirm any linkages between the company’s request to the Defense Department to take over funding for Ukraine’s government and the 1,300 Starlink satellites’ abrupt termination of services in October, such inconsistency reported in the news alone has raised doubts about SpaceX’s commitment to providing internet services for the Ukrainian people. As this instance has shown that “powerful private actors are not reliable allies” on the battlefield, it is unclear whether SpaceX would ever commit to a project to help North Koreans by mass-producing portable devices that can communicate with Starlink satellites. It is even more unclear whether the U.S. government would offer financial support to SpaceX, either overtly or covertly, to cover the expenses of relevant research and development. As with Starlink services in Ukraine, a potential overreliance on a single commercial company’s space assets for an information campaign is also an issue that must be addressed.

Another concern is the lack of the U.S. government’s clear legal response framework to adversaries’ potential kinetic attacks on U.S. commercial satellites. Per Article 52(2) of Additional Protocol I (API) of international law, the Starlink constellation is a valid military target for Russian forces as Starlink satellites’ high-speed internet and communications are used by the Ukrainian forces to “make an effective contribution to [its] military action.” China, which is closely observing Starlink’s operation in Ukraine, has also shown intent for developing “hard-kill and soft-kill capabilities for use against Starlink” if employed in a similar manner in a future Taiwan contingency. If SpaceX must bear all costs and legal responsibilities arising from its operations in North Korea without a guarantee of U.S. government intervention or protection in a kinetic conflict, no commercial company would want to pursue such a high-risk project. In the context of Starlink’s operations in Ukraine, we cannot ignore how its participation in the war has been a significant business opportunity for the company. Using Ukraine as a test bed, SpaceX was able to test out its Starlink technology in mass and scale, receive significant amounts of positive press coverage for months, and, most importantly, solidify its lead in the LEO satellite communications industry through its participation in the war. Without perceived business advantages and a legal strategy, commercial companies will refrain from formulating business plans to operate in hostile environments like North Korea.

In addition to these concerns, two U.S. geopolitical adversaries who support North Korea—China and Russia—are in the region. As China and Russia have both successfully conducted destructive ASAT tests in recent years, it is possible that Beijing or Moscow could transfer counterspace weapon technology to Pyongyang to further their own anti-access/area denial (A2/AD) strategies against the United States. Therefore, U.S. public-private partnerships to pursue an information-based strategy via satellites to undermine North Korea’s military leadership and counter their propaganda efforts should not be encouraged until some of these problems could be discussed and addressed at length. Whereas formulation of unilateral and multilateral responses to kinetic attacks in a hybrid space environment will continue to be a challenge, the issues of funding, research, and technological development may be resolved through cooperation with Japanese and South Korean commercial space companies. With Japan’s traditionally robust
commercial space industry and South Korea’s increased government investment in domestic space capabilities in recent years, a U.S.-Japan-ROK trilateral consortium could be created if the United States ever decided to pursue an information-based strategy against North Korea. In addition to historically greater government interference in business activities in Japan and South Korea, U.S. regional allies and their commercial sectors may have more vested interests than U.S. commercial companies in developing relevant satellite technology given their geographic proximity to the North Korean threat.

Conclusion

This paper examined the rise and growth of the commercial satellite imagery industry from the 1990s to the present day and assessed future opportunities for the U.S. commercial satellite industry—beyond high-resolution imagery—in the twenty-first century. Given the role that imagery intelligence has traditionally played in gathering information on the Korean peninsula and continued DPRK military provocations, many examples with regard to North Korea have been used throughout the paper to support its arguments such as the U.S. government’s monopoly on overhead imagery, dangers arising from adversary or non-government actors’ use of commercial satellite imagery, and commercial companies’ future satellite-to-cell business opportunities. Given increased global interest and activity in the space domain, the international community with the proper collaboration with commercial companies can reach a breakthrough in eliminating the North Korean threat.

About the Author: Sue Kim received an M.A. in Security Studies with a concentration in Technology and Security from Georgetown University in 2023. She also received a B.S. in Foreign Service from the Walsh School of Foreign Service in 2022, majoring in International Politics and earning the Joseph S. Sebes, S. J. Medal in Asian Studies.


4 Intelligence Revolution 1960, 65.


12 Richard Mobley, Flash Point North Korea (Annapolis: Naval Institute Press, 2003), 53.


15 Central Intelligence Agency, Black Shield Reconnaissance Missions, 8.


17 Mobley, Flash Point North Korea, 55.


19 Roblin, “Forget the SR-71 Blackbird, the AR-12 Oxcart Was Way Faster.”


22 Mobley, Flash Point North Korea, 53.

23 Roblin, “Forget the SR-71 Blackbird, the AR-12 Oxcart Was Way Faster.”


29 Broad, “Private Spy in Space to Rival Military’s.”
31 David Hawk, *The Hidden GULAG*, 88; McKinzie and Cochran, “Conducting Research on North Korea using Commercial Satellite Imagery, Map and Other Geographic Data,” 2.
38 Mapping the Fate of the Dead: Killings and Burials in North Korea, 14.
39 Mapping the Fate of the Dead: Killings and Burials in North Korea, 45
41 Mapping the Fate of the Dead: Killings and Burials in North Korea, 21.
42 Mapping the Fate of the Dead: Killings and Burials in North Korea, 23.
43 Mapping the Fate of the Dead: Killings and Burials in North Korea, 62.
45 Mapping the Fate of the Dead: Killings and Burials in North Korea, 62.
46 Mapping the Fate of the Dead: Killings and Burials in North Korea, 56.
47 Mapping the Fate of the Dead: Killings and Burials in North Korea, 56.
49 “About the Tearline Project.”
52 Bermudez Jr., Cha, Dinville and Jun, “Thermal Imagery Analysis of Continued Activity at Yongbyon.”
56 Broad, “Private Spy in Space to Rival Military’s.”
57 Broadly Accepted Practices Regarding the Use of Geospatial Technologies for Human Rights, 1.
58 Broad, “Private Spy in Space to Rival Military’s.”
63 Olbrich, “Transparency from Space,” 34.
64 Olbrich, “Transparency from Space,” 35.
65 Olbrich, “Transparency from Space,” 34.
66 Olbrich, “Transparency from Space,” 34.
72 Brooke, “Rights Group Exposes Conditions in North Korean Prison Camps.”
87 “South Korea's defense chief visits U.S. geospatial intel agency amid North Korean threats.”
93 Sandra Erwin, “Maxar’s satellite business looks to gain foothold in defense market.”
94 Sandra Erwin, “Maxar’s satellite business looks to gain foothold in defense market.”
96 Lipton, “Start-Ups Bring Silicon Valley Ethos to a Lumbering Military-Industrial Complex.”
104 Vick, “Receivers for Elon Musk’s Starlink Internet Are Being Smuggled Into Iran.”
105 Vick, “Receivers for Elon Musk’s Starlink Internet Are Being Smuggled Into Iran.”


Tara Brown, “Can Starlink Satellites be Lawfully Targeted?”
International Legal Accountability and the World Order: Book Review of Geoffrey Robertson’s *Crimes Against Humanity: The Struggle for Global Justice*

Mariam Kvaratskhelia

Geoffrey Robertson’s book, originally published in 1999 and updated in 2007, is an in-depth survey of the history and present state of the world legal order. Through telling the story of human rights from the American and French Revolutions to the Universal Declaration of Human Rights (UDHR) and up to the modern era, the author offers a brilliant insight into the evolution of human rights law as well as international criminal law. According to him, this evolution is not linear: Robertson demonstrates that the search for international morality and rule of law was and still is a process full of impediments. The author labels the Nuremberg and Tokyo trials as historic, noting that they marked the beginning of a new era in global legal thinking by a) recognizing individual’s rights vis-à-vis the state, and b) recognizing state’s obligations vis-à-vis international law. The judgment of Nuremberg, which “created an international criminal law to punish the perpetrators of crimes against humanity,”¹ provided the base of many clauses of UDHR and was also the driving force behind the establishment of rules of war. Robertson defines international criminal law as an independent and universal jurisdiction aimed to prosecute those that commit crimes “…so heinous that it is ‘against humanity.’”² Such crimes include genocide, mass murder, systematic torture as well as warfare and terror.

The principal storyline of the book revolves around how Nuremberg altered the understanding of impunity and the repercussions that should follow state-sponsored murders. If before, the discussion used to be about how democratic and autocratic states respond differently to international law, then today the question is about how any sovereign state – no matter liberal or rogue – adheres to the world legal order. The author considers the revival of the legacy of Nuremberg through two unprecedented cases. He looks at the history of war from the Hague and the Geneva Conventions to the Rome Statute of 1998 and the creation of International Criminal Court (ICC). Robertson labels the Rome Statute as a major achievement of the late 20th century because the document asserts that it is a *moral imperative* of the adherents of the law to end impunity and exercise criminal jurisdiction over perpetrators.³ In addition to this, the author tells the stories of the arrest of General Pinochet in 1998, the war crimes in Yugoslavia and Rwanda, and the Lockerbie agreement. As Robertson outlines in the prologue, he wants to build an argument for a millennial shift from appeasement to justice as the principal driving force of the current world order.

In a way, Robertson’s book serves as a guideline. His purpose seems to be to show the readers how important it is to learn lessons from the past in order to avoid future tragedies. What drew my particular attention in the book is Robertson’s overt criticism of the modern-day liberal institutions, which, according to him, give a lot of degree of leeway to global aspiring hegemons. Most crimes against humanity, the author notes, are committed by “professional soldiers, blessed by religious leaders and tacitly approved by governments.”⁴ Robertson’s assessment of the inefficacy of current international organizations, which stems from their deeply flawed delivery and
enforcement mechanisms, is astute. As history has demonstrated, “the endemic failure [of the UN] to allow for criticisms of its own members” has been a major impediment to effective decision-making. Often, judgement of these institutions is wildly biased and diluted by the interests of the superpowers. Robertson brilliantly manages to unveil the unpleasant image of these international legal organs.

The impulse to find international morality predates modern-day liberal institutions. The League of Nations, which, in the author’s words, was too conservative and “diplomatic,” sought to bring people to the same table and was one of the first attempts at establishing a global platform for states to discuss international morality. After the creation of the UN, it became clear that “a more permanent international justice system” was needed. After the Nuremberg and Tokyo trials, the UN made a reference to potential creation of an “international penal tribunal,” though the project never came to completion due to the Cold War. Australia’s Prime Minister, Dr. H.V. Evatt, proposed the European Court of Human Rights, asserting that even democratic governments could and should not be trusted with the protection of the rights of individual citizens. It became clear that finding an overarching system of justice that possessed power higher than the sovereignty of each state was gaining more momentum, and that the search for a “world constitution” was becoming more prominent on the contemporary political agenda.

Robertson criticizes the internal bureaucracy of the UN and the organs codified by the UN Charter, such as the Human Rights Commission (HRC). According to him, the HRC turned a blind eye to some of the major human rights violations of the time, such as: The CIA’s provocations in Chile and the U.S. government’s sponsorship of the military coup, the mass rape of up to 300,000 Bangladeshi women throughout Pakistan’s invasion, and others. Robertson underscores that if the HRC is to ever become credible, it needs to address major structural problems. For example, it would have to get rid of the so-called experts who, in fact, are the “mouthpieces” of certain governments. The HRC instrumentality is flawed exactly because in lieu of independent experts, the fifty-three members were representatives of governments. This allegiance is why they were committed to neutrality instead of taking clear stances — an issue Robertson deems deeply problematic. The HRC would also have to meet more often throughout the year and manage to cut ties with the UN Secretariat to gain some independence (as the budget and structure of HRC are dictated by the Secretariat). The HRC cannot pressure states to perform their duties due to the fact that there are no real legal obligations. This lack of substance is why the behavior of the superpowers of the time was often veiled “in the language of legality” during the Cold War. For example, the U.S. justified its invasion of the Dominican Republic in 1965 by asserting that it was serving the regional democratic rule, whereas the “Brezhnev Doctrine” of 1968 was formulated to frame the Soviet invasion of Czechoslovakia as the former giving “fraternal military assistance” to the latter. The absence of a binding force allowed the powerful states to “bend the law” in a way that suited their national agendas and interests.

The utter infringement of the UDHR by the states that were supposedly most in support of its creation is an indictment of the nature of the document. The UDHR was never legally binding; it only possessed the power of a declaration, or “principles without powers of implementation.” It is
interesting to note that it was the totalitarian and autocratic states, such as the Soviet Union and its puppets, that opposed the idea of making the UDHR a binding document with the power of enforceability, whereas democracies such as Britain and Australia were in favor of it. Robertson highlights that the major drawback of the declaration was its “coy phraseology [which] conceals the awkward fact that this proclamation lacks legal force.” Therefore, the big question is whether or not the UDHR can be recognized as having the same force as international law. International law comprises official, legally binding treaties, which should show “high level of compliance” in order to maintain credibility. The source of international law, besides treaties, is ratifications passed by governments. While international law requires consent of states, the UDHR compliance is optional, which turned its virtues into vices. This difference in consent does not mean, however, that international law always upholds the cause of human rights: some doctrines, such as diplomatic immunity and non-compulsory submission to the ICJ, continue to damage the principles laid out by UDHR. In addition, politics also get in the way of the high principles of judicial independence. Some examples given by Robertson include Russia, where two constitutional court judges were pressured to resign after making critical comments about the government; and Gambia, where three judges were dismissed by the President because of their decision to take up “politically sensitive cases.” Customary international law, Robertson notes, rests on state practice on the one hand and opinion juris on the other. It is exactly this “state practice” component that has been the most problematic source of international law: drawing legal measures from “practice” that serves the interests of the state will certainly not produce a set of laws that are fair.

Additionally, a point that I found to be extremely important is how international Westphalian law does not apply to transnational corporations and other non-state actors, some of whose “…global activities generate more product and greater influence than many UN member states will ever possess.” Yet another drawback of international customary law is that it sees war as a legitimate tool for enforcing a country’s national security policy. Robertson heavily criticizes this aspect as well, noting that in the future, war should be deemed as “crime of aggression” without the approval of Security Council or international law.

However, how can the decisions of the Security Council (UNSC) be trusted, since it also, in and of itself, is a largely politicized body? The “Big Five” of the UNSC that Robertson talks about are the ones in charge of the actual UN decision-making. We have seen what happens when a particular human rights violation issue does not lie on the radar of the five permanent member states. An example of this is the Rwandan genocide, which the UNSC not only disregarded but also became partially responsible for as its troops perpetuated the violence.

The traditional U.S. view of international law, Robertson rightly argues, is that it stands above it; therefore, throughout history we encounter cases where the United States or other superpowers abstain from ratifying certain treaties “until the Court was operating to [their] satisfaction.” According to the author, considering all this, the Rome Statute is a major achievement of the human rights movements, as it best deals with the “realpolitik of state power.”

Robertson’s arguments regarding contemporary liberal institutions potentially
undermines the possibility of peace. At times, the author seems too pessimistic about the role of these organizations, to the point where he devalues the purpose for which they were created. Yes, the institutions and their structural frameworks are flawed, but this does not necessarily undermine their merit in the grand scheme of things. It is better to live in a world where these organizations exist, however flawed, than in a world where they do not.

Robertson also seems to be idealistic at times: even if the judicial order lives up to his standards, the Nuremberg dilemma stays unresolved, thus leaving us with the same question of who to prosecute and who to release. All in all, Crimes Against Humanity: The Struggle for Global Justice is a must have for scholars wishing to learn about the possibilities – and limits – of legal accountability at the global level. Its incredible depth and breadth provide decades of historical analysis, and urges the audience to contemplate the trajectory of the search for international morality.

About the Author: Mariam Kvaratskhelia is a graduate student in Georgetown University’s Master of Science in Foreign Service program, with a concentration in Science, Technology and International Affairs. Her academic interests revolve around nuclear security, foreign policy and the great power competition. Mariam’s most recent research focuses on Chinese military posture at both conventional and nuclear levels.

Email: mk2027@georgetown.edu

Note: page numbers refer to those from the online edition.

2 Robertson, 12.
3 Robertson, 580.
4 Robertson, 894.
5 Robertson, 125.
6 Robertson, 77.
7 Robertson, 84.
8 Robertson, 576.
9 Robertson, 125.
10 Robertson, 105.
11 Ibid.
12 Robertson, 86.
13 Robertson, 93.
14 Robertson, 164.
15 Robertson, 213.
16 Robertson, 228.
17 Robertson, 302.
19 Robertson, 576.
20 Robertson, 578.
Soviet Nationality Policy and the Forgotten Periphery: Book Review of Kate Brown’s *A Biography of No Place: From Ethnic Borderland to Soviet Heartland*

Mariam Kvaratskhelia

*A Biography of No Place: From Ethnic Borderland to Soviet Heartland* (Harvard University Press, 2005) tells the story of a dynamic borderland region between historic Poland and Russia called *Kresy.* By employing a distinct approach to the study of time, space, and its people, Kate Brown argues that in the process of building a modern state over the course of three decades, the Soviet Union destroyed the borderlands and transformed its hybrid, multi-faceted, and nuanced culture into a homogeneous and ethnically pure nation-state. The purpose with which Brown takes on this project has particular merit: she wrote this book to bring attention to a region that was central to the formation of nationality policy in the Soviet Union, yet was always treated as a periphery, labeled as backward for its “informality” and locality. Brown’s prose almost resembles that of literary narration; it serves as her platform to raise sympathy for places and peoples that were never considered important enough to become parts of history, and whose stories cannot be excavated via traditional archival research. The author’s ethnographic approach, however, can act as both virtue and vice: while the unique narration makes her book stand out, her ultimate criticism of the nation-state loses some legitimacy and appears overly skeptical due to the very pathos with which she writes. Despite this drawback, the book presents an insightful survey of a region whose history was otherwise lost, and tells a compelling story of the movement of peoples and identities across time and space.

In order to support her research, Brown consults a myriad of non-traditional sources, such as oral histories, material culture (photographs), and archival records, ranging from central-party sources in Moscow to village documents gathered in Zhytomyr. Brown travelled around the countryside, read local newspapers, and conducted interviews with specific individuals, which is why she notes that in the process of writing this book, she “became ethnographer-journalist [herself].” It is useful to separate her methodology from her rhetoric. At the beginning of the book, Brown puts forward a disclaimer, mentioning that there are rewards for only those readers that agree to “take this journey through oral testimony, rumor and unverifiable occurrences”. Therefore, it seems that readers should engage in a type of contract with the author to fully come to appreciate this book for what its mission is. The strength of Brown’s work lies in two aspects: she tells a story of a region that was left neglected and overlooked in history, while telling it with a stylistic approach to writing that is also, in and of itself, neglected and overlooked in historical storytelling. Hence, not only does Brown give a voice to people that never had it, but she also gives attention to unusual types of sources that are deemed useless in historiography.

Brown’s argument is based on a “bottom-up” trajectory, which spotlights it in the discussion about nation-states. By starting at the micro level and then zooming out, she successfully demonstrates that the idea of a “strong Soviet state” against “weak
victimized people” did not stand the test of the culture of Kresy. Instead, we are told the story of people who dictated their own lives, who silently broke the laws of the state in attempts to save their long-lasting traditions, and who spoke a different language than the government, literally and metaphorically. In her detailed analysis of the diverse local customs of Kresy and the state’s mission to homogenize them, Brown shows that deportation of national minorities in the borderlands served the grand Soviet goal of creating a “…distilled nation-space for modern governance.” While her argument is well-supported by her evidence, to some extent Brown disregards the other side of the coin – the point of view of the state. Therefore, her ethnographic “bottom-up” approach turns to vice here: she talks about taxonomy systems and standardization techniques employed by the Soviet government; however, she fails to address why this was the modus operandi of the state to begin with. The building of nation-states, and statecraft in general, inherently entails establishing some sort of cohesion and a sense of formality. This institutionalization is how governments are able to rule their subjects, because, in Brown’s own words, “to name is to control.”

Brown is a character of the story that she tells. In a way, she appears as the representative of people that never had the voice to represent themselves. While the author’s argument about nation-states can leave the reader with clarifying questions, her book brings a unique contribution for three principal reasons. First, her discussion about identity and nationality urges us to think about how physical borders and official state documents affect our sense of individuality as well as belonging. Brown often uses words such as “Germanness” and “Polishness,” which is an interesting way to emphasize that identity in Kresy was a spectrum, and that it was exactly this absence of rigid categories that prompted the Soviet state to ascribe labels of particular nationality to locals. Secondly, Brown’s discussion on the supposed backwardness of Kresy sheds light on the standards of “progress” adopted by the Soviet state, and how these very standards dictated the borderlands’ path towards destruction. Third, the depiction of a state afraid of its unruly subjects is a powerful move on the part of the author – almost an attempt to gift long-lost agency back to the people of Kresy. This representation allows us to see that the centralized state apparatus ultimately rests on power that is “atomized” which is one of the ways to go beyond the thought paradigm of a consolidated nation-space.

Brown’s project is successful because it first and foremost showcases to other scholars the incredible merit that consulting nonconventional sources has. It provides a unique perspective on the Soviet nationality policy and augments the power of those that were continually deemed powerless vis-à-vis the state.

About the Author: Mariam Kvaratskhelia is a graduate student in Georgetown University’s Master of Science in Foreign Service program, with a concentration in Science, Technology and International Affairs. Her academic interests revolve around nuclear security, foreign policy and the great power competition. Mariam’s most recent research focuses on Chinese military posture on both conventional and nuclear levels. Email: mk2027@georgetown.edu.

2 Brown, Kate. *A Biography of No Place*, p. 16.


5 Brown, Kate. *A Biography of No Place*, p. 127.

Alt-ternative Perspective: Can Gang Classification Mitigate the Threat of Violent Extremism?

Michael P. Losacco

The United States is conflicted in its approach to tackling domestic violent extremism. Federal laws and regulations lack a dedicated statute to charge acts of domestic terrorism, while more specific laws for enforcing transnational terrorism exist. This gap has created debate on whether Congress should pass a domestic terrorism statute. Those who favor a new law argue it will attach a label to acts of domestic extremism, signaling the significance of the threat when based on American political, social, and religious discourse. Meanwhile, those against a new law worry about the potential for government agencies to abuse a new power, arguing that existing laws on the books sufficiently mitigate the threat. This paper seeks to add commentary to the latter’s argument and examine if criminal gang classification can reduce the threat of domestic extremism. Research shows that domestic extremist groups and criminal street gangs often exhibit the same characteristics. Gang classification could provide federal authorities with established tools that lead to proactive enforcement and threat reduction without expanding government powers in a manner that violates civil liberties. Upon using 18 USC § 521’s definition of a criminal street gang and comparing it to the activities of both the Proud Boys and Oath Keepers, findings indicate that criminal street gang classification of domestic extremist groups could lead to proactive enforcement measures and mitigation strategies. However, if federal authorities are too broad in their street gang classification, the risk of civil rights violations and “net widening” remains significant.

Introduction

The Department of Justice (DOJ) charged certain members of the Proud Boys and Oath Keepers with seditious conspiracy for their conduct during the January 6th Capitol Riot. But was this the best approach to reduce the threat they pose? Another option may have been to classify these extremist groups as criminal street gangs. Gang classification could provide federal authorities with established tools that lead to proactive enforcement and threat reduction without expanding government powers in a manner that violates civil liberties. These tools include using existing gang database infrastructure, Civil Gang Injunctions (CGI), and the Racketeering and Influenced Corrupt Organization Act (RICO).

This paper assesses whether the Proud Boys and Oath Keepers could qualify as a criminal street gang under 18 USC § 521 and, if so, what prosecutorial options this classification makes available. Under this general framework, the paper is organized into four parts. First, in the background discussion, I highlight how federal authorities have traditionally approached domestic extremist groups and where these efforts fall short. Second, I discuss current gang classification research and highlight its practical implications. Third, using the criminal street gang definition in 18 USC § 521, I analyze the characteristics of the Proud Boys and Oath Keepers and determine if these groups could qualify as criminal street gangs. Fourth, I examine the tools gang classification of domestic extremist groups could provide to federal authorities and outline the potential issues that could arise with this approach. In closing, I summarize the key points and offer recommendations for future research.
Background

Historically, the United States has been conflicted in its approach to mitigating the threat of domestic terrorism. Central to this debate is whether the United States needs a statute that criminalizes acts of domestic terror and provides new tools for preventative action. While federal law currently defines acts of domestic terror, it does not include a chargeable criminal offense for domestic terrorism.

In comparison, international terrorism laws under 18 USC Chapter 113B allow the federal government to designate groups as foreign terrorist organizations (FTO) and embargo material support provided knowingly in support of FTOs. For example, an individual pledging themselves servant to a designated FTO and providing resources in support of the organization could be subject to criminal prosecution under 18 USC § 2339B, or other statutory sections in chapter 113B. The scope of Chapter 113B allows prosecutors to proactively mitigate the threat of international terrorism by charging dangerous individuals that are not linked to specific terror attacks and conspiracies but are knowingly and materially involved in the general support of the foreign terroristic enterprise.

The discrepancy between domestic and international terrorism laws often forces authorities to rely on conduct-based criminal charges to hold domestic extremists accountable for their violent behavior. For example, authorities can charge domestic extremists with violating 18 USC 2332a when they unlawfully use a weapon of mass destruction (WMD) or explosive. The DOJ charged Cesar Sayoc with a violation of 18 USC 2332a in 2018 after he mailed 16 bombs to current and former elected leaders across the United States. Section 2332a, unlike the other statutory sections in Chapter 113b that apply to FTOs, only applies to individuals that use, threaten, attempt, or conspire to use weapons of mass destruction or explosives. It does not apply to the domestic terrorists’ counterparts that knowingly and materially support the domestic terror attack, or the broader terrorist organization.

Some argue that this reactive approach—charging only those individuals directly responsible for committing the criminal act after its commission—is not effective and the United States instead needs a more robust domestic terrorism statute. Mary McCord, a law professor at Georgetown University Law School, argues that current charges do not signal the significance of their crimes—a key factor, she argues, in reducing domestic terrorism. She argues the “[domestic terrorism] label carries weight…it signals to Americans that the threat of extremism is just as significant when it is based on domestic political, economic, religious, or social ideologies....”

Bruce Hoffman, a professor at the Georgetown University School of Foreign Service, similarly argues for a domestic terror law and echoes the points of McCord. Specifically, he claims a new law could lessen the sentencing disparity between individuals that are convicted of committing acts of domestic terrorism and those found to have committed violent acts of Islamic extremism. To illustrate, he found that individuals convicted of supporting the Islamic State received an average prison sentence of 13.2 years. In contrast, a member of the neo-Nazi group Atomwaffen Division received a one-year prison sentence for gathering weapons and ammunition for the group. Echoing McCord’s argument, Hoffman asserts that this sentencing discrepancy lessens the...
significance of domestic terrorism and lowers its priority among law enforcement.\textsuperscript{18}

Opponents of a domestic terrorism law argue that a new law is unnecessary and that the current federal code provides sufficient tools to mitigate the threat of domestic terrorism. Michael German, a fellow at the Brennan Center for Justice, claims that the current federal code provides the tools needed to reduce domestic terror, and the issue lies not with a lack of available charging instruments but in prioritization.\textsuperscript{19}

He argues that it is easier to charge domestic terrorists with existing laws and that creating statutes with the label of domestic terrorism does not warrant a broad expansion of government power.\textsuperscript{20} Specifically, he points to powerful tools like hate crime laws, RICO, and conspiracy charges that provide charging ability to prevent future action and support broad prosecutorial discretion.\textsuperscript{21} The use of these charges avoids politically inflammatory courtroom arguments over who qualifies as a terrorist, making it easier for juries to focus on the facts and adjudicate a conviction.\textsuperscript{22}

Others, like Robert Chesney, the dean of the University of Texas School of Law, highlight that any move to expand federal powers toward domestic terrorism will face scrutiny.\textsuperscript{23} Constitutional speech and assembly protections would make it difficult to determine what actions qualify as terrorism.\textsuperscript{24} He argues political adversaries with deep disagreements would abuse such a tool in a domestic context.\textsuperscript{25} Indeed, in the 1920s, 1960s, and 1970s, the federal government cracked down on perceived Marxist or leftist groups and individuals. This included the infamous “Palmer raids” and the FBI’s extralegal counterintelligence program (COINTELPRO).\textsuperscript{26}

As the debate over whether to codify a new domestic terrorism statute continues, authorities should examine another approach that reduces the threat of domestic extremism and prevents events such as the January 6\textsuperscript{th} Capitol Riot. Authorities could consider classifying domestic extremist groups as criminal street gangs. Research shows that domestic extremist groups and criminal street gangs share similar characteristics and that such classification could provide new tools for law enforcement in mitigating the threat domestic extremist groups pose. The following section examines this research and attempts to establish a baseline of similarities that, in turn, can help determine a potential 18 USC § 521 criminal street gang classification for the Proud Boys and Oath Keepers.

**Current Gang Classification Research**

Research suggests that street gang classification of domestic extremist groups could provide a new approach for law enforcement in preventing acts of domestic terrorism. Notably, Shannon E. Reid and Matthew Valasik point out in their book *Alt-Right Gangs: A Hazy Shade of White* that street gangs and far-right extremist groups are characteristically similar in their behavior, use of territory, and adoption of signs and symbols.\textsuperscript{27}

Reid and Valasik note that gangs commonly pursue territorial dominance and tend to loiter in specific public locations.\textsuperscript{28} Referred to as “gang set-space,” these locations provide sanctuary and protection from rival gangs and allow members to avoid confrontations with law enforcement.\textsuperscript{29}

The authors argue that far-right extremists place a similar priority on territory. Far-right groups claim territory through the concept of “natural dominion.”\textsuperscript{30} Natural dominion is the belief that members of the far-right have ownership of any space they occupy.\textsuperscript{31} Examples of far-right groups exercising this belief include conducting hate crimes.
directed at nonwhite minorities moving into predominantly white communities and participating in free speech rallies, such as the Unite the Right rally in Charlottesville, Virginia, in 2017.32

Some far-right groups also use territory through the use of “Aryan free spaces.”33 Similar to gang-set space, an Aryan free space allows far-right extremists—in this case, white supremacists—to openly express their beliefs and organize their activities.34 These spaces include white supremacist bars, music festivals, and conventions.35 For example, “Hammerfest” was an Aryan community music festival where adherents could openly wear extremist paraphernalia and listen to racist music.36 These spaces similarly provide territory without formal control and that offers a safe space for members to freely engage in inappropriate and illegal behavior.37

Far-right groups are also similar to gangs in their criminal activity. Known as cafeteria-style offending, gang crimes are typically varied and unplanned.38 A gang may participate in various crimes, from theft to murder, with little or no planning.39 Reid and Valasik contend that far-right groups participate in a similar pattern of behavior.40

Some far-right groups participate in drug trafficking, counterfeiting, and armed robbery to create a financial incentive for recruits.41 For example, in 2019, members of the New Aryan Empire were indicted for their efforts to murder, kidnap, and maim individuals.42 This indictment was coupled with other charges, including a conspiracy to distribute methamphetamine.43 Other examples of external extremist activity include authorities recently indicting members of the Atomwaffen Division for sex crimes and firearm violations.44

Reid and Valasik also claim that far-right extremist groups, similarly to gangs, adopt signs and symbols to create group solidarity.45 Many far-right groups display banners and tattoos that depict traditional Nazi iconographies, like the SS bolt and swastika.46 In other instances, far-right groups have adopted uniform clothing, mascots, and hand signals. Members of the Proud Boys typically wear black and yellow polo shirts to distinguish themselves and have adopted a cockerel mascot that serves as a symbol on patches, hats, and clothing.47 Furthermore, members of the Proud Boys have regularly displayed the “OK” hand signal to signal solidarity and white supremacist intentions.48

Despite these stark similarities, some scholars disagree with Reid and Valasik’s argument for classifying far-right groups as gangs. Specifically, they argue that authorities should not categorize far-right groups as gangs because far-right groups focus their activities on political action, whereas gang activity is financially motivated.49

In response, Reid and Valasik argue that far-right groups have other goals in conjunction with their political objectives.50 For example, some far-right groups are not interested in political action and only want to “red-pill” large parts of society. Red pilling is influencing an individual to change their beliefs, specifically to a disturbing and extreme point of view that society is controlled by a secret cabal.51 Therefore, Reid and Valasik contend that while some of these groups incorporate political messaging, many only use politics to support their broader far-right agenda.52

Scholars also point out that authorities should not classify far-right groups as gangs because they focus on hate crimes rather than financially motivated crimes.53 Reid and Valasik argue that this may not be the case. Their book demonstrates that far-right groups engage in criminal activity outside of
bias-related hate crimes. In 2010, the Attorney General of California cited a report that claimed a rise in illegal activity among white power groups. This report included reports of credit card theft, vehicle theft, and robbery. Furthermore, Reid and Valasik argue that hate crimes against people of color and criminal street gang activity are not mutually exclusive. They point out that gangs may attack fellow gangs of different races. Notably, Black and Hispanic gangs have historically attacked each other for race-based reasons.

Reid and Valasik’s analysis of criminal street gangs and far-right groups helps illustrate the reality of far-right activity while addressing common misperceptions. The similarities in criminal behavior, use of territory, and adoption of signs and symbols help inform whether categorizing domestic extremist groups as criminal street gangs is a practical approach.

Similar to Reid and Valasik’s approach in Alt-Right Gangs, David C. Pyrooz and James A. Densley also contribute to the discussion of whether authorities could categorize domestic extremist groups as criminal street gangs. In their article “On Public Protest, Violence, and Street Gangs,” Pyrooz and Densley argue that authorities could categorize the far-left extremist network, Antifa, as a street gang. The authors use the principles of violent criminal activity, collective identity, durability, and street orientation described in the California Penal Code (CPC) and the Eurogang program of research to illustrate that Antifa’s behavior warrants a street gang classification.

First, Pyrooz and Densley point out that Antifa participates in the violent criminal activity described in the Eurogang and CPC definition. Antifa implements a direct action strategy that includes illegal tactics such as arson, vandalism, and attacks on public institutions. For example, in 2017, Antifa caused considerable damage at the University of California Berkeley when a far-right activist was scheduled to conduct a speech on campus. During the disturbance, Antifa attacked police with fireworks and Molotov cocktails and used chemical irritants on individuals.

Second, Pyrooz and Densley point out that Antifa is a group with a collective identity. This identity, known as Black Bloc, encourages members to dress up in all black and display signs and symbols such as the Antifaschistische Aktion flag. Like Reid and Valasik’s argument regarding the use of signs and symbols, Pyrooz and Densley argue that Antifa is a gang in that “Bloods wear red; Crips wear blue; Antifa wear Black.”

Third, Pyrooz and Densley argue that Antifa, like other gangs, is durable and exists over long periods of time. This is a product of their ongoing presence at rallies and consistent street brawls with far-right groups such as the Proud Boys. They argue that instead of being a small movement that shows up to public protests, Antifa has frequently persisted in its goal to attack fascist adherents. Much like gangs, violence and threats against Antifa fuel its existence and maintain its continuity.

Finally, Pyrooz and Densley argue that authorities could classify Antifa as a gang due to its street orientation. Common among gangs is that their behavior typically takes place in public settings. The authors argue that Antifa is not a private group, and its actions to destroy capitalistic symbols like Starbucks and Bank of America constitute street-level violence. Moreover, Pyrooz and Densley assert that when Antifa violently brawls with its far-right opponents, it takes place in public.
The case study of classifying Antifa as a street gang provides more evidence that street gangs and domestic extremists share the same characteristics. Pyrooz and Densley’s illustration breaks down the stereotype that gangs only focus on homicides and drug sales and that their members only consist of at-risk youth. Pyrooz and Densley, and Reid and Valasik, inform this paper’s approach to applying the federal definition of a criminal street gang to the Proud Boys and Oath Keepers. Next, I will assess whether the Proud Boys and Oath Keepers can fit into this legal categorization and, accordingly, whether the DOJ should consider utilizing criminal street gang classification as a viable option for prosecuting members of these groups.

**Analysis**

The following section considers the previous research and attempts to determine if federal authorities can categorize the Proud Boys and Oath Keepers as a criminal street gang. This classification will inform whether federal authorities can employ more effective tools to reduce the threat of domestic extremism. To render this determination, I will analyze whether the Proud Boys and Oath Keepers satisfy each of the elements in 18 USC § 521.73

Section 521(a) defines a criminal street gang as:

(a) [A]n ongoing group, club, organization, or association of 5 or more persons—
(A) that has as 1 of its primary purposes the commission of 1 or more of the criminal offenses described in subsection (c);
(B) the members of which engage, or have engaged within the past 5 years, in a continuing series of offenses described in subsection (c); and
(C) the activities of which affect interstate or foreign commerce.

The offenses listed in § 521(c), referenced in subsection (a), include:

(1) a Federal felony involving a controlled substance (as defined in section 102 of the Controlled Substances Act (21 U.S.C. 802)) for which the maximum penalty is not less than 5 years;
(2) a Federal felony crime of violence that has as an element the use or attempted use of physical force against the person of another;
(3) a Federal offense involving human trafficking, sexual abuse, sexual exploitation, or transportation for prostitution or any illegal sexual activity; and
(4) a conspiracy to commit an offense described in paragraph (1), (2), or (3).

**Proud Boys**

The Proud Boys likely satisfy the requirement for criminal street gang classification under § 521(a). First, evidence shows that the Proud Boys are an ongoing group, club, organization, or association of five or more persons. Proud Boys’ rallies are typically larger than five people and consist of members who occupy leadership positions.74 Additionally, clubs and organizations typically have an entrance policy or initiation process that members
must participate in to gain membership. For example, a gang might have an initiation process for prospective members to participate in to gain entry. The Proud Boys are no different. According to the Southern Poverty Law Center (SPLC), the Proud Boys openly declare themselves as a “pro-West fraternal organization.” Moreover, Proud Boy recruits must engage in a ceremonial beating to become a member. This ceremony involves other members beating the recruit until they can yell out five breakfast cereal names.

On the other hand, the Proud Boys do not readily satisfy the second requirement in § 521(A). The stated purpose of the Proud Boys appears to fall in subsection (c). The Proud Boys argue that their goal is to advocate Western chauvinism. This entails spreading “anti-political correctness” and “anti-white guilt.” While inappropriate, this behavior is not likely to constitute any of the federal offenses listed in subsection (c).” However, one can argue that members have engaged in offenses described in subsection (c) within the last five years, particularly the offenses listed in (c)(2), federal felony crimes of violence. In 2018, Proud Boys members committed violent crimes when they physically assaulted counter-protestors at a far-right rally in New York City. Moreover, Proud Boys members violently attacked members of law enforcement during the January 6th Capitol Riot.

Next, the Proud Boys commit violent crimes across state boundaries, thus affecting interstate commerce. Along with the incidents in New York City and Washington, DC, the Proud Boys have engaged in violent behavior in other states. Specifically, the Proud Boys have engaged in violent brawls and attacks against leftist protestors in Portland, Oregon. In 2018, members of the Proud Boys physically assaulted a pedestrian in Portland after the individual began arguing with them over the Proud Boys chanting, “Build the Wall!” The same year, Proud Boys assaulted leftist activists with chemical irritants in Kentucky while sitting at a bar.

Based on the above analysis, there is a strong case that the Proud Boys fit the legal definition of a street gang in 18 USC § 521. The Proud Boys explicitly declare themselves an organization and the size of their rallies suggests their group is more significant than five persons. Moreover, while their intent is not explicitly focused on criminal offenses, members of the organization frequently engage in activities that § 521(c) defines as criminal offenses. These criminal offenses have occurred in multiple states, making them a matter of interstate commerce. By satisfying each clause in 18 USC § 521, this paper concludes that the Proud Boys likely qualify as a criminal street gang.

**Oath Keepers**

The evidence supporting the criminal street gang classification of the Oath Keepers is less clear. First, evidence supports that the Oath Keepers are an ongoing club, organization, or association of five or more persons. Similar to a fraternity, the Oath Keepers are organized into chapters at the national, state, and county levels. Moreover, its leader, Stewart Rhodes, claims the group has 35,000 dues-paying members. While it is unclear if the Oath Keepers have an initiation process like the Proud Boys, they are particular in whom they recruit. Like gangs, who focus on recruiting youth, the Oath Keepers focus on recruiting former members of the military and law enforcement. According to the SPLC, the Oath Keepers “openly and intentionally [recruit] these individuals to their ranks, mostly in an effort to capitalize on the skills and knowledge these
individuals acquired during their time of service.90

Second, it is unclear whether federal authorities could argue that the Oath Keepers’ explicit purpose is committing criminal offenses. Rhodes claims the purpose of the Oath Keepers is to reinvigorate the militia movement that relies on the US Constitution’s Second Amendment by having veterans help individuals “take back into their own hands their own personal self-defense and security.”91

Additionally, much of the Oath Keepers’ activity rests on vigilante activism that falls short of crimes of violence and could be interpreted as constitutionally protected speech and activity. The Oath Keepers’ activity usually includes individuals forming armed groups to protect businesses and individuals from whom they believe to be tyrannical elements like the federal government and racial justice protestors.92 Compared to the Proud Boys, the organized activity of the Oath Keepers tends to hide behind constitutionally protected speech and activity. But, with proper evidence, federal authorities could argue that their activity could fall under subsection (c)(4), conspiracy to commit a federal felony crime of violence.

Third, if there was evidence of a conspiracy to commit a felony crime of violence, it is plausible that it would affect interstate commerce. Given the national organization of the Oath Keepers, the group transcends state boundaries, making it a federal matter. Moreover, its constant involvement in standoffs with federal agents, such as the one with Clive Bundy in 2014, automatically creates a situation that affects interstate commerce because of their obstruction of federal enforcement matters.93 The case to classify the Oath Keepers as a criminal street gang is weaker than the case for the Proud Boys. The Oath Keepers are an organized group of five or more people whose activity transcends state boundaries. But, their stated intent, along with the typical behavior of their members, makes it hard to classify them as a gang because their activity usually stays within the boundaries of constitutionally protected speech and activity. As a result, the Oath Keepers likely fail to satisfy the second and third requirements for §521 classification.

This is not to say federal authorities cannot ever classify the Oath Keepers as a gang. If the evidence establishes a conspiracy to commit a federal felony crime of violence, the group would satisfy § 521 and authorities could classify them as a street gang. Notably, Rhodes and other members were convicted of seditious conspiracy in 2022 for their role in planning to oppose the federal government forcefully during the January 6th Capitol Riot.94 Thus, like the evidence available for the Proud Boys, these isolated criminal incidents could support characterizing the Oath Keepers as a criminal street gang.

Options & Issues

A gang classification of the Proud Boys and a potential classification of the Oath Keepers could lend new tools to reduce the threat domestic extremists pose in the United States. Specifically, gang classification could allow authorities to be proactive and engage in anticipatory enforcement. These tools can improve early warning and prevent events like the January 6th Capitol Riot from occurring again. The following section discusses these tools and explains how they can reduce the harm extremist groups pose.

First, a gang classification of the Proud Boys or Oath Keepers could allow law
enforcement to track their smaller-scale acts of violence through existing gang databases.\textsuperscript{95} For example, suppose a Proud Boy frequently participates in violent assaults against leftist protestors. In that case, that information should be made available to law enforcement and prosecutors so they can aggressively prosecute that member for their more minor crimes.\textsuperscript{96} This can degrade the threat the Proud Boys pose by taking members off the street before they can participate in serious acts of violence like the January 6\textsuperscript{th} Capitol Riot.

Additionally, gang databases can help law enforcement track and monitor arrested Proud Boys and Oath Keepers as soon as they are released from prison.\textsuperscript{97} A prison could pass information and intelligence to the FBI field office, including where the inmate is going to live.\textsuperscript{98} This could allow for targeted monitoring and intervention programs designed to reduce the chance of an individual quickly rejoining the extremist organization.\textsuperscript{99}

Classifying the Proud Boys and Oath Keepers as a criminal street gang could also allow federal authorities to disrupt their activity using civil gang injunctions (CGI).\textsuperscript{100} A CGI is a restraining order prohibiting gang members from activity in a designated area, bringing criminal penalties if violated.\textsuperscript{101} Historically, this tool has been used to disrupt illegal gang activity, namely graffiti, drug sales, and firearm possession.\textsuperscript{102} A CGI can also be easier and quicker to obtain because it is applied through civil courts.\textsuperscript{103} Compared to criminal courts, civil courts do not require high standards of evidence to adjudicate legal motions.\textsuperscript{104} A CGI applied to the Proud Boys and Oath Keepers could allow preventative action by expressly enjoining their violent behavior in sensitive locations, such as federal buildings, and at scheduled protests and other high-profile events.\textsuperscript{105} This would allow these groups to participate in constitutionally protected protest, but carry the added weight of criminal charges if members pursued violence.

Finally, gang classification of the Proud Boys and Oath Keepers can expose these groups to criminal liability under 18 USC CH 96, Racketeer Influenced and Corrupt Organization (RICO). Originally codified to prosecute the mafia for their organized crime activity, this statute allows prosecutors to charge individuals associated with a criminal enterprise, even if they did not directly participate in the enterprise’s criminal activity.\textsuperscript{106} Central to a RICO charge is establishing an enterprise, which can include crime families, drug cartels, and street gangs.\textsuperscript{107}

Given the possibility that the Proud Boys and, to some extent, the Oath Keepers qualify as a criminal street gang under 18 USC § 521, prosecutors could use RICO as a powerful tool to reduce the harm these groups pose. Upon establishing a pattern of racketeering activity—which could include assault, intimidation, and, more recently, seditious conspiracy—prosecutors could declare the Proud Boys were engaging in a “racket,” allowing them to charge individuals who claim membership.\textsuperscript{108} This could enable law enforcement to proactively dismantle the group and reduce its ability to conduct future acts of violence.\textsuperscript{109}

Notwithstanding the benefits derived from classifying the Proud Boys and Oath Keepers as criminal street gangs, this approach also presents several issues. First, gang classification could present issues regarding constitutionally protected speech and activity. For example, a CGI could infringe on an individual’s freedom to associate if the injunction’s scope was not narrow enough to only prohibit certain individuals from committing offenses in
Moreover, courts may still interpret 18 USC § 521 as overly broad. However, gang definitions at the state level have routinely been upheld.\textsuperscript{111}

Second, gang classification could result in disproportionate harm and lead to a rise in incarcerations, also known as “net widening.”\textsuperscript{112} Indeed, the DOJ inappropriately engaged in net widening when it classified the “Juggalos” as a criminal street gang in 2017.\textsuperscript{113} The Juggalos are a group of Insane Clown Posse fans who dress up as clowns to show their support.\textsuperscript{114} Members often engage in inappropriate behavior involving violent imagery and discussion of murder scenarios.\textsuperscript{115} After Juggalo members engaged in a pair of crimes in 2014, the DOJ moved to classify the Juggalos as a criminal street gang.\textsuperscript{116} It became apparent, though, that a few individuals’ activities did not involve the rest of the group, leading the American Civil Liberties Union (ACLU) to file a lawsuit successfully challenging the gang classification.\textsuperscript{117}

Finally, the FBI’s unwillingness to prioritize far-right extremism like it prioritizes other forms of domestic extremism could render a gang classification approach moot. For gang classification to work and afford prosecutors the enforcement tools listed in this paper, the FBI must prioritize investigating far-right groups. Historically, the FBI has been reluctant to do this.\textsuperscript{118} Between 2004 and 2005, the FBI declared eco-terrorism the number one domestic terror threat, even though the number of deaths from eco-terrorism was zero.\textsuperscript{119} Moreover, when eco-terrorists raided a Smithfield pork factory in 2017, the FBI launched an extensive investigation viewed as disproportionate to a crime involving the theft of two piglets worth $40 each.\textsuperscript{120} A continuation of this trend—deflecting attention from the seriousness of far-right extremism and focusing on other politically popular objectives—will reduce the benefits gang classification can yield and fail to mitigate the harm domestic extremists pose.

Conclusion

The January 6th Capitol Riot has renewed the debate over how the United States should mitigate the threat of domestic extremist groups. Unlike transnational terrorism, the United States does not have a domestic terrorism law that authorities can use to charge individuals for acts of domestic terrorism. Instead, prosecutors rely on existing criminal statutes to hold far-right groups accountable for their violent actions. This approach precludes anticipatory enforcement. The absence of a charging practice specifically tailored to curtail domestic terrorism inhibits law enforcement from preventing violent events like the January 6th Capitol Riots because extremist groups often operate within constitutionally protected speech and activity.

Street gang classification under existing federal statutes may provide an enforcement option to offset these deficiencies. Street gang classification can allow for better information sharing with existing gang databases and enable prosecutors to use RICO charges and CGIs to dismantle groups that commit violent crimes across state boundaries.

While gang classification is a promising solution for domestic extremist groups, it does not come without issues. Gang classification could easily lead to the “net widening” of groups on the edge of gang classification and increase unjust incarcerations. Moreover, CGIs could be ruled unconstitutional if prosecutors do not narrowly tailor the injunction to violent activity. Future research should investigate...
if and how current federal gang enforcement disproportionately harms minority groups and determine the likelihood that such harm could transfer onto efforts to counter domestic extremists.

About the Author: Michael P. Losacco is a Georgetown University Security Studies Program (SSP) graduate student. Before pursuing a master’s degree, he served as an active-duty armor officer in the U.S. Army, where in 2017, he deployed to Afghanistan and supported NATO’s Resolute Support mission in Paktia province.

2 Ibid.

3 Ibid.

4 Ibid and 18 USC Ch. 113B. https://www.law.cornell.edu/uscode/text/18/part-I/chapter-113B.

5 Ibid.

6 Ibid.

7 Ibid.

8 Ibid.

9 Ibid.

10 Ibid.

11 Ibid.


13 Ibid.

14 Andrea Ho, Professor Bruce Hoffman on Domestic Terrorism, Georgetown Journal of International Affairs Dialogues, April 21, 2021. https://gjia.georgetown.edu/2021/04/21/professor-bruce-hoffman-on-domestic-terrorism/.

15 Ibid.

16 Ibid.

17 Ibid.

18 Ibid.


20 Ibid.

21 Ibid.

22 Ibid.

23 Chesney, “Should We.”

24 Ibid.

25 Ibid.


29 Ibid, 83.

30 Ibid, 86.

31 Ibid.

32 Ibid, 87.

33 Ibid, 87-90.

34 Ibid, 88.

35 Ibid.

36 Ibid.

37 Ibid.

38 Ibid, 112.

39 Ibid, 112-113

40 Ibid.

41 Ibid, 113.

42 Ibid, 113.

44 Reid and Valasik, “Gangs,” 113.
46 Ibid.
48 Ibid.
49 Ibid, 38.
50 Ibid, 39.
51 Ibid.
52 Ibid.
53 Ibid, 35.
54 Ibid.
55 Ibid.
56 Ibid.
58 Ibid, 232.
60 Ibid.
61 Ibid.
62 Ibid.
63 Ibid, 234.
64 Ibid.
65 Ibid.
66 Ibid, 234.
67 Ibid.
68 Ibid.
69 Ibid.
70 Ibid.
71 Ibid.
72 Ibid.
75 Reid and Valasik, “Gangs,” 4.
76 Extremist Files, “Proud Boys.”
77 Ibid.
78 Ibid.
79 Ibid.
80 Ibid.
81 Ibid.
82 Ibid.
83 Ibid.
84 Ibid.
85 Ibid.
86 Ibid.
88 Ibid.
90 Extremist Files, “Oath Keepers.”
91 Ibid.
92 Ibid.
93 Ibid.

93 Reid and Valasik, “Gangs,” 130.
94 Ibid.
95 Ibid.
96 Ibid.
97 Ibid.
98 Ibid.
99 Ibid.
100 Ibid, 85.
101 Ibid.
102 Ibid.
103 Ibid.
104 Ibid.
105 Ibid.
107 Ibid.
108 Ibid.
109 Ibid.
112 Ibid, 234.
113 Ibid.
114 Ibid.
116 Ibid.
117 Ibid.
118 German, “New Laws.”
119 Ibid.
120 Ibid.
The Worst Menace to Society: Turkey’s Approach to Digital Censorship

Jessica Maksimov

In 2013, Turkey’s ruling Justice and Development Party (AKP) was faced with the Gezi Protests—a movement consisting of over three million protestors who were primarily mobilized via online platforms. For over three weeks, the protests grew as people gathered to voice their discontent with the AKP’s growing authoritarianism. This paper deems the Gezi Protests as a critical juncture in Turkey's technology policy, as it forced the ruling party to reexamine the Internet’s ability to mobilize and organize dissent. Consequently, the AKP used its legislature and executive agencies to centralize Internet control and limit Internet freedoms. Through an analysis of Turkey’s Internet censorship laws and Freedom House’s “Freedom on the Net” reports, this study finds that between 2013-2018, Turkey’s government doubled the country’s Internet controls. This increase in legislative and punitive measures was aimed at fomenting self-censorship online by prosecuting and jailing online activists and restricting Internet freedom and anonymity. Given that Turkey is a party to international conventions upholding peoples’ rights to online privacy and freedom of expression, the government’s lack of commitment to these principles is noteworthy. The AKP's pursuit of digital authoritarianism calls for increased scrutiny from the United States and fellow NATO allies. Washington should communicate that there will be consequences for Ankara if it continues adhering to unacceptable global Internet practices that are at odds with the democratic principles tying the NATO alliance together.

Introduction

In May 2013, plans to demolish Istanbul's Gezi Park drove a group of activists to hold peaceful sit-ins at the site. In response, the ruling Justice and Development Party (AKP) ordered their evacuation, using excessive police force. As the government cracked down on the protestors, live-stream channels and social media accounts across Turkey lit up with reports of the police’s brutality. For the next three weeks, a leaderless movement—consisting of approximately three million protestors—swept across Turkish streets voicing anti-government sentiments. As the protests grew, the Turkish participants’ list of grievances widened, ranging from limited LGBTQ+ rights to repressed freedoms of expression and assembly. In this flurry of discontent, a ubiquitous theme emerged: opposition to the AKP’s increasingly authoritarian rule. When the Gezi Park protests erupted, government-controlled mass media channels refused to broadcast the protests, airing cooking shows and penguin documentaries instead. This censorship prompted Internet users and citizen journalists to broadcast the events via online platforms. At the time of the protests, Turkey had an Internet penetration rate of 47 percent among its population—with Facebook’s penetration rate totaling around 41 percent, Twitter’s 31 percent, and YouTube’s 16 percent. These social media platforms proved to be an effective means for mobilizing the population to march in protest of the government’s brutality in Gezi. In response to the Gezi protests, the government reoriented its domestic cyber posture to resemble digital authoritarianism—which consists of digital information technology that authoritarian regimes use to surveil, repress, and manipulate domestic and foreign online audiences.

This paper traces the Turkish government’s altered approach to Internet freedom.
following the Gezi Protests in 2013 and determines that the AKP’s response to the protests triggered a path toward digital authoritarianism in Turkey.9 Through an analysis of Freedom House’s “Freedom on the Net” reports, which annually score Turkey’s Internet and digital media restrictions, this study calculates the state’s implementation of domestic Internet controls, which limited Internet freedom in Turkey.

Literature Review

In accordance with diffusion theory, anti-government protests can encourage comparable uprisings in neighboring states with similar government structures and popular grievances. As a result, power structures in states near these movements grow fearful of anti-government mobilization within their borders.10 Accordingly, when the Arab Spring spread through Turkey’s neighboring states, the AKP grew increasingly wary of similar protest movements arising and overthrowing the party.11 According to scholar Zafir Yilmaz, “after the Gezi Uprising, internal opposition activities were defined as the greatest threat to security by the National Security Council,” indicating a shift in policy following a grassroots crisis that challenged the rulers’ authority.12

The Gezi Protests constituted a “critical juncture,” or a crisis that evoked a reorientation in policy objectives.13 Recognized political scientists Giovanni Capoccia and Daniel Keleman argued in 2007 that these junctures are “brief phases of institutional flux” during which the ruling party’s decisions close off alternative options and establish “institutions that generate self-reinforcing path-dependent processes.”14 The Gezi protests reoriented the AKP’s understanding of the Internet’s mobilizing and organizational capacities. It is also relevant to consider the Internet in the context of globalization, which in the 2010s exposed Internet users to an international catalog of real-time reporting.15

Turkey is a party to international conventions and agreements that reinforce peoples’ rights to online privacy and freedom of expression, especially Article 10 of the European Convention on Human Rights. This convention specifically protects one’s freedom to hold opinions, receive, and impart information without interference from public authorities.16 However, starting in the early 2010s, Turkey implemented legislation blocking, filtering, surveilling, and storing the data of websites that the government perceived as oppositional to the ruling party’s authority.17 These efforts aimed to increase the state’s Internet governance: control over the flow of information on the Internet and the online structure that maintains this online traffic.18 To counteract the Internet’s potential for facilitating anti-government mobilization, the AKP began restricting Internet freedoms through legislative and punitive measures. Consequently, Turkey used its domestic laws to increase its Internet sovereignty by treating the digital space as a territory and controlling the flow of information within its borders.19 As the data below show, these Internet restrictions expanded further after the government faced terrorist attacks and a failed coup d’état attempt in 2016.20

Internet Freedom Restrictions Begin

In March of 2007, the AKP instituted its first legal mechanism for Internet regulation—Law No. 5651—implemented with the pretext of protecting children and preventing access to illegal and harmful
online content. This legislation allowed government officials to block websites containing content of “child sexual abuse, drug use, the provision of dangerous substances, prostitution, obscenity, gambling, suicide promotion, and crimes against Mustafa Kemal Atatürk.” The Telecommunications Communication Presidency (TİB) and the judiciary were put in charge of censoring these topics with varying degrees of transparency. Even after the passage of Law No. 5651, however, courts continued to uphold the freedom of speech online and reverse the TİB’s site-blocking decisions. Undeterred, activists increasingly relied on social media to run their operations.

The onset of the Gezi protests shifted the government’s understanding of the role of the Internet in Turkey’s society. The state’s first National Cyber Security Strategy and Action Plan—published one month after Gezi—acknowledged that Information and Communication Technology (ICTs) usage was “spreading rapidly” in Turkey and played “important roles in all aspects of our lives.” The government’s new understanding of the Internet’s prevalence in people’s lives and its contribution to anti-government mobilization led to legislative changes that restricted Internet freedoms from 2014 onward.

### Methodology

To unearth the state of Turkey’s Internet freedom after Gezi, this paper analyzes online restrictions preceding and succeeding the 2013 protests. Additionally, it briefly assesses the AKP’s response to the 2016 attempted coup, which further challenged the government’s grasp on power. The data show that the government perceived the protests and coup, organized via social media, as challenging its authority. The aftermath of Gezi saw increased restrictions online as the government tried to deter anti-government sentiments from spreading.

Using the Freedom House “Freedom on the Net (FOTN)” reports, this study assesses the state of Turkey’s domestic Internet freedoms using three indicators: 1) obstacles to Internet access, 2) limits to online content, and 3) violations of user rights. This paper uses Freedom House’s definitions below:
A) **Obstacles to Access:**

Assesses infrastructural and economic barriers to access; government efforts to block specific applications or technologies; and legal, regulatory, and ownership control over the Internet and mobile phone access providers.

B) **Limits on Content:**

Examines filtering and blocking of websites; other forms of censorship and self-censorship; manipulation of content; the diversity of online news media; and usage of digital media for social and political activism.

C) **Violations of User Rights:**

Measures legal protections and restrictions on online activity; surveillance; privacy; and repercussions for online activity, such as legal prosecution, imprisonment, physical attacks, or other forms of harassment.

The FOTN reports aggregate all three indicators to determine a state’s overall Internet and digital media freedom score, which ranges from 0-100 (0 = the most free and 100 = the least free). This study tracks these scores starting from 2009 (when the first FOTN report was released) and ending with 2018 (when the FOTN scoring scale was altered). It then charts the AKP’s implementation of “Key Internet Controls” and analyzes the FOTN Reports’ explanations behind each year’s Internet freedom scores.

**Findings**

The data show that the AKP reoriented its approach to Internet freedoms following the 2013 protests by increasing its control over Internet governance and increasing punitive measures against oppositional online voices. Figure 1 displays Turkey’s annual FOTN scores—showing that the largest increases in online restrictions occurred between the years 2013-2014 (from 49 to 55 points) and 2016-2017 (from 61 to 66 points). While the 2013 FOTN report’s time period does not encompass the Gezi protests, or their aftermath, within its date range, they are included and analyzed in the 2014 report. According to the 2014 FOTN Report, in the five years that elapsed since the first FOTN Report was released in 2009, Turkey’s Internet freedom declined by a total of 13 points—with a decline of 6 points within one year of the Gezi protests.
Figure 2 shows that Internet controls spiked from 3 to 7 (out of 9 possible controls) between 2013 and 2014. In 2013, the three key Internet controls in Turkey constituted: 1) the blocking of social media and/or communication applications, 2) the blocking of political, social, and/or religious content, and 3) arrests, imprisonment, and/or prolonged detention of online journalists, bloggers, or ICT users for having posted political or social content. In the 2014 report, four more Internet controls were added to this list, including: 1) the passage of new laws that increased censorship and surveillance, 2) technical attacks against government critics, and 3) physical attacks on or 4) killings of online journalists, bloggers, and ICT users. Over the following years, the number of controls in Turkey remained high, with arrests and detentions of journalists, bloggers, and ICT users occurring every year. The government’s use of its institutions to detain, prosecute, and imprison oppositional voices showed the AKP’s increasing hold on power and its determination to silence criticism.
**Figure 2: Turkey’s Key Internet Controls (2009-2018)**

<table>
<thead>
<tr>
<th>Year</th>
<th>Social media and/or communications apps (Web 2.0) blocked</th>
<th>Political, social, and/or religious content blocked</th>
<th>Localized or nationwide ICT shutdown</th>
<th>Pro-Government commentators manipulate online discussions</th>
<th>New law/directive increasing censorship or punishment passed</th>
<th>New law/directive increasing surveillance or restricting anonymity passed</th>
<th>Online journalist/blogger/ICT user arrested, imprisoned, and/or in prolonged detention for political or social content</th>
<th>Online journalist/blogger/ICT user physically attacked or killed</th>
<th>Technical attacks against government critics and human rights organizations</th>
<th>TOTAL # of Key Internet Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>2011</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>2012</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2</td>
</tr>
<tr>
<td>2013</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>3</td>
</tr>
<tr>
<td>2014</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>7</td>
</tr>
<tr>
<td>2015</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>7</td>
</tr>
<tr>
<td>2016</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>6</td>
</tr>
<tr>
<td>2017</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>7</td>
</tr>
<tr>
<td>2018</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>5</td>
</tr>
</tbody>
</table>

This data shows that the critical juncture in 2013 set Turkish Internet controls on a self-reinforcing, path-dependent course for the next five assessed years. As Figure 3 (below) shows, the Internet restrictions continued to steadily increase until two of Turkey’s defining criteria were downgraded from “Partly Free” to “Not Free” in 2016. To understand the de facto measures and controls that the scoring reflects, the next section analyzes the changes that amounted to these declines in Internet freedom.
Figure 3: Turkey’s Freedom on the Net Scores Table

<table>
<thead>
<tr>
<th>Year</th>
<th>FOTN Score (0-100)</th>
<th>Obstacles to Access (0-25)</th>
<th>Limits to Content (0-35)</th>
<th>Violations of User Rights (0-40)</th>
<th>Category (F=Free, PF=Partly Free, NF=Not Free)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>40</td>
<td>11</td>
<td>13</td>
<td>16</td>
<td>PF</td>
</tr>
<tr>
<td>2011</td>
<td>45*</td>
<td>12</td>
<td>16</td>
<td>17</td>
<td>PF</td>
</tr>
<tr>
<td>2012</td>
<td>46</td>
<td>12</td>
<td>17</td>
<td>17</td>
<td>PF</td>
</tr>
<tr>
<td>2013</td>
<td>49</td>
<td>12</td>
<td>18</td>
<td>19</td>
<td>PF</td>
</tr>
<tr>
<td>2014</td>
<td>55</td>
<td>14</td>
<td>18</td>
<td>23</td>
<td>PF</td>
</tr>
<tr>
<td>2015</td>
<td>58</td>
<td>13</td>
<td>20</td>
<td>25</td>
<td>PF</td>
</tr>
<tr>
<td>2016</td>
<td>61</td>
<td>13</td>
<td>21</td>
<td>27</td>
<td>NF</td>
</tr>
<tr>
<td>2017</td>
<td>66</td>
<td>13</td>
<td>23</td>
<td>30</td>
<td>NF</td>
</tr>
<tr>
<td>2018</td>
<td>66</td>
<td>11</td>
<td>25</td>
<td>30</td>
<td>NF</td>
</tr>
</tbody>
</table>

*Though the FOTN score increased by 5 points between 2009 and 2011, this increase occurred over the course of two years, rather than one year. This is due to the Freedom House’s data release timeline of its initial FOTN reports. The biggest one-year jumps in scores are demarcated in bold.

The Freedom on the Net Scores, Explained

The government’s response to the 2013 antigovernment protests included assaults on online journalists, increased blocks on social media platforms, new amendments to Law No. 5651, and increased powers granted to the TİB and National Intelligence Organization (MIT). Moreover, then Prime Minister Recep Tayyip Erdoğan himself insulted the protesters, calling them “çapulcular” (translated as looters and marauders) in an effort to downplay their credibility, revealing his disapproving sentiments towards the movement.31 During the Gezi protests, many citizen journalists covering the movement were physically assaulted and jailed. One specific case included a freelance journalist named Ahmet Sik, who was deliberately hit on the head with a tear gas canister.32 In addition, approximately 59 journalists were fired from their publications for criticizing the government.33

Aside from the protests, in February 2014, recordings implicating high-level Turkish officials, including Erdoğan, in a corruption scandal were posted on YouTube and SoundCloud. The state then blocked YouTube to suppress the audio clips’ circulation and attempted to suspend oppositional Twitter accounts. Twitter, however, refused to follow government orders. In response, PM Erdoğan vowed to “wipe out Twitter” and called social media the “worst menace to society.”34 The Turkish government then blocked access to the platform. Moreover, following this perceived challenge to the AKP’s authority, new legislation equipped the state regulator, BTK (Information and Communication Technologies Authority), with the power to block websites without a court order—withdrawning site blocking from the judiciary’s jurisdiction.
In the legislative realm, Law No. 6532 and Law No. 5651 were amended to increase Internet Service Providers’ (ISP) accountability to government agencies. In February 2014, Turkey’s Grand National Assembly (GNA) passed new amendments to Law No. 5651, which broadened two of the Telecommunications Communication Presidency’s (TİB) powers. First, online hosting providers were required to store all hosting-related data for up to two years, provide it to the TİB upon request, and perform any actions requested by the TİB. Secondly, the TİB was granted the power to direct ISPs to block specific URLs without a preceding judicial review if the content was seen as violating an individual’s right to privacy. Under the amendments, the TİB staff also received impunity from criminal investigations if they committed crimes while carrying out their duties. As Freedom House reported in 2016, “criminal investigations into TİB staff [could] only be initiated through an authorization from the TİB director, and investigations into the director [could] only be initiated by the relevant minister. This process casts serious doubt on the functioning and accountability of the TİB.” Furthermore, the amendments required that ISPs create a single Association of Access Providers in which membership was mandated by law. Only Association members could then obtain “activity certificates” from the BTK allowing them to legally provide Internet access within Turkey.

In April 2014, the GNA also passed new amendments to Law No. 6532, titled “Amending the Law on State Intelligence Services and the National Intelligence Organization.” These amendments granted MIT agents unlimited access to online communication data (without requiring a court order) and limited their accountability for any wrongdoing. According to Freedom House, “the law force[d] public and private bodies… to provide the MIT any requested data, documents, or information regarding certain crimes, such as crimes against the security of the state, national security, state secrets, and espionage. Failure to comply [was] punishable by prison.” This further centralized the online space under government authority.

It is also important to note that in 2016, the ruling party experienced new challenges to its authority in the form of a failed coup and a series of terrorist attacks. After surviving the failed coup, the AKP imposed a State of Emergency, through which it then solidified Internet control by issuing emergency decrees to Law No. 5651, further restricting Internet freedoms. In April 2017, the government used an amendment to Article 159 of the Constitution to restructure the previously independent Council of Judges and Prosecutors (CJP). The amendment to the CJP, which admits, appoints, transfers, promotes, and supervises judges and prosecutors, allowed the President to appoint six of thirteen CJP members and the Parliament to appoint the remaining seven. Over 3,500 judges and prosecutors—labeled oppositional Fethullahist Terrorist Organization members—were dismissed following the coup. The AKP replaced those dismissed with pro-government appointees, decreasing the independence of the judiciary, which led to “hundreds of arbitrary arrests, detentions, and unfair trials of journalists and civil society actors.” The data shows that Turkey’s amendments have coincided with moments of political turmoil and a resulting government-perceived need to suppress dissent.
Conclusion

In 2013, online restrictions were not a novelty in Turkey. However, the data shows a steep increase in the legislative and punitive measures Turkey’s government implemented after Gezi’s Internet-mobilized movement challenged the AKP’s authority.

At this critical juncture, the flow of online information presented an existential challenge to the state’s ability to control the national narrative about transpiring events. As a result, Internet controls experienced a nearly two-fold increase. The government’s actions increased its sovereignty and control over the Internet’s flow of information. The state mandated ISPs to store and share their operational data (subject to government oversight); tightened the government’s Internet governance controls by expanding the TİB’s authority to filter and block objectionable content and the MIT’s surveillance powers; and attempted to evoke online self-censorship by using arrests, assaults, and job terminations of oppositional online voices. Through the combination of all these measures, Turkey’s digital authoritarianism grew dramatically in the years following Gezi and the failed coup attempt, and oppressive measures continued to increase thereafter.

Although this study only examines one aspect of the intricate puzzle comprising Turkey’s approach to technology policy, it sheds light on the state’s efforts to control the flow of information to and from its domestic audience. Turkey’s international commitments to protecting freedoms of opinion and the impartation of information without government interference starkly contrast its domestic policies. Ankara’s practice of digital censorship at home demonstrates that this NATO member has prioritized its retention of domestic power over its citizens’ Internet freedoms. Turkey’s pursuit of digital authoritarianism requires increased scrutiny from the United States’ national and cyber security agencies. The United States should communicate clearly with Ankara that there will be consequences if it keeps adhering to unacceptable global Internet practices that are at odds with the democratic principles tying the NATO alliance together. Inaction could lead to a more polarized online regulation space between the United States and Eurasia and a further deterioration of Internet freedoms globally.

About the Author: Jessica Maria Maksimov is an M.A. candidate in Eurasian, Russian and East European Studies at Georgetown University. She received her B.A. in Government, with minors in Russian and Spanish, from Georgetown University’s College of Arts & Science. Her research is primarily centered around U.S., Russian, and Eastern European security as well as global technology policy.


6 Ibid.


22 Ibid.


25 Ibid.

26 Ibid, 23.


28 The 2013 report cuts off in April, 2013 a month before the Gezi protests began.


30 Figure 1 demonstrates the increases in Internet control measures, displaying all three indicators aggregated into the final scores.


37 Ibid.


40 Ibid, 3.
Lessons Learned and Neglected: Germany and the Viability of the Offensive Before World War I

Christian Trotti

As the United States prepares for the future of large-scale conventional warfare, policymakers and military planners must heed the warnings of the past. Accordingly, the pre-World War I period deserves greater attention as a case of innovation failure. While the conflicts of the late 19th and early 20th centuries were revealing a shift toward defensive advantages, European armies maintained a strong bias toward offensive operations and decisive battles. Due to this incongruity, their offensive doctrines produced bloody stalemates on the battlefields of World War I. To prevent similar misinterpretations of contemporary conflicts (e.g., in Ukraine) before the next major war, this paper asks the following question: why did the European militaries—particularly Germany—fail to learn the right lessons prior to World War I?

By evaluating three antebellum conflicts (the Franco-Prussian, Spanish-American, and Russo-Japanese Wars), and through historical inspection of the prewar German General Staff and naval leadership, this paper finds that the primary culprit was not irrational officers, insufficient technological adoption, or institutional resistance to change. Rather, it was a failure of vision, caused by a learning environment in which the historical and contemporaneous case studies offered only inconsistent ‘lessons learned’ about the future of warfare. The pre-World War I conflicts had all demonstrated both the increasing potency of defensive firepower, on the one hand, and the continued viability and decisiveness of offensive doctrine, on the other. In this confusing context, preexisting institutional and cognitive biases toward the offensive became the only available means with which German officers could interpret evolving military trends with relative certainty, leading them to bet everything on the rapid decisiveness of offensive operations on land and major naval engagements at sea. To avoid such an outcome today and facilitate more prescient innovation, this paper recommends improvements to U.S. military training, education, and career trajectories.

Introduction

Military innovation in peacetime has long been a topic of great concern for scholars and policymakers alike. Rarely are the stakes so great, and the prospects of success so daunting. While technological, operational, and organizational trends in military affairs are constantly evolving—thereby complicating the ability to predict the future of warfare—the geopolitical actor that best envisions and prepares for that future can often seize a first-mover advantage in the next war.1 Today, the complexity of adopting and integrating emerging technologies like artificial intelligence, quantum computing, and hypersonic weapons has raised concerns about the United States’ readiness for a potential conventional war with great-power competitors China or Russia. Therefore, leading figures in academia and the U.S. Department of Defense (DoD) are actively mining military history for lessons about peacetime innovation that are applicable to contemporary security challenges.2 The Interwar Period between World Wars I and II, for example, when Germany pioneered modern armored warfare and the United States transformed naval and amphibious combat, is one of the most popular Revolutions in Military Affairs to study because it was a resounding success.3 However, there are two problems with the existing literature on innovation: 1) it primarily focuses on technological innovation at the expense of the operational,
doctrinal, and organizational changes, as well as the historical learning, necessary to leverage such technology; and 2) it is overly biased toward success, even though the primary concern among policymakers is failure.

Therefore, this paper posits that the pre-World War I period deserves greater attention as a case of innovation failure. While this disastrous conflict is already essential to the academic literatures on both the offense-defense balance and the causes of war, it is often neglected in the literature on military modernization in peacetime. The primary failure among European armies prior to World War I was a strong bias toward offensive operations and decisive battles (i.e., the “cult of the offensive”), while the conflicts of the late 19th and early 20th centuries were increasingly revealing a shift toward defensive advantages. However, the most disturbing feature of this period was that many of the belligerents—especially Germany—had already anticipated these defensive trends based on close study of the wars at the time, and yet they still believed that their prewar offensive doctrines were not only possible, but preferable. Yet on the battlefields of World War I, these doctrines produced bloody stalemates, trench warfare, and attrition on the Western Front, as well as a lack of decisive strategic effects on the more maneuverable Eastern Front. Similarly, European navies expected to achieve decisive strategic victories in major battleship engagements at sea, rather than adequately preparing for the smaller-scale naval action necessary for maintaining or undermining long-term blockades. This inability to innovate novel operational and tactical approaches before World War I was not a failure of technological adoption, but rather a failure of historical learning and a failure of vision. And most importantly, it is a warning for policymakers and military planners today. If the professional officer corps of those pre-World War I states—who were often astute students of military history—failed to grasp the ‘right’ lessons from conflicts like the American Civil War, the Franco-Prussian War, the Boer War, the Spanish-American War, the Russo-Japanese War, and the Balkan Wars, then it is entirely possible that current DoD officials can overlook or misinterpret the lessons of today’s conventional conflicts like the Second Nagorno-Karabakh War and the Russo-Ukrainian War, thereby laying the foundation for failure in the next great-power conflict. Therefore, this paper strives to answer the following question: why did the European militaries—particularly Germany—fail to learn the right lessons prior to World War I?

In addressing this question, many political scientists paint German military officials such as Generals Helmuth von Moltke, Alfred von Schlieffen, and Helmuth von Moltke the Younger—the architects of German offensive operations in the army—and Admiral Alfred von Tirpitz—one of the leaders of German naval modernization—as largely irrational or at least short-sighted. Indeed, academics attribute Germany’s failure to almost inevitable structural processes such as: the inability of these military officials to adapt to systemic technological trends that favored only defensive operations; institutional biases toward offensive operations despite their “operational infeasibility”; shattered civil-military relations that did not sufficiently constrain offensive-minded military leaders; or ambitious political objectives and binding alliance commitments that tied the hands of military leaders. These arguments are premised on the idea that defensive trends were so obvious that only structural processes could have prevented these men from recognizing and adapting to the technological circumstances of the time.
However, a closer historical inspection of the German General Staff and naval leadership reveals a more complicated story. Upon evaluating several of the wars prior to World War I—many of which served as case studies for German military planners—it is evident that there were not any uniform ‘lessons learned’ about the future of warfare. Rather, the problem was that wars ranging from the American Civil War and Franco-Prussian War to the Spanish-American War and Russo-Japanese War bore a variety of conflicting lessons—and the primary tension was between the increasing potency of defensive firepower on the one hand, and the continued viability and decisiveness of offensive doctrine on the other. Indeed, offensive operations had achieved major strategic victories in all those conflicts, albeit at great cost, despite the defensive advantages accrued by artillery, rifles, and field fortifications. And major naval engagements often decisively complemented these offensive ground operations. It is easy in hindsight to distinguish the ‘right’ lessons of those conflicts (e.g., the reign of artillery, the necessity for combined arms and decentralized command, and the difficulty in achieving decisive results on land or at sea) from the ‘wrong’ lessons (e.g., the ability of offensive initiative to prevail through superior morale, advanced preparations, and operational envelopment)—but only because ‘right’ and ‘wrong’ are judged in accordance with how appropriate those lessons were for the eventual World War I. In reality, however, all of those lessons were equally true at the time, and German military planners were faced with the unenviable task of using those cases as historical analogies to choose the right lessons and extrapolate the relevant trends for the next major continental war—without understanding in advance how those conflicts would be similar to, and different than, the Great War. Unfortunately, they were not confronted with evidence that was clear and consistent enough to disabuse them of their prevailing offensive vision of war.9

It is only in this ambiguous historical context that scholars can appreciate the institutional and cognitive limitations which biased European armies toward offensive, decisive engagements. For the most part, German military planners were not irrational; rather, they were confused by a variety of conflicting lessons about the future of warfare. When confronted by this environment, they had little choice but to remain tied to their offensive vision due to preexisting institutional and cognitive biases that had conditioned them to prefer the offensive as the best available means for achieving decisive results. Among the ‘lessons learned’ from the conflicts of the era, these institutional and cognitive biases led them to prioritize and ‘cherry-pick’ the lessons that were most in accordance with their offensive vision of warfare (i.e., the ‘wrong’ lessons, in hindsight)—while making only minor adjustments to accommodate the more important lessons (i.e., the ‘right’ lessons) about defensive firepower.10 This led to flawed assumptions, insufficient preparations, and unrealistic expectations regarding the prospect of decisive victory both on land (through offensive operations) and at sea (through major naval engagements)—thereby confounding the offensive Schlieffen Plan which was to serve as their war plan for a two-front war with France and Russia. And poor civil-military relations did not pressure the military to change these outdated views.11 Thus, a rational response to confusing ‘lessons learned’ produced irrational decisions.

Ultimately, through the thesis above, this paper makes a novel contribution to the innovation literature by: 1) analyzing World War I as a case of innovation failure, rather
than focusing on a successful case like the Interwar Period; 2) exploring innovation through the lens of historical ‘lessons learned,’ rather than solely through technological adoption; and 3) blending the political science focus on structural factors (i.e., systemic technological trends, institutional and cognitive biases, civil-military relations, etc.) with a historical focus on the state and context of military art and science prior to World War I, especially as individual decision-makers perceived them at the time (i.e., in the form of ‘lessons learned’). Accordingly, this paper argues that the structural factors identified by political scientists are not a primary cause, nor a sufficient cause, of the offensive bias—rather, they are still necessary causes, but are secondary to the presence of conflicting ‘lessons learned’ from the wars of the time. Indeed, it was the difficulty of grappling with these lessons that led German military planners to revert to their preconceived offensive biases. From this perspective, institutional and cognitive biases are not so much a hindrance to rationality as much as they are the only recourse for decision-makers struggling to decipher key trends.

The paper is organized as follows. First, in order to explore this era of industrial-age warfare and establish the context in which German military planners evaluated history, this paper begins with an operational analysis of the major lessons from three influential wars before World War I: 1) the Franco-Prussian War, 1870-71; 2) the Spanish-American War, 1898; and 3) the Russo-Japanese War, 1904-05. It would be ideal to include other conflicts like the American Civil War, the Boer War, and the Balkan Wars, which also shaped perceptions of military trends in the 50 years preceding the Great War; however, limited space prohibits this. Accordingly, the three case studies above were chosen using John Stuart Mill’s “Method of Agreement”—since they reflect wars in three different theaters (i.e., Europe, the Americas, and Asia, respectively) between three different dyads (i.e., France-Prussia, Spain-America, and Russia-Japan, respectively), any common ‘lessons learned’ among such different cases would be particularly strong and therefore apparent to pre-World War I military planners in Germany. Moreover, these three cases were studied by German military commanders to varying extents: 1) the Franco-Prussian War was examined extensively by both Moltke the Elder, who reflected on ‘lessons learned’ to continually adapt his war plans for a future great-power conflict, and Schlieffen, who used it as the basis for wargames, staff rides, and studies of tactical and strategic problems; 2) Schlieffen studied the Russo-Japanese War to explore the changing technological character of warfare and the consequences for Russian military readiness, but did not scrutinize the operational and tactical details to the same extent as the Franco-Prussian War; and 3) the Spanish-American War was largely neglected by senior leaders, but it was observed and reflected upon by more junior German naval officers in the Philippines. Second, the paper then analyzes why Germany failed to learn the ‘right’ lessons—first by comparing the ‘lessons learned’ (i.e., the ‘wrong’ lessons about the viability of the offensive) and the ‘lessons neglected’ (i.e., the ‘right’ lessons about defensive firepower); then by comparing the three case studies above to World War I in order to demonstrate why these lessons were so inconsistent; and finally by exploring the institutional, cognitive, and civil-military factors which biased German officers toward offensive doctrine within this context. While the paper seeks to provide insight into all of the European armies prior to World War I, it focuses on Germany specifically due to
limited space. Germany was chosen over France, Britain, Russia, and Austria-Hungary because Germany is often blamed for causing the war, and the political pressure to act exerted by the rigid mobilization timelines of the Schlieffen Plan is considered a contributing factor. But more importantly, it was Germany’s decisive success in the Franco-Prussian War that served as a model of modern warfare in the late 19th and early 20th centuries, which was adopted by the other European states. Therefore, it will be instructive to track how Germany perceived the evolution of warfare from the Franco-Prussian War to World War I. Third and lastly, the paper concludes with policy implications for the United States today in order to improve historical learning, mitigate institutional and cognitive biases, and foster innovation in preparation for the next great-power war.

Case Study #1: The Franco-Prussian War

The Franco-Prussian War of 1870-71 was the third and final German War of Unification, as Prussia used external war with France to rally and consolidate disparate German states into a new nation. Chancellor Otto von Bismarck and Chief of the General Staff Helmuth von Moltke identified the French army as the enemy’s Clausewitzian “center of gravity”—“the hub of all power and movement, on which everything depends”—and Moltke devised a relatively simple strategy to destroy it: converge with superior manpower against the French army, wherever it was. The sequence of military operations can be organized into three distinct phases. First, in July 1870, Moltke’s superior prewar mobilization plans allowed Germany to rapidly field a massive army and prevent a French spoiling attack across the Rhine. Second, from July to September 1870, Prussia conducted offensive operations and achieved decisive success against the French army. After a series of encounter battles at Spicheren, Froeschwiller, Colombey, and Vionville, the Prussians engaged one of the two French armies at the Battle of Gravelotte-St. Privat on August 18, 1870, and despite suffering heavy casualties, compelled the French to retreat to a fortress at Metz—to which the Prussian army laid siege until the French surrender in October. The second French army attempted to relieve the besieged forces at Metz but “was intercepted, driven against the Belgian frontier at Sedan, and forced to surrender on September 1”—representing the end of conventional combat operations. The final phase was the Prussian siege of Paris from September 1870 to January 1871. Despite Prussia’s overwhelming conventional victory, an insurgency emerged in the French provinces to resist the occupying Prussian forces, while the new Government of National Defense in Paris continued to hold out despite the siege. While Moltke sought to continue prosecuting a war of extermination and occupation, Bismarck wanted to end the war quickly due to deteriorating Prussian logistics and morale. Therefore, in order to compel surrender, Bismarck prevailed upon the army to execute a strategic bombardment of Paris—and while the bombardment itself did not secure victory, it motivated the Parisians to make one last desperate attack against their besiegers at Buzenval, which failed miserably and convinced the French populace of the need to surrender. Bismarck then negotiated a peace settlement with the new French government.

The first ‘lesson learned’ from this conflict was that superior preparation was essential to operational success on the modern battlefield. Moltke was able to execute a relatively simple strategy of concentration and envelopment because his prewar mobilization plans generated an army far
larger than that of France (740,000 vs. 350,000 troops) in a shorter period of time, thereby securing a first-mover advantage. By drawing troops from local districts, maintaining a corps organization in peacetime that would match wartime units and maneuvers, and leveraging the relatively new technology of railroads, “the Prussian general staff proved that it had mastered the problems of mass organization and movement.” This superior planning was only made possible by the revolutionary institution of the Prussian General Staff. As the “brains and nerve center of the army,” it sought to mitigate the Clausewitzian chance and uncertainty inherent in modern warfare by preparing mobilization timelines and initial concentration of forces far in advance of hostilities, since these elements could be calculated. The French army, which had not yet completed its own military reforms, was woefully unprepared.

This superior preparation enabled novel operational concepts, thereby creating new possibilities for the offensive in spite of the era’s increasing defensive advantages. Railroads and improved roads allowed armies to maneuver separately and then concentrate on the field of battle, rather than being massed days in advance, as in the Napoleonic era. Thus, concentric operations to converge upon the enemy using exterior lines of operation became just as advantageous as isolating separate enemy columns using interior lines, and the Prussian army could alternate between either of these operational approaches depending upon the circumstances—thereby enabling a flexible “strategy of expedients” rather than a strict ascription to Jominian principles. Moltke observed that time was now more important than position—as long as a commander can seize the initiative, it would be possible to exploit either interior or exterior lines to close upon an enemy quickly, strategically and operationally envelop their flanks at any distance, and nullify the advantages of their defensive position and technology. But this was only made possible by improved communications, primarily through telegraph, which enabled a decentralized command structure of mission command, or “Auftragstaktic.” Due to Clausewitzian chance, uncertainty, and probability on the battlefield, Moltke believed that “a dogmatic enforcement of the plan of operations was a deadly sin and great care was taken to encourage initiative on the part of all commanders, high or low.” Therefore, while the Prussian General Staff centralized control over the initial concentration and maneuver of forces in the Franco-Prussian War, they decentralized tactical control over individual army groups and formations. This decentralization yielded the initiative to Prussian field commanders, who were able to march to the sound of the guns and rapidly converge upon the enemy without waiting for orders. While this resulted in some friction—primarily through unnecessary encounter battles—it made a quick offensive envelopment of the enemy possible. The result was a conventional victory over French forces.

Of course, there were troubling ‘lessons learned’ as well. Defensive advantages were increasingly undeniable. Armed with the superior Chassepot rifle, the French inflicted severe casualties upon Prussian advances during the Battle of Gravelotte-St. Privat. One assault in particular produced 8,000 Prussian casualties within 20 minutes, and it was clear that dense Napoleonic waves of attacking infantry would no longer be possible—instead, dispersal into skirmish lines was necessary. But these lessons were undermined by the fact that the offensive was still possible. Moltke’s doctrine of converging upon and enveloping the enemy’s flanks allowed Prussians to mostly avoid deadly frontal attacks—and
this approach succeeded in driving the French from their fortified positions at Gravelotte-St. Privat, and surrounding and annihilating them at Sedan.\textsuperscript{34} And while some technological trends favored the defensive, others—like railroads, telegraph, and even artillery when used in advance of an attack—seemed to favor the offensive.\textsuperscript{35} More concerning, however, was the inability to achieve militarily decisive results, despite offensive operational success. Even after French conventional forces were beaten, the French national spirit was not. In the face of a French insurgency that threatened to break Prussia’s siege of Paris, it was clear that Prussia had reached its “culminating point of victory.”\textsuperscript{36} The fearsome Prussian army was suddenly demoralized, insufficiently supplied, and vulnerable—and if Moltke was allowed to continue prosecuting his war of extermination, ultimate victory may have eluded the Prussians. This situation on the ground was exacerbated by France’s control of the sea, which allowed them to import supplies from abroad and sustain the insurgency.\textsuperscript{37} Only Bismarck’s compellent approach and adroit diplomacy produced a decisive political victory before the military situation became worse. Thus, it was clear that decisive victory necessitated connecting military success to political success.\textsuperscript{38} In an era of rising nationalism and the foreboding prospect of absolute and existential war, it was increasingly apparent that enemy armies were not the true center of gravity—rather, it was the enemy’s national will, which would be considerably harder to break.

Unfortunately, however, the subsequent cases demonstrate that this lesson about the difficulty of decisive victory was neglected.

Case Study #2: The Spanish-American War

The Spanish-American War of 1898 was a major inflection point in the history of American empire, yielding colonies and great-power status to the rising United States. It was a global conflict, fought in two theaters: 1) the Atlantic Ocean, particularly in Cuba and Puerto Rico; and 2) the Pacific Ocean, particularly in the Philippines. U.S. planners identified the Spanish fleets in these respective oceans as the enemy centers of gravity since they were the primary means by which Spain effected its remaining imperial control and reinforced its colonial armies. Therefore, the United States sought to use its better-modernized and equipped navy as the primary power projection force to pin down or destroy the two Spanish fleets, thereby establishing command of the sea and enabling U.S. invasion forces to achieve victory in Cuba and the Philippines, in tandem with Cuban and Filipino rebels. The war proceeded in three phases. First, while the U.S. army struggled to mobilize, the U.S. navy blockaded the northern coast of Cuba in April 1898, and Commodore George Dewey’s U.S. Asiatic Fleet decisively destroyed Spain’s fleet at Manila Bay in the Philippines on May 1.\textsuperscript{39} This early naval success accelerated the mobilization timeline for the army, and although it was unprepared, it entered the second phase: difficult joint operations in Cuba from May to July 1898.\textsuperscript{40} With Spanish naval forces pinned down at Santiago de Cuba, the U.S. army landed at Daiquirí and attempted to lay siege to Santiago, but suffered significant casualties in the Battles of San Juan Hill and El Caney on July 1. However, victories in these battles laid the foundation for the third and final phase: strategic success from July to December 1898. The Spanish fleet was destroyed as it tried to break out from Santiago; Puerto Rico was invaded in “a
nearly bloodless campaign of rapid maneuver”；and the Spanish garrison in the Philippines was compelled to surrender.\textsuperscript{41}

Despite American victory, there were some painful ‘lessons learned’ from this conflict. In contrast to Prussia’s ruthless efficiency and first-mover advantage in the previous case, the United States was unprepared for modern warfare. The army’s readiness was atrocious—upon declaration of war in April 1898, there were only 28,000 regular soldiers, and it required months to mobilize, train, and equip a volunteer force that would eventually number 290,000.\textsuperscript{42} The navy was better prepared due to recent modernization efforts by Assistant Secretary of the Navy Theodore Roosevelt and his predecessors, but it lacked the means to implement a full blockade of Cuba, thereby creating gaps.\textsuperscript{43} More importantly, this case demonstrated the necessity of a general staff by its absence. The United States suffered from a poor chain of command, an inability to coordinate joint operations and combined arms, and a lack of an institutional basis for resolving disputes among military services and commanders.\textsuperscript{44} Consequently, the army and navy often implemented their own respective war plans, failed to immediately exploit naval success in the Philippines, and lacked logistical preparations for Cuba.\textsuperscript{45} Most egregiously, in the absence of a general staff, interservice and personal rivalry between General William Shafter and Rear Admiral William Sampson resulted in an inability to decide upon a military aim in Cuba (i.e., destruction of the Spanish fleet at Santiago, or the siege and occupation of Santiago itself).\textsuperscript{46} Fortunately, the United States was so much stronger than Spain that these weaknesses did not produce defeat. However, they impeded success and even opened opportunities for potential Spanish counterattack, especially if Spain had tried to spoil the amphibious landings at Daiquiri.\textsuperscript{47}

At sea, the Spanish-American War demonstrated the viability of theorist Alfred Thayer Mahan’s views on naval warfare. As a result of extensive modernization, the United States fielded fleets of faster and better-armored battleships and steel-plated cruisers.\textsuperscript{48} In accordance with Mahanian strategy, the United States prioritized destroying Spanish fleets in order to establish sea control and economically strangle the enemy, while leveraging concentration, interior lines of operation, and naval logistical preparations like coaling stations to achieve success.\textsuperscript{49} The ability to outrange and outgun the Spanish fleets resulted in decisive victories at Manila Bay and Santiago.\textsuperscript{50} Despite this success, the lack of amphibious doctrine and coordination between the army and navy nearly ruined a strategy that demanded joint operations.\textsuperscript{51}

On land, the results indicated strong advantages for the defender, which had increased since the Franco-Prussian War.\textsuperscript{52} At the Battles of San Juan Hill and El Caney, artillery and rifle fire by entrenched Spanish positions inflicted severe casualties upon the advancing Americans—at El Caney, “American commanders were at a loss to explain how 520 Spaniards could resist 5,400 Americans for more than nine hours.”\textsuperscript{53} While U.S. black powder ammunition made its cannons an easy target, Spanish smokeless powder heavily impeded U.S. counterbattery fire. And U.S. deficiencies in artillery preparation and combined arms coordination—especially at San Juan Hill, where advancing Americans were struck down by friendly artillery fire—served as an ominous warning for the campaigns of 1914.\textsuperscript{54} These trends almost precipitated a failure to achieve decisive breakthrough at the San Juan Heights.

However, like the Franco-Prussian War, the Spanish-American War demonstrated that the offensive was still viable. The
Americans had been conditioned since the Union’s 1864-1865 campaigns in the Civil War to believe that sheer numerical superiority and morale could prevail against defensive positions—and these factors helped them persist at the San Juan Heights. Moreover, once artillery was prepared and firepower superiority was achieved, offensives in Cuba, Puerto Rico, and the Philippines were far more successful. Certain technologies were also viewed as equally offensive in nature as they were defensive—ironically, Gatling guns (a precursor to World War I-era machine guns that were entirely defensive) ensured success for the U.S. advance on San Juan Hill. And lastly, for units that decentralized command authority, this war demonstrated the effectiveness of American “extended order” (based upon Prussian/German mission command, or Auftragstaktik) to seek tactical advantages and avoid frontal assaults.

Case Study #3: The Russo-Japanese War

The Russo-Japanese War of 1904-05 was a struggle for imperial supremacy in East Asia between a declining Russia and a rapidly modernizing Japan. Japanese forces had equivalent technology and better training due to their ties with Germany and Britain, but Russia was still a daunting power with superior resources and manpower reserves. Therefore, Japanese military planners like Marshal Oyama Iwao, General Kodama Gentaro, and Admiral Togo Hiroheichi decided to establish sea control early in the conflict and then land an invasion force to strike at the Russian army in Manchuria, based at Liaoyang. Given that Russian forces were operating far from their capital and from the majority of Russian forces in Europe, their Manchurian army was likely their center of gravity in the East, and its defeat would ideally compel Russian policymakers to decide that sending a new army to the East would not be worth the cost. The war proceeded in three phases. First, a surprise Japanese naval strike on Port Arthur in February 1904 divided Russia’s Port Arthur and Vladivostok fleets, thereby establishing command of the sea—while unimpeded landings of Japanese troops at Inchon and a major land victory at the Battle of the Yalu in April 1904 helped seize the initiative and provided an early psychological advantage. However, the next phase involved greater difficulty as Japanese forces sought to eliminate Port Arthur as a threat and simultaneously advance to Liaoyang. The loss of Japanese battleships in May and the failure to break through the Port Arthur defenses in August threatened to delay operations. Meanwhile, the Battles of the Mo-tien Pass, Makura Yama, and Liaoyang from July through September were tactical victories for Japan, but they incurred heavy casualties due to modern firepower and failed to achieve decisive strategic victory due to Russia’s ability to replace losses. Fortunately for Japan, the final phase demonstrated success. The eventual collapse of Port Arthur and the decisive victories on land at Mukden and at sea at Tsushima in early 1905 destroyed Russian morale, leading to revolt at home and an eventual peace agreement.

During the Russo-Japanese War, defensive firepower demonstrated its fullest potential. Indeed, at the Nanshan, at the Mo-tien Pass, and especially during the ground assault on Port Arthur, the “shrapnel fire [of artillery] as well as infantry rifle-fire made any movement within sight and range of the enemy out of the question, and put an end to all idea of close formations maneuvering on the battlefield.” Indirect fire in particular was revolutionary. As a result, Japan’s advance on land was costly, bloody, and time-consuming. However, consistent with the preceding case studies, Japanese
advances prevailed in almost every battle. Like Moltke and the Prussian army, Japan constantly seized the initiative through aggressive offensives, thereby setting the tempo of operations, fighting on ground of its own choosing, paralyzing Russian decision-making and morale, and preventing Russian commanders from leveraging their superior numbers. Thus, despite having the same technology and inferior numbers, Japan made the offensive possible—primarily through a variety of tactical innovations and inherent advantages. First, artillery preparation allowed the Japanese to achieve firepower superiority prior to advancing. Second, during engagements like the initial naval strike on Port Arthur and the Battles of the Yalu, Nanshan, and Sha Ho, Japanese commanders relied on deception (e.g., through feints) and tactical surprise (e.g., through night attacks) to circumvent the defensive advantages of modern firepower. Finally, superior Japanese morale—even “fanatical courage”—ensured perseverance in the face of heavy casualties.

However, despite the incredible extent of these tactical victories, decisive strategic victory remained elusive. After the Battle of Mukden, the Japanese had reached their culminating point of victory. The Russian army was badly beaten and in retreat, but not destroyed; it continued to replace its losses. Japan lacked the manpower and supplies to pursue it—and despite the fact that sea control was supposed to be an operational enabler for the ground offensive against the enemy center of gravity in Manchuria, command of the sea did little to ameliorate this poor logistical situation. Mukden was not nearly as ‘decisive’ as is popularly believed—the prospect of a much longer war, which Japan could not sustain, reared its ugly head. In what amounted to be a prescient glimpse into the grim future of World War I, the tactical defeat and even destruction of enemy armies did not produce victory as long as the enemy’s national will persisted. In such an era, no single army could be the center of gravity. Decisive political victory for Japan only came when Moscow started feeling the domestic pressure of the Revolution of 1905, and when the United States pressured both belligerents to cease hostilities.

Like the Spanish-American War, the Russo-Japanese War demonstrated the importance of Mahanian theories about sea control. Command of the sea ensured a continuous offensive campaign on land, without fear of counterattacks against the Japanese homeland. Additionally, Admiral Togo’s ability to both concentrate his fleet and leverage interior lines of operation against an overstretched opponent allowed him to divide and individually destroy Russia’s Port Arthur and Vladivostok fleets early in the war, and eventually the 2nd Squadron of the Russian Pacific Fleet, which had taken months to arrive from the Baltic Sea only to be destroyed at the Battle of Tsushima. However, there were also troubling lessons. Battleships—the capital ships so prized by Mahan and other navalists—were exceedingly vulnerable to mines and torpedoes. Moreover, despite the Japanese attempt to coordinate navy and army strategies, command of the sea was not sufficient to mitigate the aforementioned logistical difficulties on land. And while the Battle of Tsushima was widely portrayed as a decisive naval engagement like the Battle of Trafalgar a century before it, it was not very ‘decisive’ after all—it was a victory over inferior Russian ships, and it was not as important in ensuring Russian capitulation as the domestic Revolution of 1905.
Germany: Lessons Learned and Neglected Prior to World War I

‘Right’ and ‘Wrong’ Lessons

Why did the German General Staff learn the ‘wrong’ lessons from the wars of the late 19th and early 20th centuries? As demonstrated by the three cases above, there were no ‘right’ or ‘wrong’ lessons at the time. There were consistent lessons across all three cases, but ironically the lessons of each case were not logically consistent with each other—namely, while defensive firepower was increasingly destructive, the victors were all able to use offensive operations to achieve major victories. In every case, the trends that supported a narrative of defensive warfare were juxtaposed with the trends that supported a narrative of offensive warfare—and it was not clear whether warfare was truly defensive, while offensive operations remained an exception to the norm (i.e., the ‘right’ but neglected interpretation); or whether warfare was still truly offensive, while defensive firepower was a slight modification to traditional views (i.e., the ‘wrong’ but accepted interpretation at the time). ‘Right’ and ‘wrong’ lessons can only be determined according to how applicable they eventually would be to the attrition of World War I.

In hindsight, there are several significant lessons from the three case studies that were very applicable to World War I—these lessons were ‘right,’ but they were neglected at the time. First, modern artillery, rifles, and machine guns were rendering offensive advances, especially over relatively open terrain, increasingly untenable. And most importantly, these trends were becoming stronger with every successive case. Indeed, there were two waves of defensive technology in the 19th century. The first wave in the middle of the century yielded “mass-produced, rifled, breech-loading firearms” and incrementally better artillery, resulting in increased use of defensive trenches and offensive skirmish lines in the American Civil War and the Franco-Prussian War. However, at the turn of the century, the advent of new technologies like indirect, recoiling, and quick-firing artillery, along with smokeless powder, high explosive rounds, and machine guns, resulted in a more decisive defensive advantage. This was evident in the increasing volume, range, and accuracy of fire in the Spanish-American and Russo-Japanese Wars. These realities lead to a second neglected lesson: while operational and tactical offensives were not necessarily dead (as evident by the case studies), they necessitated far greater care. Indeed, European powers prior to World War I should have better emulated the Japanese by adopting certain tactical measures, such as: artillery preparations for firepower superiority, along with better-combined arms coordination between infantry and artillery; dispersal of infantry into skirmish lines; infiltration and siege tactics rather than frontal assaults; and decentralized mission command to allow individual commanders the opportunity to exploit potential breakthroughs. But all of these potential adjustments point to a third, more pessimistic lesson: while the offensive might have remained possible, it was unlikely to be decisive in and of itself. Given nationalist sentiment and trends toward total war, even if a state’s armies were destroyed, “a resolute government with untapped resources at its disposal normally could raise other forces and continue to fight.” Annihilation by offensive means was no longer possible—limited wars combining offensive and defensive tactics were likely a better approach to achieving political victory.

In the 50 years preceding World War I, the Chiefs of the German General Staff—
Moltke the Elder, Schlieffen, and Moltke the Younger—largely respected these trends. Indeed, Schlieffen believed that the Russo-Japanese War portended “more instances of trench warfare” in the future, where mass attacks and numerical superiority would be insufficient “to overcome the murderous fire of an entrenched enemy.” Like their contemporaries in other European armies, German strategists “were neither blind to the likely consequences of their attacks nor ill-informed about the defensive powers of twentieth-century weapons.” However, in evaluating these case studies, they consistently found proof “that an attack against an enemy’s front can, despite all difficulties, succeed very well.” Therefore, their primary ‘lessons learned’ from the wars of the time were various mechanisms to keep the offensive not only viable, but *decisive* in the face of defensive advantages. They believed that offensive initiative could prevail under three conditions. First, it had to be fast enough at the strategic and operational levels to envelop enemies before they had time to build field fortifications and trenches, thereby corresponding to Moltke’s prioritization of time over position. This lesson would imply a preference for both rapid mobilization, as in the Franco-Prussian War, and control over time itself on the battlefield. For the latter, Moltke preferred a flexible “strategy of expedients,” by which decentralized mission command would allow officers to quickly strike at enemy weak points in a series of encounter battles; while Schlieffen rejected mission command and preferred set-piece battles, or “manoeuvre a priori,” to establish strict and rapid timelines for attacking the enemy. The second condition—conceived by Schlieffen in response to the increasing army sizes, dispersal of attacking infantry, and the advent of multi-day battles evident in the Franco-Prussian and Russo-Japanese Wars—was the expansion of the battle front into a ‘complete battle’ (*Gesamtschlacht*) across the entire theater, where frontal assaults could be avoided and the enemy could be enveloped at various weak points. And the third condition for decisive offensives was that they were well-coordinated and motivated at the tactical level. Given that the Prussians had defeated the better-equipped French, the Americans had defeated the better-equipped Spanish, and the Japanese had defeated the more numerous Russians, the German General Staff believed that the enemy’s defensive positions and weapons could be overcome by a variety of evolutionary (rather than revolutionary) measures and adjustments: continuous maneuver to envelop the enemy; artillery preparation to achieve firepower superiority; dispersal of advancing forces; and most importantly, vigorous morale. The final ‘lesson learned’—which was especially important for a new naval power like Germany—was that sea power could be decisive in itself, especially in the construction of empire. Kaiser Wilhelm II strongly believed in Mahanian ideas about sea control, and his personal investment in naval modernization as a key tenet of Germany’s *Weltpolitik* imbued the navy with greater prestige than it otherwise would have enjoyed in a more conservative and aristocratic land power like Germany.

Therefore, Germany counted on the offensive Schlieffen Plan (subsequently modified by Moltke the Younger) to achieve success by envelopment, using many of the same concepts that Moltke the Elder had refined to win the Wars of German Unification—while simultaneously leveraging a modernized navy to prevail at sea. This vision of warfare was a failure, not so much because it relied upon viable offensives—indeed, as the Eastern Front of World War I would demonstrate, offensives and maneuvers were still possible—but rather because the vision bet everything
upon the probability of the offensive achieving decisive victory on land and sea, despite trends that were proving just the opposite.

*Imperfect Cases and Conflicting Conclusions*

Thus, the German General Staff had chosen the ‘wrong’ lessons learned. This outcome still seems surprising, as the General Staff was an institution that prided itself on historical learning. This section will explore why the learning environment was so confusing by answering the following questions: why were the ‘lessons learned’ from previous cases so inconsistent in their applicability to World War I; and why did they leave room for misinterpretation? This section aims to disprove the accounts that blame German (and European) military officials for their lack of recognition of supposedly ‘obvious’ defensive trends.

While the three cases in this paper shared similarities with the eventual Great War, there were some key differences that explain why offensives were more viable in the discussed cases than in World War I. First, the three case studies all exhibited unequal distributions of power and skill. The Prussian, American, and Japanese victors were able to succeed against the defensive advantages of their enemies, in part, because they were stronger and more competent. Indeed, French military readiness was extremely poor and they relied on a warrior ethos rather than operational and logistical planning; Spain was not sufficiently powerful or predisposed to seize the initiative and spoil slowly developing U.S. operations; and Russian forces were so incompetent at every stage of the conflict that a full list of their operational and tactical mistakes would merit an extensive research project in its own right. In World War I, by contrast, many of the belligerents were more equally matched—they had all learned the importance of mobilization and readiness, and they had significant resources at their disposal. Therefore, offensives could no longer prevail based on strength and competence alone.

Second, the three case studies were mostly limited wars. In the Spanish-American and Russo-Japanese Wars in particular, the vanquished belligerents did not give up because they could no longer fight, but rather because it was no longer worth the struggle. And even the Franco-Prussian War, which was relatively existential from the French perspective, was still limited by Bismarck’s desire to achieve rapid victory and moderate territorial and security concessions. Therefore, the political decisiveness of these limited wars masked a dearth of military decisiveness—battles like Sedan, Manila Bay, San Juan Hill, Mukden, and Tsushima were only militarily ‘decisive’ in hindsight. Like blitzkrieg in World War II, offensive operations in these wars were most “[e]ffective against poorly prepared and often poorly commanded adversaries and within a limited theater of operations… [and] could not be sustained over longer distances or bring final victory against an enemy who could trade space for time and disposed of ample reserves.” The enemy’s military power was not yet broken—they just gave up. This surrender would not be the case in an existential, total war like World War I, where national will had supplanted enemy armies as the true center of gravity. Thus, even if the Schlieffen Plan had worked operationally, it was unlikely to achieve strategically and politically decisive effects.

Third, the terrain was different. For the Spanish-American and Russo-Japanese Wars, the theaters were spread out over greater distances, complicating the ability of a more distant country (Spain and Russia, respectively) to replace losses, thereby
allowing the United States and Japan to achieve more ‘decisive’ results in a short period of time. But on the Western Front of World War I, the terrain was so narrow and troop density was so high that frontal assaults could not be avoided, complicating Moltke’s and Schlieffen’s planned offensive envelopment strategy in the absence of vulnerable flanks to attack. Additionally, because the Spanish-American and Russo-Japanese Wars were predominantly fought on islands and peninsulas in the Atlantic and Pacific Oceans, naval warfare was incredibly important—indeed, navies were essential for transporting and supplying troops, thereby serving as an operational enabler to ground forces fighting far away from their capitals and logistical bases. Sealift capabilities ensured that warfare was more mobile, and flanks on land were more vulnerable to amphibious landings.

However, in a European land war like World War I, operational maneuvers and critical supply lines would not be as dependent on command of the sea. While economic strangulation through a naval blockade could be effective (and would eventually play a critical role in compelling Germany to surrender in 1918), it would require a significant amount of time—far more than that anticipated by Mahan and the other advocates of naval ‘decisiveness’—because essential supplies were already accumulated in the theater. Unlike the Spanish-American and Russo-Japanese Wars, a land war in Europe would not feature a rapid and decisive role for European navies—and while naval action could be helpful, it would need to be tied to ground operations through better joint coordination, which did not exist between the German army and navy prior to World War I.

Lastly, while these cases portended technological trends that would feature heavily in World War I, these trends had not yet reached their zenith. The technology of the time was continuously changing, and “[i]t was hotly contested, among military experts, whether these developments on the whole favored the attack or the defense.”

For example, improvements in artillery armed the defender with devastating firepower, but effective artillery preparation by advancing forces like the Japanese allowed them to prevail against Russian defensive positions; while machine guns would eventually cut down troops crossing ‘No Man’s Land’ in World War I, their predecessor—the rapid-fire Gatling gun—enabled advancing U.S. forces to lay suppressive fire against the Spanish position at the San Juan Heights; and the same railroads and telegraphic communications that enabled the reinforcement of defensive, interior lines in the Great War were used to enable offensive operations by the Prussians and Japanese in their earlier conflicts. As explained by Robert Jervis’s concept of “offense-defense differentiation,” it was difficult to distinguish the inherently defensive or offensive character of these weapons. Thus, offensives were still somewhat possible in those earlier case studies—and some technologies were viewed as equally viable for offensive or defensive purposes—because the defensive technologies would continue to grow stronger up until 1914, and their true defensive potential would only become ineluctably clear during the Great War.

All of these differences would eventually magnify the advantages of defensive warfare in World War I, but they are only apparent in hindsight. At the time, it would have been nearly impossible for German military commanders to foresee how the multiple cases available for study were different from the coming Great War—and therefore just as impossible to know which lessons would be most applicable to the future of warfare, and which should be neglected.
Institutional and Cognitive Biases and the Failures of Civil-Military Relations

The absence of clear ‘lessons learned’ meant that military commanders had no choice but to rely on their preconceptions and experiences to interpret the ambiguous world around them. Due to institutional and cognitive biases toward the offensive, these leaders prioritized and ‘cherry-picked’ the lessons that best supported their existing vision of warfare, resulting in both rational and irrational choices.

The literature on institutions reveals that organizations tend to routinize tasks, promote standard operating procedures, and cultivate a specific organizational culture in order to maximize efficiency in the approach to different challenges. Military institutions in particular are often biased toward offensive doctrine because it allows them to preempt ‘inevitable’ threats, establish a structured environment by determining the time and place to fight, and secure greater organizational autonomy over the conduct of military affairs. These insights certainly apply to the German General Staff prior to World War I. Indeed, Moltke the Elder ensured standardization so that junior officers could be trusted with decentralized mission command. Moreover, German officers were strongly biased toward offensive doctrine, decisive victory, and wars of annihilation (Vernichtungskrieg, enabled by prostration strategies, or Niederwerfungsstrategie) because they had worked, and worked well, in the General Staff’s formative experiences of the Wars of Unification—particularly the Franco-Prussian War. This vision persisted as theorists and historians like Carl von Clausewitz and Hans Delbrück cautioned that limited wars (Ermattungsstrategie) and defensive operations bore certain advantages. And, more importantly, it persisted even in the face of contradictory evidence from the Russo-Japanese War, which demonstrated the difficulty of achieving decisive results. Due to the institutional bias toward wars of annihilation, Schlieffen justified this contradiction by arguing that the Japanese were successful in achieving minor breakthroughs due to their offensive approach, but they had failed to achieve decisive results because they did not go far enough (as Moltke had) in annihilating the enemy along all fronts. Therefore, as offensives continued to succeed in the wars of the time, German military planners were not confronted with sufficient evidence to substantially alter their vision of future warfare. Instead, while Moltke the Elder and Schlieffen recognized the defensive trends that complicated offensive doctrine, they undertook minor, evolutionary changes—for example, by making “existing forces more effective”—to accommodate and overcome these defensive trends, rather than revolutionary changes to fight in an altogether different way. Such changes in the name of efficiency are often embraced by organizations, while true reform that runs against the grain of institutional biases is significantly more difficult. Thus, Schlieffen and Moltke the Younger certainly improved the prewar army by integrating emerging technologies and concepts, including heavier and indirect artillery, machine guns, mortars, grenades, modern signal equipment, and motorized vehicles. But they were only adopted in the name of continuing the viability of the offensive—they were ‘cherry-picked.’ Without a new vision for the revolution underway in warfare (which would have required a pairing of offensive and defensive warfare, and an acceptance of limited war rather than decisive battles of annihilation), the potential of these new technologies could never be fully realized. This myopia greatly impeded the modern development of combined arms. In 1914, for
example, advancing German soldiers often neglected to leverage machine guns, mortars, and grenades because they seemed cumbersome to their advance, and they were not well trained in their use. They were not at all prepared for the defensive warfare that would define World War I.

German naval doctrine was also shaped by institutional biases. As opposed to the longstanding tradition and ethos of the Prussian (and subsequently German) army, the German navy was young, nationalist, and politically supported by the Mahan-inclined Kaiser. Given this institutional threat to the army, as well as the young navy’s failure to achieve decisive victories in the Wars of Unification, the German General Staff did not establish a role for the navy in the Schlieffen Plan. In order to justify its existence and vie for resources (in accordance with Graham Allison and Philip Zelikow’s model of bureaucratic politics), the navy was bureaucratically biased toward Mahanian ideas about the importance of sea control. Thus, they cherry-picked historical lessons about the decisiveness of sea power and procured powerful battleships, even though navalists continually struggled to define what ‘decisiveness’ meant at sea; there had been only a handful of truly decisive naval battles like Trafalgar and (questionably) Tsushima over the preceding century of warfare; and, even if naval decisiveness was clearly defined and theoretically attainable, by 1914 Germany could not hope to decisively defeat the superior British navy. This bias toward decisive naval victory led Admiral Alfred von Tirpitz and the navy to neglect a better (and more prescient) joint strategy, originally advocated by Chancellor Leo von Caprivi and Admiral Friedrich von Hollman, which would have complemented the Schlieffen Plan on land by protecting the German coastline from blockade—thereby employing the navy in a secondary, economy-of-force role to prevent the kind of economic strangulation that eventually contributed to German collapse in 1918. Instead, Germany built a warfighting navy that was ill-suited to the land-based terrain and relevant centers of gravity in the European theater, while joint planning remained essentially nonexistent. Thus, due to bureaucratic politics and organizational biases, the navy’s conception of decisive victory was divorced from Germany’s overall military strategy and operations.

Within this institutional context, there were also cognitive biases that predisposed individual army and navy officers toward the offensive and decisive victory, resulting in unrealistic war-planning assumptions. According to the literature on cognitive biases, “decision-makers are apt to err by being too wedded to the established view and too closed to new information”—especially when the available data offers ambiguous insights and when the individuals are very confident about their theories. The literature on prospect theory also posits that individuals “tend to overweight losses with respect to comparable gains” and thus are “risk-averse with respect to gains and risk-acceptant with respect to losses.” Upon applying this literature to the German General Staff, it is evident that German commanders were in the ‘worst of all worlds’ in terms of their cognitive ability to see alternative futures and test their prevailing theories. Indeed, these officers were confounded by ambiguous and inconsistent ‘lessons learned’; shared institutional confidence in the continued viability of decisive victory through offensive doctrine; and were confronted by a geopolitical environment that increasingly threatened the prospect of strategic encirclement by the Triple Entente, thereby rendering their plans for a two-front war less tenable while prompting fears of military and geopolitical loss. This troubling
situation was compounded by the professionalism of the German military, by which officers were trained to accept that they must make the best of a bad geopolitical situation—rather than exercise a recourse to diplomatic or other means which could have improved the threat environment without resorting to war.\textsuperscript{109}

Since German commanders were not presented with sufficient evidence to change their offensive views, they were compelled to double down on them in order to mitigate uncertainty and avoid potential defeat. Thus, the cognitive flexibility of very skilled German commanders deteriorated over time.

For example, in extrapolating from the Franco-Prussian War to plan for a future great-power conflict, Moltke the Elder originally grappled with a more appropriate defensive-offensive war strategy that would begin with an advance into favorable French territory to spoil their mobilization plans, and then force the French to counterattack against newly established German defensive positions. This strategy combined the advantages of both offensive initiative and defensive firepower while seeking limited war aims rather than total victory.\textsuperscript{110}

Moreover, Schlieffen’s study of the Russo-Japanese War and the increasing likelihood of trench warfare led him to assume a strategically defensive approach during his last Kriegsspiel wargame in 1905.\textsuperscript{111}

However, the growing military power of, and entente diplomacy between, Germany’s French and Russian adversaries created a more dire geostrategic situation, and the only hope for winning such a two-front war appeared to require the rapid defeat of one of the belligerents.\textsuperscript{112} Consequently, in such an uncertain environment, the resulting Schlieffen Plan began to rest on increasingly unrealistic assumptions, expectations, and hopes in order to remain viable—no longer respecting the endogenous friction and exogenous chance, probability, and uncertainty inherent in war.\textsuperscript{113} For example, Schlieffen expected that the British Expeditionary Force could be easily defeated, that Belgian and French railroads would be intact for the offensive, that an \textit{ad hoc} logistical system could prevail, that France would not redeploy to counteract the envelopment, and that the racial ‘superiority’ of the German nation would ensure greater morale in the face of firepower.\textsuperscript{114} Perhaps most outlandishly, Moltke the Elder and Schlieffen assumed they could control time itself, either through rapid maneuver enabled by mission command or the set-piece battles of the \textit{manoeuvre a priori}, respectively.\textsuperscript{115}

This wishful thinking was perpetuated by a selection bias in the analysis of ‘relevant’ wars. Schlieffen studied the campaigns of Hannibal, Frederick the Great, Napoleon, and Moltke the Elder (e.g., the Franco-Prussian War) more than he studied the Spanish-American or Russo-Japanese Wars, likely for two reasons related to cognitive bias: 1) older campaigns demonstrated spectacular offensive victories, representing confirmation bias; and 2) Schlieffen preferred cases that he had studied or experienced in earlier stages of his life over more recent wars, representing the “availability heuristic.”\textsuperscript{116} Thus, the criteria for case selection represented attempts to derive certainty from an uncertain learning environment, and avoided the more troubling and mutually reinforcing lessons of “bothersome” analogies like the Spanish-American and Russo-Japanese Wars.\textsuperscript{117}

Schlieffen was so predisposed toward his certainty in the offensive that, even when he picked better cases to study, he failed to learn the most appropriate lessons. For example, rather than interpreting the increasing difficulty of offensive operations in the Franco-Prussian or Russo-Japanese Wars as a harbinger of future defensive warfare, he blamed such difficulties on
Moltke’s decentralized mission command in the former (even though it had actually helped); and on Japan’s failure to envelop in the latter (even though Japan constantly enveloped Russian forces). Thus, Schlieffen respected some of the key lessons from the modern case studies, but he saw them as either defensive modifications or aberrations to the truly offensive character of warfare.

Admiral Tirpitz and naval leaders were also hindered by cognitive biases. For example, navalists believed in the prospect of decisive victory at sea but failed to define what ‘decisiveness’ meant at the grand strategic, strategic, and tactical levels of war. As a result, the primary object of naval action was assumed to be the destruction of the enemy’s fleet, often without detailed consideration of how that outcome would affect war on land. Consequently, despite the supposedly central role of sea power, naval doctrine, and war planning were primitive compared to such processes in the army. Moreover, like their army counterparts, naval strategists harbored unrealistic expectations about the true extent of their naval capabilities—indeed, they expected to achieve decisive victories against the British navy even though, in 1914, German naval forces were outnumbered and dispersed around the globe, complicating the concentration of naval forces necessary for Mahanian sea control (demonstrated by U.S. and Japanese naval operations in the Spanish-American and Russo-Japanese Wars, respectively). Additionally, the German naval officers strongly believed in the power of the battleship, but the Russo-Japanese War had demonstrated that mines and torpedoes could offset battleships at a far lower cost.

As argued by political scientist Jack Snyder, such offensive biases are often exacerbated by weak civil-military relations. In the Franco-Prussian War, when the French insurgency stretched German forces to their culminating point, Bismarck was able to provide a check on Moltke’s bias toward decisive victory, and he leveraged coercive diplomacy to secure a negotiated peace. However, there was no Bismarckian figure to constrain Schlieffen or Moltke the Younger. By 1914, the General Staff had become more powerful in its proximity to the Kaiser and its control over the military lever of statecraft, and Chancellor Theobold von Bethmann Hollweg—ascribing to what scholars would now term Huntingtonian norms of civil-military relations—refused to stand in the way of its rigid offensive doctrine. Indeed, “[w]hat was special about the period before World War I was that the state of civil-military relations in each of the major powers tended to exacerbate that normal offensive bias” of the armed forces, and in the case of Germany in particular, it was “because the lack of civilian control allowed it to grow unchecked.” Thus, there were no impediments to the implementation of the Schlieffen Plan.

Conclusion

German military innovation prior to World War I did not fail because of irrational officers, insufficient technological adoption, or faceless structural processes that produced an inevitable result. Rather, it was a failure of vision, caused primarily by a learning environment in which historical case studies offered only inconsistent ‘lessons learned.’ Thus, institutional and cognitive biases toward the offensive became the only available means with which officers could interpret evolving military trends with relative certainty, leading them to bet everything on the rapid decisiveness of offensive military operations on land and major naval engagements at sea. In the face
of such an uncertain environment, and without sufficient evidence in any direction to discard their prevailing visions of warfare, these highly capable leaders doubled down on their offensive biases—while slightly accommodating what they regarded as evolutionary (rather than revolutionary) defensive trends through relatively marginal technological and doctrinal adjustments. Thus, rational decision-making broke down and yielded the irrational assumptions inherent in the Schlieffen Plan.

Based on the preceding analysis, this paper offers several policy recommendations for U.S. defense officials to prevent similar results in the future. First, contrary to popular expectations, peacetime innovation is not solely, or even mostly, about technological adoption—rather, it is about vision. New technologies like artificial intelligence will provide only incremental improvements unless DoD develops a new vision for employing these capabilities through greater innovation in operational concepts, tactics, and organizational reform. Second, since this vision is informed by the clarity or ambiguity of historical and modern ‘lessons learned,’ DoD needs to provide military officers with a stronger basis for evaluating such lessons. This training will require pairing the study of military history with other disciplines like technology forecasting and political science, which can offer insights into extrapolating future trends. Such approaches could be particularly insightful when applied to the Second Nagorno-Karabakh War and the Russo-Ukrainian War, which may be laboratories for the future of warfare. Third, DoD must address the institutional and cognitive biases that inhibit the adaptability of officers’ visions by rotating these officers among new private and non-government institutions, thereby imbuing them with diverse experiences. Fourth and finally, the Joint Staff must empower younger officers to have a voice in major strategic debates in order to avoid the crystallization of views among a small cadre of senior officers. One potential solution is to establish new roles for junior officers that provide greater input into strategic and operational decision-making. Through these recommendations, the United States can chart a path toward more successful historical learning and therefore more transformative innovation prior to the next major conflict.

About the Author: Christian Trotti is an M.A. candidate and Graduate Teaching Assistant with Georgetown University’s Security Studies Program. His research focuses on defense strategy, military operations, and technological and organizational innovation. He was formerly an Assistant Director of Forward Defense at the Atlantic Council, where he led its defense modernization portfolio.


11 Snyder, 108-46.


16 For more on the diffusion of military innovations to other states, see Michael Horowitz and Shira Pindyck, “What Is Military Innovation and Why It Matters,” *Journal of Strategic Studies* (March 2022).
18 Rothenberg, 303.
20 Rothenberg, 304.
22 Ibid., 348-373.
23 Ibid., 303.
25 Holborn, 283.
27 Holborn, 288.
28 Rothenberg, 300; Holborn, 290.
29 Holborn, 293-294; Rothenberg, 296.
30 Rothenberg, 296; Holborn, 291.
31 Holborn, 290-91.
32 Rothenberg, 300-04.
38 Clausewitz, 75-89.
41 Cosmas, 110; The Great War.
42 The Great War.
46 Cosmas, 109.
47 Vagts, 456; Smith, 131-32.
50 Smith, 78-147.
51 Vagts, 455; Cosmas, 104.
52 House, 17.
53 Smith, 136.
54 Brereton, 91-92; Smith, 140.
56 Brereton, 91-95; Smith, 139; The Great War.
57 Brereton, 82-94.
60 Ibid., 34-57.
61 Ibid., 57-76.
63 Howard, “Men against Fire,” 518.
64 Howard, “Men against Fire,” 518; Jukes, 32-57.
65 Howard, “Men against Fire,” 518; Streich and Levy, 505.
68 Jukes, 24-76; Crowl, 455-58.
69 Jukes, 34-35.
70 Menning, 154-55.
71 Jukes, 68-76.
72 House, 16.
73 House, 17; Howard, “Men against Fire,” 511.
74 House, 13-30; Howard, “Men against Fire,” 510-526; Jukes, 24-76.
75 Rothenberg, 324.

Howard, “Men against Fire,” 510.


Rothenberg, 319; House, 24.


Breemer, 2.

Crowl, 473-474; Park, 132-33.

Holborn, 289-90.

Showalter, 221-225; Vagts, 455; Smith, 79-150; Jukes, 21-70.

This is partly due to the “diffusion” of insights from previous wars. For more on diffusion in innovation, see Horowitz and Pindyck.

Rothenberg, 297.

Ibid., 320.

House, 25.

Park, 134.

Howard, “Men against Fire,” 511.


Snyder, 118-21.

Rothenberg, 301-02.

Clausewitz; Craig, “Delbrück,” 341-43.


Rothenberg, 313.


Rothenberg, 318-319.


For more on bureaucratic politics, see Allison and Zelikow; Park, 138; Crowl, 473-74.

Breemer, 1-4.


Rothenberg, 311-324.

Ibid., 306-310.

112 Rothenberg, 306-310.

113 For more on friction, chance, probability, and uncertainty, see Clausewitz.

114 Rothenberg, 319; Snyder, 116-17.


119 Breemer, 4.

120 London, 12-20.

121 Snyder, 108-46.

122 Rothenberg, 298.

123 Craig, “The Political Leader as Strategist,” 482-84.

124 Snyder, 109-110.

125 For more on this approach, see Jensen, Whyte, and Cuomo; CSIS Events.

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