Routledge
Taylor & Francis Group

Security Studies

ISSN: 0963-6412 (Print) 1556-1852 (Online) Journal homepage: www.tandfonline.com/journals/fsst20

Implausible Deniability and Escalation in the Gray Zone

Lauren Sukin & Kathryn Hedgecock

To cite this article: Lauren Sukin & Kathryn Hedgecock (21 Apr 2026): Implausible Deniability and Escalation in the Gray Zone, Security Studies, DOI: 10.1080/09636412.2026.2616810

To link to this article: https://doi.org/10.1080/09636412.2026.2616810

SECURITY STUDIES
https://doi.org/10.1080/09636412.2026.2616810

Routledge
Taylor & Francis Group

Implausible Deniability and Escalation in the Gray Zone

Lauren Sukin and Kathryn Hedgecock

ABSTRACT

As gray-zone conflict emerges as the global norm for strategic engagement, plausible and implausible deniability are increasingly critical to competition. The general view is that states use deniability—obscuring which actors took which actions—to limit the extent of accountability for their aggression in the international system. Using survey experiments among US military cadets, we examine how two strategies of deniability common to the gray zone—cyber operations and the use of proxy organizations—influence willingness to respond with force. Scholars have debated whether these strategies are escalatory. We instead argue that the use of deniability strategies makes escalation more likely but also lowers the intensity of the escalation that occurs. Understanding how deniability operates in the gray zone will be increasingly significant as states continue to shift from traditional combat to an environment of strategic competition in the gray zone.

Introduction

The landscape of conflict has evolved, with states increasingly turning to aggression below the threshold of warfare—from cyber operations to sponsoring proxy actors—in the so-called "gray zone."[1] One of the most common features of the gray zone is the use of plausibly (or, in some cases, implausibly) deniable methods.

Lauren Sukin is the John G. Winant Associate Professor in US Foreign Policy at Nuffield College and the Department of Politics and International Relations at the University of Oxford. She is also a Fellow at the Peace Research Center Prague in the Faculty of Social Sciences at Charles University.
Maj. Kathryn Hedgecock previously served as Assistant Professor of International Affairs at the United States Military Academy in West Point, NY. The views expressed in this article are those of the authors and do not necessarily represent the official policy or position of the United States Army, Department of War, or US government.

[1]Rory Cormac and Richard Aldrich, "Grey Is the New Black: Covert Action and Implausible Deniability," *International Affairs* 94, no. 3 (1 May 2018): 477–94; Michael Mazarr, "Front Matter," U.S. Army War College Press, Advancing Strategic Thought Series, 2015, i–vi; Erik Gartzke and Jon Lindsay, *Cross-Domain Deterrence: Strategy in an Era of Complexity* (New York, NY: Oxford University Press, 2019); Andrés Gannon et al., "The Shadow of Deterrence: Why Capable Actors Engage in Contests Short of War," *Journal of Conflict Resolution* 68, no. 2–3 (April 2023): 230–268; Hal Brands, Paradoxes of the Gray Zone, SSRN Scholarly Paper ID 2737593, January 2016; Austin Carson and Keren Yarhi-Milo, "Covert Communication: The Intelligibility and Credibility of Signaling in Secret," *Security Studies* 26, no. 1 (January 2017): 124–56;

Supplemental data for this article can be accessed online at https://doi.org/10.1080/09636412.2026.2616810.

By denying involvement, states hope to evade the costs of aggression. How effective is this strategy? Scholars disagree on whether deniable tactics—such as cyber operations or proxy organizations—help or hinder escalation management. While some scholars argue these tactics induce caution, others say they introduce incentives for aggression and create opportunities for miscalculation.[2] We offer a new approach, arguing deniable attacks raise the likelihood of retaliation while tempering its intensity.

Deniability is an under-theorized area of study. Much scholarship on the gray zone focuses on related features, such as uncertainty, failing to adequately explore deniability in its own right. Yet high certainty and denial can co-exist. The United States initially denied the Bay of Pigs invasion; Israel remained silent for over a decade on its 2007 strike on Syria's al-Kibar nuclear facility; Iran denied involvement in the 2019 attacks on Saudi oil facilities, despite video evidence to the contrary and credit-claiming by Iranian-aligned Houthi rebels.

In this manuscript, we explore deniability in gray-zone warfare by examining the types of actors who operate in the gray zone and the means of attack they employ. Specifically, we investigate how proxy actors and cyber operations—two prominent types of activities in the gray zone—affect escalation. We articulate deniability as a strategy of political cover for aggressive actions, enabled by tactics such as proxies and cyber operations.

We proceed as follows. First, we assess scholarship on the (de)escalatory potential of the gray zone. We motivate our study by conceptualizing deniability as distinct from the more commonly studied phenomenon of uncertainty. We theorize deniability may increase the chance of escalation occurring but, at the same time, result in lower-intensity actions. Next, we outline our empirical approach. We design and conduct an experiment among a novel sample of US Military Academy cadets. We present our results in three parts. First, we demonstrate that cadets react to deniable attacks with a greater preference for escalatory responses. Second, we show that the preferred responses to deniable attacks are lower in intensity than the preferred responses to undeniable attacks. Third, we assess the mechanisms of our theory, leveraging open-ended responses to demonstrate how attuned cadets are to deniability. Our findings show that deniability

Frank Hoffman, "Conflict in the 21st Century: The Rise of Hybrid Wars," 2007; Frank Hoffman, "Examining Complex Forms of Conflict: Gray Zone and Hybrid Challenges," Institute for National Strategic Security, National Defense University, PRISM 7, no. 4 (2018): 30–47; Michael Poznansky, "The Psychology of Overt and Covert Intervention," *Security Studies* 30, no. 3 (May 2021): 325–53.

[2]Gannon et al., "The Shadow of Deterrence"; Jonathan Wilkenfeld et al., "Escalation Management in Gray Zone Crises: The Proxy Factor," *International Studies Quarterly* 66, no. 3 (2022): 143–173; Scott Williamson, "Do Proxies Provide Plausible Deniability? Evidence from Experiments on Three Surveys," *Journal of Conflict Resolution* 68, no. 2–3 (2023): 322–347.

is not simply escalatory nor de-escalatory. Instead, deniability increases the likelihood of escalation, but reduces the intensity of retaliation. Finally, we conclude with a discussion of implications and avenues for further research.

Deniability and (De)escalation in the Gray Zone

Defining Deniability

Gray-zone operations refer to cases "when a military capable challenger intentionally limits the intensity and capacity with which they conduct military operations."[3] Other scholars similarly point to such limitations as crucial to the gray zone. For example, Maass writes that gray-zone operations "stay below the threshold of war … . Whereas conventional understandings of war entail one state's direct application of military force against another … gray zone aggression relies on alternative and unconventional means such as propaganda, sabotage, cyberattacks, proxy forces, and insurrections."[4] Indeed, states can use many different strategies to operate in the gray zone, from sponsoring mercenaries to attack adversaries in their stead to engaging emerging means of conflict, such as cyber or influence operations. These tactics replace or supplement traditional military operations. A critical task in studying the gray zone is to examine the conditions under which these limited operations might escalate.

While there is a very wide variety of activities that can occur in the gray zone, several features are frequently present. For example, one common element of activity in the gray zone is the existence of attribution challenges. Attacks can be difficult to attribute if they are hard to discover or if, once discoverable, their origins are uncertain. Much scholarship on the gray zone has focused on this type of uncertainty.

However, this approach is limiting. Uncertainty is only one component of the complex attribution challenges present in the gray zone. Rather than focusing on whether accurate attribution is possible, some scholars have instead highlighted when or how attribution occurs.[5] For example, attribution can be slow or fast, partial or complete, public or private. Importantly, even in cases with little uncertainty about whodunit, attribution may still fail to materialize if the attacker has made their involvement

[3]Gannon et al., "The Shadow of Deterrence," 5.
[4]Richard W. Maass, "Legal Deterrence by Denial: Strategic Initiative and International Law in the Gray Zone (Summer 2025)," 8, no. 3 (2025).
[5]Kathryn Hedgecock and Lauren Sukin, "Responding to Uncertainty: The Importance of Covertness in Responding to Cyber and Kinetic Attacks," *Journal of Conflict Resolution* 67, no. 10 (2023): 1873–1903; Milton Mueller et al., "Cyber Attribution," *The Cyber Defense Review* 4, no. 1 (2019): 107–22.

deniable, especially if the victim of an attack faces strategic incentives to avoid pointing fingers.

Deniability is an increasingly important and distinct component of attribution in the gray zone. There was hardly any question where the "little green men" in Ukraine came from, and Russia's "patriotic hackers" were always a myth. China's rarely acknowledged maritime militia is no mystery; the "fishing boats" that compose it are purpose-built. Iranian proxies, such as the Houthis, operate throughout the Middle East, although Iran consistently denies supplying weapons to many of these groups. When states contrive ways to establish distance between themselves and their agents, they can deny their involvement in aggression, thus complicating attribution and affecting whether and how retaliation might occur.

Most conflict literature focuses on the conventional realm, where it is typically difficult to deny that attacks have occurred or to deny who has conducted them. However, this represents only a subset of political behavior. As gray-zone activity becomes more common, it will become increasingly important to understand its strategic dynamics, deniability among them.

Here, we explicate how states create deniability, distinguishing the concept from the more commonly studied one of uncertainty. While certainty and uncertainty in attribution refer to two ends of a spectrum of confidence in correctly identifying an attacker, deniability refers to the condition in which an attacker has created the political conditions to deny their involvement in an attack. Denial may be easier when uncertainty is high, but deniability is nonetheless present in many prominent low-uncertainty cases, making it an important phenomenon to study.

Two common methods for creating deniability involve choosing deniable actors to conduct attacks or implementing attacks using deniable means. While denial can be enabled by other approaches as well—such as false-flag and misinformation operations—the selection of a deniable actor or means of attack are two common ways states can distance themselves from their actions. Deniable actors and means allow for political separation between the state and the attack, either by routing aggression through a third party or by employing a sub-conventional method of action.

Figure 1 visualizes this dynamic, showing how states signal deniability by selecting different types of actors or means for aggression. Figure 1 uses examples of Russian or Russian-backed aggression against Ukraine to illustrate. The figure shows multiple ways to achieve deniability; thus, in this study, we assess not only an attack's means but also the type of actor that conducts it.

Figure 1 shows that attacks can be both certain and deniable. For example, aggressors can try to obfuscate responsibility by using detectable but covert actors, such as by sending non-uniformed military personnel instead of uniformed personnel. With this, aggressors can try to deny

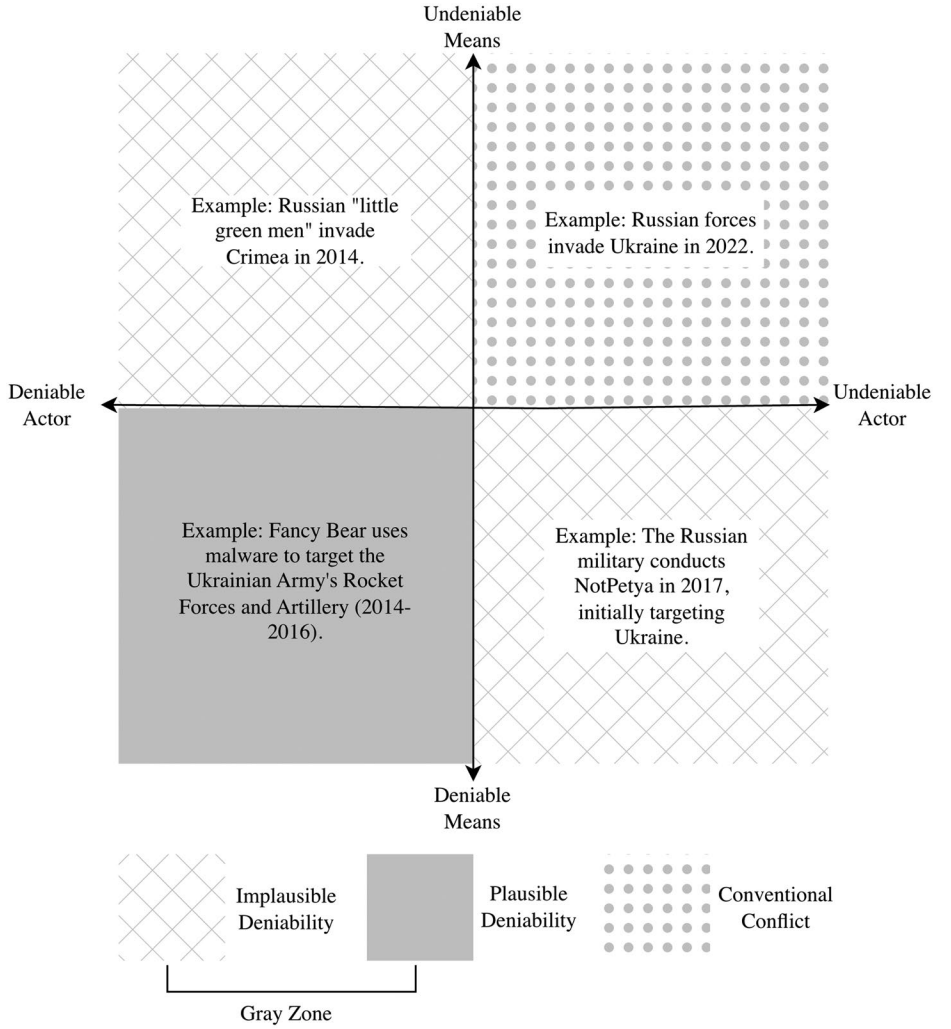

Figure 1. The actor and means of an attack affect its deniability.

involvement to sympathetic audiences who lack "proof" of foul-play. States may do so because deniability lowers visibility to the attacker's own public and reduces the chances of reputation costs from third parties.

Although it may be known that the proxy has a relationship with a state sponsor, an imperfect principal-agent relationship makes it unclear how much the principal (state) directed any action of the agent (proxy). When accusations of misconduct arise, the principal can use this dynamic to contest attribution, claiming noninvolvement with the agent, asserting it acted in defiance of the principal's wishes, or simply by failing to accept attribution. (In this way, deniability is distinct from pure uncertainty; even if the principal's involvement is known with certainty, implausible deniability still shapes the political dynamics.) Our concept of deniability captures both explicit denial (e.g., public statements of denial) and implicit

denial (e.g., silence); however, future work could fruitfully examine the effects of these different denial forms.

States can also use novel and complex means to deny even highly observable aggression. For example, scholars have worried about the military use of artificial intelligence to shift culpability for aggression.[6] In the cyber domain, aggressors can lean on poor non-expert understanding to raise doubts in their own and foreign publics about the veracity of a victim's accusations. This is further enabled by the fact that victims of cyber aggression often cannot fully justify attribution, either because revealing information publicly could bring attention to unpatched vulnerabilities or because it could reveal detection capabilities that are better kept secret.[7]

Denial is a strategic choice. States might choose to deny aggression to reduce audience costs and diminish political risk.[8] Non-state organizations also sometimes strategically deny. They may do so when there is a strain on their principal-agent relationships or when perceptions of the effects of the attack are politically costly.[9] States strategically use secrecy in other ways as well, such as by choosing to reveal or conceal capabilities.[10]

Our approach differs from the existing literature by focusing on deniability rather than uncertainty, thus offering a new approach for understanding the effects of common actions in the gray zone. In addition, we differ from previous studies by highlighting a key similarity across multiple types of gray-zone tactics rather than reviewing them in isolation. In particular, our approach differs from previous work in the cyber domain, which has largely examined how the means of an attack (e.g., whether it is cyber or kinetic) and the effects of an attack (e.g., the scale of damage or the choice of targets) determine responses.[11] We argue that an important, but not unique, feature of cyber means—its ability to create

[6]James Johnson, "The AI Commander Problem: Ethical, Political, and Psychological Dilemmas of Human-Machine Interactions in AI-Enabled Warfare," *Journal of Military Ethics* 21, nos. 3–4 (2022): 246–71.

[7]Thomas Rid and Ben Buchanan, "Attributing Cyber Attacks," *Journal of Strategic Studies* 38, no. 1 (2 January 2015): 4–37; Florian Egloff, "Public Attribution of Cyber Intrusions," *Journal of Cybersecurity* 6, no. 1 (19 May 2020): 1–12.

[8]Michael Poznansky and Evan Perkoski, "Rethinking Secrecy in Cyberspace: The Politics of Voluntary Attribution," *Journal of Global Security Studies* 3, no. 4 (1 October 2018): 402–16.

[9]Max Abrahms and Justin Conrad, "The Strategic Logic of Credit Claiming: A New Theory for Anonymous Terrorist Attacks," *Security Studies* 26, no. 2 (2017): 279–304; Hoffman, "Conflict in the 21st Century: The Rise of Hybrid Wars."

[10]Brendan Rittenhouse Green and Austin Long, "Conceal or Reveal? Managing Clandestine Military Capabilities in Peacetime Competition," *International Security* 44, no. 3 (2019): 48–83.

[11]Sarah Kreps and Jacquelyn Schneider, "Escalation Firebreaks in the Cyber, Conventional, and Nuclear Domains: Moving Beyond Effects-Based Logics," *Journal of Cybersecurity* 5, no. 1 (2019): 1–11; Marcelo Leal and Paul Musgrave, "Hitting Back or Holding Back in Cyberspace: Experimental Evidence Regarding Americans' Responses to Cyberattacks," *Conflict Management and Peace Science* 40, no. 1 (2023): 42–64; Hedgecock and Sukin, "Responding to Uncertainty."

deniability—is essential to understanding how cyber operations will be received. We further trace the effects of deniability for the use of proxies.

In the sections that follow, we explore disagreement in the scholarly literature on whether deniable operations in the gray zone are escalatory or de-escalatory. We then offer a third perspective, arguing that deniable attacks should increase the likelihood of escalation, while decreasing the likely intensity of retaliation, before testing our arguments.

The Escalatory Gray Zone

Deniable attacks may incentivize escalation. When states use deniability to evade culpability for their actions, targets face a hard choice. Backing down could signal to the aggressor that they have free rein, so long as they keep their movements below a certain threshold. Proportional or in-kind retaliation can make further competition somewhat costly, but as gray-zone competition is "cheap," this is likely insufficient to deter. Imagine, for example, that an aggressor sponsors an attack by a proxy; retaliating against that proxy does a minimal amount of harm to the sponsor state. Weakening the proxy could make further aggression more complicated, but the aggressor's incentives for action hardly change. A third option is to escalate. Domestic and foreign audiences see this risk as disproportionate, invoking political backlash. However, by raising the costs of action and bringing consequences to the aggressor, this option could deter. From this perspective, deniability may force a target's hand by requiring escalation for deterrence purposes. This underlies a key weakness of deniable attacks: while they can undermine an adversary's capabilities or signal interests, they often do little on their own to deter or coerce.

States have sometimes taken this route. The Bush Doctrine's approach of holding states responsible for the actions of non-state organizations operating in their territory was, among other things, intended to deter states from using deniability to sponsor aggression. A series of attacks against US forces by Iranian-backed militias purported to be for the purposes of reestablishing deterrence against Iran preceded the killing of Qasem Soleimani. Israel linked several of its attacks against Iranian officials connected to the Quds Force to Iranian support for Hamas' 7 October 2023 attack on Israel, although Iran has denied its involvement.

Scholarship on common gray-zone tactics, such as proxies and cyber operations, suggests these tactics may be both deniable and escalatory. For example, Williamson finds no less desire for forceful retaliation against proxies than against states that sponsor them.[12] However, the use of proxy

[12]Williamson, "Do Proxies Provide Plausible Deniability?"

organizations decreases the blame individuals place on sponsor states. This suggests that states are effectively using proxies to achieve deniability, but that this does not reduce escalation. Wilkenfeld et al. similarly show proxy use is more likely to escalate, even when a proxy's initial incursion is nonviolent.[13] This points to a deterrence-based logic, where punishment is necessary to prevent further proxy use. In the cyber domain, Leal and Musgrave determine Americans are more likely to support retaliation against a state or terrorist organization that conducts a cyberattack than unaffiliated individuals who do so, pointing to a deterrence-by-punishment logic to explain this outcome.[14]

States relying on cyber operations for deniability make use of the fact that novel and complex means can raise questions about responsibility. It is easier to deny involvement in a poorly understood phenomenon shrouded in secrecy. Even among expert communities, the prominence of false flag operations in the cyber domain might raise skepticism about the accuracy of attribution. This is, in part, because even if decision-makers fully know who is at fault for a cyberattack, they may be unable to reveal how they know this information or feel they will be unable to convince publics of proper attribution. In a world of persistent disinformation, the "shadow of a doubt" created by the complexity of cyber operations may be enough to establish deniability.[15]

Many scholars have pointed to these same features of cyber operations to argue that their use is escalatory. Publics may simply lack the expertise to adequately distinguish between cyber and kinetic operations and may, as a result, demand similar responses to both kinds of aggression.[16] Even relatively well-informed decision-makers may not fully understand how a certain exploit works; this can lead individuals to overestimate their ability to manage new threats and crave certain, decisive action to regain "control."[17] Scholars have argued the novelty of cyberattacks may drive aggression, pointing to misunderstanding and fear that accompanies new systems that are not fully understood by relevant decision-makers.[18] As a result,

[13]Wilkenfeld et al., "Escalation Management in Gray Zone Crises."
[14]Leal and Musgrave, "Hitting Back or Holding Back in Cyberspace."
[15]Vulnerability to disinformation is not isolated to civilians; Russian disinformation campaigns have frequently targeted veterans and service-members by using patriotic narratives and creating fake military and military-adjacent profiles: Dana Weinberg et al., "How the Russian Influence Operation on Twitter Weaponized Military Narratives," 2021. OSF Preprint. https://doi.org/10.31235/osf.io/b9a2m.
[16]Ryan Shandler, "Cyber Conflict & Domestic Audience Costs," *International Interactions* (2025): 1–25; Hedgecock and Sukin, "Responding to Uncertainty."
[17]Jacquelyn Schneider et al., "Hacking Nuclear Stability: Wargaming Technology, Uncertainty, and Escalation," *International Organization* (2023): 1–35; Jennifer Mitzen and Randall Schweller, "Knowing the Unknown Unknowns: Misplaced Certainty and the Onset of War," *Security Studies* 20, no. 1 (March 2011): 2–35.
[18]Michael Gross et al., "The Psychological Effects of Cyber Terrorism," Taylor & Francis, *Bulletin of the Atomic Scientists* 72, no. 5 (September 2016): 284–91; Barry Posen, *Inadvertent Escalation: Conventional War and Nuclear Risks* (Ithaca, NY: Cornell University Press, 1991); Schneider et al., "Hacking Nuclear

states may make the "safe" assumption that their adversary's capabilities are to be feared. Military organizational culture that prefers offense may make states with professional militaries more likely to make such assumptions.[19] States may also rely on such assumptions because faster decision-making timelines created by new technologies reduce the decision space for de-escalation.[20]

This camp of scholarship suggests gray zone operations—such as cyber operations or attacks using proxies—are escalatory, as the very features that make these attacks deniable may also incentivize strong responses. This scholarship expects that:

Hypothesis 1 (H1): Deniability increases the likelihood of escalation.

The De-Escalatory Gray Zone

In contrast, deniability could be de-escalatory by preventing or undermining attribution. This matters because punishment may only be seen as "right" or "just" if it occurs against the correct actor.[21] Punishment against an incorrect actor could be seen as morally wrong and thus damage a state's reputation. States should therefore seek to avoid escalating when deniability is at play.

For example, if a target does not attribute an attack against it—perhaps because the target does not wish to reveal how they were able to detect the attack or attacker—then the target would suffer costs if it retaliated anyway. Particularly in the cyber domain, states face significant incentives to avoid attribution to maintain secrecy.[22] Strategic dynamics between adversaries can also lead to states avoiding (public) attribution, as mutual deniability can help states keep from activating

Stability"; Nadiya Kostyuk and Carly Wayne, "The Microfoundations of State Cybersecurity: Cyber Risk Perceptions and the Mass Public," *Journal of Global Security Studies* 6, no. 2 (2020): 1–25; Erik Lin-Greenberg, "Evaluating Escalation: Conceptualizing Escalation in an Era of Emerging Military Technologies," *The Journal of Politics* 85, no. 3 (July 2023): 1151–55.

[19] Jeffrey Legro, "Military Culture and Inadvertent Escalation in World War II," *International Security* 18, no. 4 (1994): 108–42; Robert Jervis, "Cooperation under the Security Dilemma," *World Politics* 30, no. 2 (1978): 167–214; Scott Sagan, "1914 Revisited: Allies, Offense, and Instability," *International Security* 11, no. 2 (1986): 151–75.

[20] James Johnson, "Inadvertent Escalation in the Age of Intelligence Machines: A New Model for Nuclear Risk in the Digital Age," *European Journal of International Security* 7, no. 3 (August 2022): 337–59.

[21] Aaron Brantly, "The Cyber Deterrence Problem," in *2018 10th International Conference on Cyber Conflict (CyCon)*, 2018 10th International Conference on Cyber Conflict (CyCon) (2018), 31–54; Rose McDermott et al., "'Blunt Not the Heart, Enrage It': The Psychology of Revenge and Deterrence," *Texas National Security Review* 1, no. 1 (24 November 2017); Rid and Buchanan, "Attributing Cyber Attacks."

[22] Egloff, "Public Attribution of Cyber Intrusions"; Erica Borghard and Shawn Lonergan, "Cyber Operations as Imperfect Tools of Escalation," *Strategic Studies Quarterly* 13, no. 3 (2019): 122–45; Rid and Buchanan, "Attributing Cyber Attacks."

domestic hawks.[23] States can sometimes use private attribution to help manage escalation.[24]

If the target does attempt to attribute and to retaliate against an attack, but its narrative is not believable because of denials by the aggressor, then audiences that believe the denial might punish the target, even if it is, in truth, acting "in the right." That is, even if the identification of the attacker is correct, contestation over attribution could lead to an interpretation of unjust punishment by third-parties. Indeed, Abramson and Baram find states resort to face-saving, de-escalatory strategies after cyberattacks, even when they have high confidence in the identity of the attacker.[25]

This logic also helps explain why many scholars view de-escalation as typical of proxy use. Proxy wars are sometimes even defined with this in mind; Mumford, for example, categorizes proxy wars as "arm's-length 'effects-based operations' whereby a specific objective is desired … without risking foreseen consequences (conflict escalation with a rival super power, for example)."[26] Much empirical scholarship finds proxies reduce the chances of escalation.[27] Wilkenfeld et al. write: "most academic literature on this subject presumes … the challenger seeks to lower the risks it would otherwise incur by taking direct action."[28] Because proxies give their sponsors a(n) (im)plausible shield behind which to deny their involvement, retaliation may be seen as less appropriate, making escalation less likely. If the involvement of the state sponsor is contested—even in cases where links to the proxy are well-known—retaliators may incur reputation costs for seemingly inappropriate retribution.

Scholars have also argued that cyber means can be de-escalatory by pointing to the same mechanisms—novelty and complexity—that enable denial in the cyber domain. While some scholars argue novelty spurs fear, for others, the use of new technologies can induce caution.[29] A failure to understand how a cyber operation worked may lead to hesitancy that restrains decision-making.[30] Accordingly, cyber operations may have a

[23]Conversely, highly visible and consequently less deniable attacks may mobilize the public to demand a significant response: Poznansky, "The Psychology of Overt and Covert Intervention"; Austin Carson, *Secret Wars: Covert Conflict in International Politics* (Princeton, NY: Princeton University Press, 25 September 2018). Or such attacks may galvanize the attacker's public, enhancing the adversary's stakes and resolve. Denied attacks avoid these pressures.
[24]Poznansky and Perkoski, "Rethinking Secrecy in Cyberspace."
[25]Yehonatan Abramson and Gil Baram, "Saving Face in the Cyberspace: Responses to Public Cyber Intrusions in the Gulf," *Contemporary Security Policy* 45, no. 2 (2024): 210–38.
[26]Andrew Mumford, "Proxy Warfare and the Future of Conflict," *The RUSI Journal* 158, no. 2 (2013): 40–46.
[27]Daniel Byman and Sarah Kreps, "Agents of Destruction? Applying Principal-Agent Analysis to State-Sponsored Terrorism," *International Studies Perspectives* 11, no. 1 (2010): 1–18; Bertil Dunér, "Proxy Intervention in Civil Wars," *Journal of Peace Research* 18, no. 4 (1981): 353–61; Michael Klare, "Subterranean Alliances: America's Global Proxy Network," *Journal of International Affairs* (1989): 97–118.
[28]Wilkenfeld et al., "Escalation Management in Gray Zone Crises," 3.
[29]Schneider et al., "Hacking Nuclear Stability."
[30]Miguel Alberto Gomez and Eula Bianca Villar, "Fear, Uncertainty, and Dread: Cognitive Heuristics and Cyber Threats," *Politics and Governance* 6, no. 2 (11 June 2018): 61–72; Monica Kaminska, "Restraint under

lower propensity to escalate.[31] Importantly, this novelty-derived fear exists regardless of whether attribution is certain or uncertain, as it is instead inherent to the complex and new nature of cyber tools. Cyber operations may even be employed as an intentional off-ramp to conflict for this reason.[32] Shandler argues it is the public's lack of ability to fully understand and distinguish cyber operations that allows decision-makers to "back away" from a conflict by employing cyber means to satisfy the public's demands for retaliation while, in practice, staying below the conventional threshold.[33] A key strand of scholarship suggests the cyber domain may act as a "firebreak," where any chain of escalation occurring in response to a cyber event remains in-domain.[34]

From this perspective, deniable attacks—including those conducted via proxies or using cyber means—should be de-escalatory. This is because these strategies encourage decision-makers to act cautiously, such that:

Hypothesis 2 (H2): Deniability decreases the likelihood of escalation.

Escalation Likelihood and Retaliation Intensity

How can we reconcile competing perspectives on the (de)escalatory potential of deniability? Do deniable strategies—whether through the use of covert actions and/or covert means—create aggressive competition? Or do they induce caution? We argue that a third perspective is needed.

We posit that, compared to conventional attacks, aggression in the gray zone is more likely to result in escalation—but, at the same time, that retaliation is likely to be lower in intensity. In the remainder of this section, we explain our third hypothesis, which states:

Hypothesis 3 (H3): Deniability increases the likelihood of escalation but decreases the intensity of retaliation.

Conditions of Uncertainty: Why the United States Tolerates Cyberattacks," *Journal of Cybersecurity* 7, no. 1 (2021): 1–15.

[31]Ryan Maness and Brandon Valeriano, "The Impact of Cyber Conflict on International Interactions," *Armed Forces & Society* 42, no. 2 (April 2016): 301–23; Jacquelyn Schneider, "Cyber and Crisis Escalation: Insights from Wargaming," in *USASOC Futures Forum*, Vol. 43 (2017); Kreps and Schneider, "Escalation firebreaks in the cyber, conventional, and nuclear domains"; Brandon Valeriano and Ryan Maness, "The Dynamics of Cyber Conflict between Rival Antagonists, 2001–11," *Journal of Peace Research* 51, no. 3 (May 2014): 347–60; Rid and Buchanan, "Attributing Cyber Attacks"; Joseph Nye, "Deterrence and Dissuasion in Cyberspace," *International Security* 41, no. 3 (January 2017): 44–71; Brantly, ""The Cyber Deterrence Problem."

[32]Brandon Valeriano and Ryan Maness, "Deescalation Pathways and Disruptive Technology: Cyber Operations as Off-Ramps to War," in *Cyber Peace: Charting a Path Toward a Sustainable, Stable, and Secure Cyberspace,* ed. Scott Shackelford et al. (Cambridge University Press, 2022), 64–94.

[33]Shandler, "Cyber Conflict & Domestic Audience Costs."

[34]Kreps and Schneider, "Escalation firebreaks in the cyber, conventional, and nuclear domains"; Lin-Greenberg, "Evaluating Escalation"; Kaminska, "Restraint Under Conditions of Uncertainty"; Kostyuk and Wayne, "The Microfoundations of State Cybersecurity"; Ryan Shandler et al., "Cyber Terrorism and Public Support for Retaliation—A Multi-Country Survey Experiment," 52, no. 2 (2022): 850–68; Gross et al., "The Psychological Effects of Cyber Terrorism."

First, we offer a definition of escalation that focuses on threshold-crossing. We delineate the likelihood of escalation from the intensity of retaliation. Next, we consider how deniability—separate from uncertainty—affects the strategic dynamics of action-and-response along the escalation ladder.

Defining Escalation

Following an act of aggression, states have the option to de-escalate, respond proportionally, or escalate. Proportional responses often involve an in-kind or like-for-like response, using similar weapons, focusing on similar targets, or causing similar amounts of damage as the initial attack. Alternatively, if states choose to escalate or de-escalate, they have a wider menu of options. De-escalation ranges from backing down altogether to a retaliatory operation just slightly less intense than the initial attack. Escalation could increase scale, political significance, or visibility relative to the initial attack by changing the retaliatory attack's outcome, target, or means.

Often, escalation involves crossing certain thresholds, such as moving from conventional to nuclear war. Kahn, for example, designs an escalation ladder broken into six thresholds whose crossing represent "very sharp changes in the character" of conflict.[35] Many of these thresholds relate to the means of attack, such as the use of nuclear weapons or choice of targets.[36] While many iterations of the ladder have emerged since Kahn's seminal work, threshold-crossing has remained a core component. In an important modern work, Morgan et al. define escalation as occurring when an action "crosses thresholds considered significant by one or more of the participants," such as violating territorial boundaries.[37]

Many other scholars similarly refer to crossing thresholds (or "bright" or "red" lines) as crucial to the concept of escalation.[38] Some iterations fragment into dimensions such as vertical escalation—related to impact or scale—and horizontal escalation—indicating an expansion into new physical areas or domains.[39] Such multidimensional models offer an

[35]Herman Kahn, *On Escalation: Metaphors and Scenarios* (New York, NY: Routledge, 1965).
[36]Some versions of the ladder include non-military actions; for example, Kahn *On Escalation* assigns economic measures as the second rung on a 44-rung escalation ladder.
[37]Forrest Morgan et al., *Dangerous Thresholds: Managing Escalation in the 21st Century* (Rand Corporation, 2008), 8.
[38]Bruno Tertrais, "Drawing Red Lines Right," *The Washington Quarterly* 37, no. 3 (2014): 7–24; Daniel Altman, "Red Lines and Faits Accomplis in Interstate Coercion and Crisis" (PhD diss., Massachusetts Institute of Technology, 2015); Martin Libicki, *The New Calculus of Escalation: Avoiding Armageddon in Great Power Conflict* (Washington, D.C.: Georgetown University Press, 2025); Martin Libicki and Olesya Tkacheva, "Cyberspace Escalation: Ladders or Lattices?," *Cyber Threats and NATO 2030: Horizon Scanning and Analysis* 60 (2020): 60–72.
[39]Joshua Epstein, "Horizontal Escalation: Sour Notes of a Recurrent Theme," *International Security* 8, no. 3 (1983): 19–31.

alternative conceptualization of escalation as movement across multiple thresholds. The threshold-based articulation of escalation persists for a wide variety of actions. In cyberspace, Lin explains escalation as crossing certain "lines in the sand."[40] Johnson illustrates an escalation ladder for covert action, placing more visible and "intrusive" actions higher.[41] While the activities that appear on the ladder and their precise order might change, the general principle of the ladder is that each sequential rung involves a more intense action.[42]

Intensity is a function of both the effects of an action and its means. Lin-Greenberg argues a "modern escalation ladder" not only has actions with bigger effects (such as higher casualties) at higher rungs but also has "actions that are more physically present and visible on the battlefield fall at higher rungs."[43] Lin-Greenberg presents cross-national data from political elites in the United States, Singapore, and India, showing visibility is an important factor in assessing the intensity of an action. In this perspective, two attacks with equal real-world effects sit at different levels of intensity if one is more deniable.

Much of the aforementioned literature has treated (de)escalation as a fairly simplistic feature of policymaking, where attacks produce responses sitting "higher" or "lower" on the escalation ladder (or at the same rung). We argue that perspective misses a critical nuance: which step of the ladder one is on also matters. On a 10-step ladder, moving up a single rung from step 1 results in a very different outcome than moving up a single rung from step 9. That is, the likelihood of escalation is not the only salient feature. We also need to consider the intensity of the retaliatory action.

For this paper, we are interested in a largely monadic view of escalation. That is, given State A has used force, State B faces a decision about where on the ladder its retaliatory response will sit. To do this, it must consider both the past (i.e., what State A has done) and the future (i.e., how State A will react to State B's retaliation). This quantity is of interest for two key reasons. First, State B's decision to move "up" or "down" (or to stay at the same level) on the ladder has critical effects on whether and how conflict will end or continue. The ability to effectively deter, for example, relies on this choice. States can choose to escalate in order to signal their resolve to continue moving up the ladder in a bid to deter future aggression. Escalatory responses then represent a calculated effort to out-bid the

[40]Herbert Lin, "Escalation Dynamics and Conflict Termination in Cyberspace," *Strategic Studies Quarterly* 6, no. 3 (2012): 52.
[41]Loch Johnson, "On Drawing a Bright Line for Covert Operations," *American Journal of International Law* 86, no. 2 (1992): 284–309.
[42]States have different views of where thresholds lie: Nadiya Kostyuk et al., "Determinants of the Cyber Escalation Ladder," *The Cyber Defense Review* 3, no. 1 (2018): 123–34.
[43]Lin-Greenberg, "Evaluating Escalation."

other side's willingness to endure pain.[44] But the ability to effectively resolve a conflict also depends on where on the ladder State B decides to respond to State A. States can decide their risk-tolerance and thus de-escalate to evade the costly consequences at the top. State B's decision about how to respond to State A therefore shapes both the process and outcome of the exchange. Additionally, it is important to understand State B's decision-making because this process is of interest to State A. Like State B, at each decision-point, State A must consider the past (i.e., State B's actions) and the future (i.e., State B's likely reactions). While the overall course of a conflict's (de)escalation is inherently a reciprocal process, it can be broken down into a series of strategic decision-points for each actor.

The Countervailing Effects of Deniability

We now discuss how states approach retaliation, identifying how key differences between deniable and undeniable attacks shape decisions about whether and how retaliation should occur. We start with the argument that deniability is part of the escalation ladder, such that deniable attacks sit at lower-intensity rungs than undeniable ones.

We argue that deniable attacks result in more frequent escalation because of a simple feature of the escalation ladder: the costs to both the attacker and the target are highest at the top of the ladder and lowest at the bottom. Thus, de-escalation should be more common at higher rungs, while escalation should be more common at lower rungs. This is because every retaliatory action risks escalation by one's adversary. Since the cost of this increases at higher rungs of the ladder, states should be more cautious at higher rungs than at lower rungs. This is the strategic logic that underpins the "stability-instability" problem in nuclear deterrence. Because of the immense cost of a nuclear exchange, states should be very cautious about taking high-intensity actions. At the same time, anticipated caution at the top of the ladder makes it easier for states to aggress on lower rungs, as they know their adversary won't easily jump up toward more costly punishments. A similar, albeit less strict, dynamic distinguishing conventional and gray-zone conflict may exist.

Even though escalation may be more common at lower rungs, retaliation to lower-rung attacks should remain less intense than retaliation to higher-rung attacks. This is because jumping up too many rungs at once may undermine deterrence signaling. Imagine a state responded to a

[44]Herman Kahn, *On Thermonuclear War* (Piscataway, NJ: Transaction Publishers, 1960); Herman Kahn, *Thinking about the Unthinkable* (London, UK: Horizon Press, 1962); Ernst Fehr and Simon Gächter, "Fairness and Retaliation: The Economics of Reciprocity," *Journal of Economic Perspectives* 14, no. 3 (2000): 159–82; Robert Axelrod, *The Evolution of Cooperation: Revised Edition* (New York, NY: Basic Books, 2009).

limited cyberattack by bombing a major military facility belonging to that adversary. This extreme reaction may signal that conflict is inevitable, thereby altering the adversary's calculations and making them less likely to back down. Because more deniable attacks sit at the lower end of the escalation ladder, any further escalation in response to them is also likely to remain at low intensities. Moreover, states typically lack the strategic incentive to escalate significantly after a deniable operation; they choose gray-zone operations over lower-stakes issues, where the costs of massive escalation are unlikely to be worthwhile. Escalation at the higher end of the ladder, if it occurs, is also likely to only involve moving a few rungs at a time because of the higher costs of high-intensity actions. Yet, even though the pace does not differ, the outcome does: any escalation from high-intensity points on the ladder will necessarily result in high-intensity actions. (In addition, de-escalation is unlikely to move down by many rungs, as states retain an incentive to deter by communicating that further actions will still be costly.)

Our argument departs from existing literature on the gray zone. For example, recent work applying the stability-instability paradox to the gray zone has highlighted the cyber-nuclear nexus, exploring the slim likelihood that cyber conflict will escalate into nuclear conflict.[45] Instead, we focus on more common types of competition and conflict. Other scholars have made analogous arguments that build on the stability-instability logic; for example, a common explanation for escalation dynamics within the gray zone centers on the idea that gray-zone strategies do not escalate beyond certain "firebreaks."[46]

The firebreaks argument suggests that the high costs of conventional conflict make sub-conventional conflict more likely and more frequent. However, there are two key concerns with the firebreak approach. First, while the logic explains the rise of gray-zone strategies, it does not explain the mechanisms for escalation within the gray zone. Indeed, many scholars have argued the features of the gray zone, such as deniability, reduce the risk of escalation. Instead, we expect deniable attacks to be more likely to trigger escalation. However, we offer an important addendum: retaliation against deniable attacks will nonetheless tend toward lower-intensity actions

[45] Jon Lindsay and Erik Gartzke, *Coercion through Cyberspace: The Stability-Instability Paradox Revisited*, ed. Kelly Greenhill and Peter Krause (New York, NY: Oxford University Press, 2018); Jacquelyn Schneider, "Emerging Technologies and International Stability," in *The Capability/Vulnerability Paradox and Military Revolutions: Implications for Computing, Cyber, and the Onset of War*, ed. Todd Sechser et al. (Routledge, 2021): 841–863.

[46] Kreps and Schneider, "Escalation firebreaks in the cyber, conventional, and nuclear domains"; Henry Farrell and Charles Glaser, "The Role of Effects, Saliencies and Norms in US Cyberwar Doctrine," *Journal of Cybersecurity* 3, no. 1 (1 March 2017): 7–17; Michael Fischerkeller and Richard Harknett, "Deterrence Is Not a Credible Strategy for Cyberspace," *Orbis* 61, no. 3 (1 January 2017): 381–93.

more than retaliation against higher-intensity attacks. Additionally, the firebreak argument suggests clear divisions between sub-conventional and conventional conflict, such that gray zone activities should not escalate to the conventional domain. For example, in the cyber domain, this literature claims states manage escalation by responding in kind, such as by using only cyber operations to respond to cyberattacks. In contrast, we argue that escalation moves slowly, meaning it may often stay below certain firebreaks, but that it can also cross them.[47] Indeed, we will demonstrate significant potential for cross-domain escalation.

Methodology

Cadet Sample

Studying gray-zone activity is empirically challenging. First, covert operations are a particularly common feature of the gray zone, making escalation difficult to observe and measure. Second, actions in the gray zone are often transnational, multi-actor, and slow, making them difficult to identify and analyze as discrete events. Our approach, therefore, instead relies on participant observation, using a survey experiment administered to a sample of US Military Academy cadets.[48] This allows us to circumvent the empirical problems previously described by clearly specifying a scenario.

Our sample is composed of 735 cadets at the United States Military Academy.[49] Cadets are informed, highly educated, and attentive. They are actively taught about gray-zone warfare, cyber operations, targeting decisions, and escalation dynamics.[50] Cadets are military trainees and should become junior officers upon graduation.[51] Junior officers, while unlikely to be involved in strategic decision-making, are regularly involved in

[47]Other scholars have similarly shown limited pathways for cross-domain escalation: Borghard and Lonergan, "Cyber Operations as Imperfect Tools of Escalation"; Gartzke and Lindsay, *Cross-Domain Deterrence: Strategy in an Era of Complexity*; Jason Healy and Robert Jervis, "The Escalation Inversion and Other Oddities of Situational Cyber Stability," *Texas National Security Review* 3, no. 1 (2020): 31–53.

[48]See Online Supplemental Appendix 1 for the survey text. We received ethics approval from the United States Military Academy (USMA) Human Research Board IRB, with approval number CA-2022-162.

[49]94% of respondents completed the survey. Categorization of responses was incomplete if respondents did not get to the final question of the survey. Incomplete rates are similar between treatments. The completion rates are 92 percent, 96 percent and 95 percent for Insurgents, Iranian-Sponsored Insurgents, and Iranian Military treatments, respectively.

[50]Francesca Spidalieri and Jennifer McArdle, "Transforming the Next Generation of Military Leaders into Cyber-Strategic Leaders: The Role of Cybersecurity Education in US Service Academies," *The Cyber Defense Review* 1, no. 1 (2016): 141–64.

[51]89 percent of our respondents are juniors or seniors, with the remaining 11 percent being sophomores. This means 89 percent will commission as officers in the U.S. Army unless academically, physically or medically disqualified for service (typically quite rare). The remaining respondents are not yet bound to become officers but, in general, students are very unlikely to transfer from West Point before their affirmation ceremony. We dropped 37 respondents in the full model because they did not report their class year.

targeting decisions and planning platoon operations. Upon graduation, many cadets will serve in locations or roles that involve gray-zone operations, either in an advisory capacity or in a direct position of responsibility. For example, some cadets will become cyber officers on the National Mission Force. In this role, they would inform responses to cyber operations around the globe.

Researchers are often significantly limited in access and opportunities to administer surveys to cadets. The small scholarship on cadets is highly varied, with no other studies, to our knowledge, investigating deniability, cyber operations, or gray-zone conflict. In the social science realm, scholars have investigated cadets' motivations for service, professionalization, attitudes toward gender roles, and social values.[52] Very few studies explicitly explore cadets' thinking on military, foreign policy, or strategic issues. A notable exception is Jost et al., which shows that, like military officers in general, cadets are more hawkish than civilians.[53] (Cadets mirror military officers in other ways as well; for example, they are more likely to be male and Republican.)

A very small number of studies have explored foreign policy attitudes among Reserve Officers' Training Corps (ROTC) cadets. A lower percentage of ROTC cadets than West Point cadets become military officers. However, ROTC samples may nonetheless provide some insight into military socialization. Scholars have examined ROTC cadets' views on military restraint, showing both the role and limitations of military socialization.[54] Lushenko and Sparrow similarly study ROTC cadets' views on military policy, examining their degree of trust in artificial intelligence (AI).[55]

Beyond cadets, some studies have assessed the foreign policy preferences of military officers undergoing graduate-level Professional Military Education (PME). Hanson and Knuppe finds opposition among Naval War College students to the use of military force to crack down on political

[52]Risa Brooks et al., "What Makes a Military Professional? Evaluating Norm Socialization in West Point Cadets," *Armed Forces & Society* 48, no. 4 (2022): 803–27; Amy Wrzesniewski et al., "Multiple Types of Motives Don't Multiply the Motivation of West Point Cadets," *Proceedings of the National Academy of Sciences* 111, no. 30 (2014): 10990–95; John Hammill et al., "Self-Selection and Parental Socioeconomic Status as Determinants of the Values of West Point Cadets," *Armed Forces & Society* 22, no. 1 (1995): 103–15; Volker Franke, "Generation X and the Military: A Comparison of Attitudes and Values Between West Point Cadets and College Students," *Journal of Political & Military Sociology* (2001): 92–119; Jerome Adams, "Women at West Point: A Three-Year Perspective," *Sex Roles* 11, nos. 5-6 (1984): 525–41.

[53]Tyler Jost et al., "The Character and Origins of Military Attitudes on the Use of Force," *International Studies Quarterly* 66, no. 2 (2022).

[54]Andrew Bell, "Combatant Socialization and Norms of Restraint: Examining officer training at the US Military Academy and Army ROTC," *Journal of Peace Research* 59, no. 2 (2022): 180–96; Andrew Bell et al., "The Moral Foundations of Restraint: Partisanship, Military Training, and Norms of Civilian Protection," *Journal of Peace Research* 59, no. 5 (2022): 694–709.

[55]Paul Lushenko and Robert Sparrow, "Artificial Intelligence and US Military Cadets' Attitudes about Future War," *Armed Forces & Society* (2024): 1–28.

opponents.[56] Notably, Schneider et al. examine the results of a wargame conducted at the Naval War College, showing officers were cautious in their selection of cyber strategies because they feared adversary retaliation.[57] Other scholars discuss the methodological and educational implications of wargames using studies from war colleges.[58] PME samples represent valuable insight into military decision-making, but they are distinct from studies of cadets in that the individuals in the sample are already serving as intermediate or senior officers.

Cadets are not actively serving and lack the field experience of officers, but studying cadets nevertheless provides insight into military decision-making. Research on cadets shows they are socialized and professionalized into military norms from very early on in their training.[59] Those who select into West Point tend to have different social and political preferences than their peers at civilian educational institutions. These traits—such as a higher degree of nationalism and greater hawkishness—remain prominent among US military service-members overall.[60] Many of cadets' unique traits, such as a tendency toward heroism and conservatism, tend to be influential in shaping political preferences. In Online Supplemental Appendix 4, we provide further information about the views of our sample of cadets, such as that they are moderately globalist and somewhat vengeful. Cadets also match the demographics of the officer corps.[61]

Study Design

We asked cadets to read a description of an attack on US military forces, starting with a brief scene-setter stating: "The United States has a small number of military forces stationed in Somalia, an African country on the Indian Ocean. US forces are responsible for training and assistance to Somali anti-piracy and anti-terrorism operatives. In recent years, piracy and terrorism organizations have been operating in Somalia." This background information ensures comparability.

[56]Kolby Hanson and Austin Knuppe, "Polarization Versus Professionalism: Military and Civilian Views on the Domestic Use of the Military," *Political Science Research and Methods* (2024): 1–18.
[57]Schneider, "Cyber and Crisis Escalation."
[58]Amanda Rosen and Lisa Kerr, "Wargaming for Learning: How Educational Gaming Supports Student Learning and Perspectives," *Journal of Political Science Education* 20, no. 2 (2024): 318–35; Benjamin Schechter et al., "Wargaming as a Methodology: The International Crisis Wargame and Experimental Wargaming," *Simulation & Gaming* 52, no. 4 (2021): 513–26.
[59]Brooks et al., "What Makes a Military Professional?."
[60]Franke, "Generation X and the Military"; Hammill et al., "Self-Selection and Parental Socioeconomic Status as Determinants of the Values of West Point Cadets"; Bell, "Combatant Socialization and Norms of Restraint"; Jost et al., "The Character and Origins of Military Attitudes on the Use of Force."
[61]See Online Supplemental Appendix 3 for a Comparison of Respondent Demographics to the U.S. Army Officer Corps.

Relative to conflicts in highly contested regions—such as the Middle East—or with a great power competitor—such as Russia or China, we expect cadets to have more neutral opinions about the Somalia mission. Cadets generally support the mission; a majority report believing in its underlying values and agree the United States "needs to play an active role in solving conflicts around the world."[62] As the United States continues to expand its anti-terrorism missions, protracted conflicts involving multiple insurgent actors, like the conflict in Somalia, are increasingly important.

Each participant received a vignette describing an attack with two randomized parameters: the *attacker* and *attack type*. This allows us to vary how deniability is achieved. The vignette reads:

> Imagine that [*attacker*] has just crossed the Somali border and engaged in [*attack type*] conflict with U.S. forces in a coastal city, resulting in casualties to four U.S. soldiers and eight Somali civilians.

We vary whether the source of deniability comes from the attack's actor or means. Table 1 displays the treatments. First, we randomly informed respondents about the *attack type*, which varied between a "cyberattack" or an "armed attack."[63] This dimension introduces deniability because cyber operations are less well-understood and more difficult to detect and trace, making them less visible than an equivalent kinetic attack.

We also exploit a second major source of deniability in gray zone operations: the use of proxy actors. We include three potential treatments for the *attacker*: the Iranian army, an Iranian-sponsored insurgent group, or an insurgent group.[64] Each introduces further complications for attribution by making it unclear who has conducted the attack or what the depth of state involvement is. The least deniable option involves the Iranian army acting directly. When the attack comes from a proxy with known connections to Iran, Iran's involvement is at least implausibly deniable. Even under relatively high-information conditions, the extent of Iran's involvement cannot be truly known due to the principal-agent problem. Even if Iran exerts significant control over a proxy, it cannot fully control

[62]We name a specific conflict because we expect cadets to have sufficient information about U.S. military posture. We anticipate they would make strong inferences about a model country if the treatment were not real-world, skewing results. Thus, we prioritize internal validity over generalizability.

[63]Earlier, we told respondents: "Cyber conflict is a form of conflict that exploits cyberspace and cyber-enabled technologies. Examples of cyber conflict include disabling or destroying infrastructure and networks, denying access to communication systems, or infiltrating critical infrastructure. Armed conflict is a form of conflict that involves physical engagement with kinetic weaponry. Examples of armed conflict include two armies fighting with artillery, a non-state actor taking up small arms against a government, or drone strikes used to target an adversary's bases."

[64]The treatment language reads: "the Iranian army," "a non-uniformed insurgent group sponsored by Iran," or "a non-uniformed insurgent group."

Table 1. Treatments by factor.

Actor	Means
The Iranian Army	Kinetic
Iranian-Sponsored Insurgent Group	Cyber
Insurgent Group	

its actions. Moreover, Iran could publically contest the attribution of the attack, using the proxy as a shield, even if its involvement is only implausibly deniable. The third actor type is an insurgent group without clear state ties. Here, there is significant ambiguity about the presence and extent of state involvement, creating highly plausible deniability. That the group has "crossed the Somali border" suggests, in all cases, that the attacker is not indigenous to Somalia, such that there may be a state sponsor. Indeed, Online Supplemental Appendix 5 shows 58% of respondents guessed there was a state sponsor.

By varying the means and the actor behind the attack, we determine whether the attack is an example of conventional or gray-zone conflict.[65] Furthermore, we vary how plausibly deniable the attack is. The most deniable attacks use cyber means and insurgent groups without known state sponsors to conduct the attacks. Less-deniable attacks that still occur in the gray zone include cyberattacks conducted by the Iranian military and kinetic attacks conducted by a proxy organization with known ties to Iran. Our study design therefore allows us to investigate how (de) escalation occurs from a variety of gray-zone and non-gray-zone starting points, assessing how and why conflicts might move into or out of this zone.

While the focus on this singular scenario inherently limits our study, its selection is appropriate for three reasons. First, Iran boasts a strong military but is not a great or nuclear power. This means conventional escalation against Iran is plausible. Indeed, there have been both large-impact cyber and kinetic attacks against Iran in recent years, from the targeting of Qasem Soleimani to Stuxnet. If our aggressor were even more powerful than Iran, we might bias our theory in favor of limited-intensity retaliation. Second, Iran is one of few countries with robust, offensive cyber capabilities and one of very few that has, in recent years, sponsored proxies to directly attack US forces. This makes Iran a plausible candidate for our scenario. Third, US-Iran

[65]We are interested in the effects of deniability, not the process of attribution. Other approaches, such as indicating the degree of uncertainty about what actor was involved in the attack, would largely capture confidence in intelligence estimates. We argue deniability is inherent to proxy use, regardless of whether the connections between the proxy and state sponsor are unknown, known with low confidence, or known with high confidence. Previous work has examined the relationship between confidence in attribution and public support for retaliation: Hedgecock and Sukin, "Responding to Uncertainty."

competition is a prevalent feature of US politics and an important object of study in its own right.

The scenario includes a small number of casualties to set the stakes; the described outcomes reflect plausible real-world effects. An event of this scale would be notable and raise concerns, so respondents should take it seriously.[66] At the same time, the described attack is not so large as to require any sort of immediate, highly visible, or clearly widespread response.

Importantly, a decision about retaliation against an attack of this type is likely to originate at the level of officers in the field. Within a set of options constrained by the rules of engagement, lower-ranked officers are likely to be instrumental in deciding whether low-intensity retaliation should occur or whether high-intensity options should pass up the chain. (Such assignments are likely outcomes for cadets in the early years of their career.)

Upon reading the vignette, cadets are asked whether and how they would most prefer the United States to respond. This allows us to measure both the likelihood and intensity of escalation. Table 2 displays the response options, and cadets choose one. They can choose a high-intensity option, such as a kinetic operation, a lower-intensity option, such as a cyber operation, or a nonmilitary option, such as sanctions.[67] Cadets can also choose a high-intensity target (e.g., Iran) or a lower-intensity one (e.g., insurgents).[68] Respondents could alternatively choose not to respond.

[66]Public opinion research suggests aversion to civilian casualties, although force protection and strategic concerns may outweigh this: Janina Dill et al., "Inconstant Care: Public Attitudes Towards Force Protection and Civilian Casualties in the United States, United Kingdom, and Israel," *Journal of Conflict Resolution* 67, no. 4 (2023): 587–616; Janina Dill et al., "Kettles of Hawks: Public Opinion on the Nuclear Taboo and Noncombatant Immunity in the United States, United Kingdom, France, and Israel," *Security Studies* 31, no. 1 (2022): 1–31.

[67]Existing literature shows sanctions are used as a punishment for a wide variety of activities that would not reach the threshold to justify the use of force; sanctions can occur in concert with or as an alternative to military conflict in some cases (Robert A. Pape, "Why Economic Sanctions Do Not Work," *International Security* 22, no. 2 (1997): 90–136; William Kaempfer and Anton Lowenberg, "The Political Economy of Economic Sanctions," *Handbook of Defense Economics* 2 (2007): 867–911). The literature has generally not found sanctions to be effective deterrents or costly signals, unlike the use of force (T. Whang and H. J. Kim, "International Signaling and Economic Sanctions," *International Interactions* 41, no. 3 (2015): 427–52). Although sanctions can raise tensions, they are typically not seen as significantly escalatory unless used in conjunction with military action (A. C. Drury, "Sanctions as Coercive Diplomacy: The US President's Decision to Initiate Economic Sanctions," *Political Research Quarterly* 54, no. 3 (2001): 485–508). Relatedly, the use of sanctions has not, in and of itself, led to recourse to force. Note a military officer faced with this situation could not, on her own, choose to impose sanctions, but including this option allows us to assess whether cadets preferred a political or military solution to the crisis.

[68]If an insurgent group without a named sponsor carried out the attack, respondents are asked to predict whether the group had a state sponsor and then are offered a choice between retaliating against the expected sponsor or its proxy. Of the 245 respondents who received this treatment, 104 said they "do not think [the group] receives support from a state." The most common countries selected as possible sponsors were Iran (44 respondents) and China (29), followed by Russia (20) and Somalia (9). See Online Supplemental Appendix 5.

Table 2. Response options.

Action	Target
Kinetic Operation	State
Kinetic Operation	Insurgents
Cyber Operation	State
Cyber Operation	Insurgents
Sanctions	State
Sanctions	Insurgents
Denounce	–
Do Nothing	–

While a simplification of the universe of options available in the real-world, this menu includes military, economic, and diplomatic options. It allows us to assess who cadets think should be held responsible and how.

Results

In this section, we first show that cadets are more likely to prefer escalation in response to more deniable attacks. Second, cadets prefer less-intense responses to more deniable attacks. These two findings lend support to our theory, H_3, and highlight the importance of delineating between escalation likelihood and retaliation intensity. Third, we leverage open-ended questions to illustrate the logic of cadets' preferences.

Deniability Makes Escalation More Likely

First, we assess how deniability strategies affect the likelihood of escalation. Building on our prior conceptualization of escalation, we code retaliatory actions as escalatory if they cross a major threshold. In our study, this is observed through the retaliatory attack's target and means.

First, we code retaliations that match both the treated action type and actor as being proportional responses; for example, a cyberattack against an insurgent group that occurred as a response to a cyberattack by an insurgent group would be neither escalatory nor de-escalatory. Second, we code as escalatory attacks that cross the threshold from covert to overt along either the actor or means dimension. Thus, we code as escalatory retaliation to a cyberattack by a kinetic attack, and we code as de-escalatory retaliation to a kinetic attack by a cyberattack. Similarly, we code as escalatory retaliation to an attack from a proxy by targeting its state sponsor and as de-escalatory retaliation to an attack from a state by targeting its proxy.[69] We consider de-escalatory attacks that show movement from overt

[69]An action is coded as escalatory if the threshold is crossed from covert to overt along at least one dimension, even if there is movement along the other dimension from covert to overt. For example, if a response to a cyberattack by a state was to use kinetic force against a proxy, then that is coded as escalatory. Crossing the threshold is, in and of itself, an escalatory decision: Lin-Greenberg, "Evaluating Escalation."

to covert along one of the dimensions (actor or means), so long as the other dimension either remains static or also moves toward covertness.

An important feature of this coding scheme is that cadets receive the same menu of response options regardless of what treatment they received. As a result, there are more options available for de-escalation and fewer options for escalation when cadets consider retaliating to attacks occurring at higher rungs of the ladder. We argue that this is, itself, a feature of the decision-making space. The ladder is not infinite. Although we believe this is an important, real-world dynamic, it nevertheless presents a measurement challenge. That is, if our respondents were simply selecting their preferred retaliation option at random, we would expect to find evidence of a greater likelihood of escalation in response to gray-zone operations than kinetic ones. We thus take care to demonstrate below that our results are nonrandom. We also provide a more granular breakdown of cadets' preferences in the next section, which allows us to assess our arguments about (de)escalation against other operationalizations.

Figure 2 reveals that strategies of deniability increase the likelihood of escalation. A total of 40.7% of respondents escalated in response to cyberattacks, compared to just 5.0% for kinetic attacks. Similarly, 13.5% of respondents escalated in response to an attack by Iran, compared to 27.5% for attacks by insurgents.[70] The results support H_1 and H_3, which both predict an increased likelihood of escalation against deniable attacks. These findings contradict H_2, which predicts that deniability is de-escalatory.

Importantly, these results are nonrandom. For example, consider cadets responding to an attack by an Iranian-sponsored insurgent group that used cyber means. If cadets were choosing their preferred response at random, 43% should de-escalate, 14% should choose a proportional response, and 43% should escalate. Figure 2 shows that the actual proportions are 16%, 30% and 54%. That is, escalation in response to this highly deniable attack is substantially more likely than random. Similarly, consider cadets responding to an attack by an Iranian-sponsored insurgent group that used kinetic means. If acting randomly, 57% would de-escalate, 14% would choose a proportional response, and 29% would escalate. Instead, these proportions are 43%, 48%, and 9%. That is, escalation in response to this less-deniable attack is substantially less likely than random. For each treatment, the preferred responses systematically differ from those expected by randomness alone.

Table 3 demonstrates the escalatory potential of deniability, showing correlations between the treatments and a tripartite ordinal measure

[70]The inverse is also true: 26.1% de-escalated after cyberattacks, compared to 43.0% after kinetic attacks. 51.2% of respondents de-escalated after attacks by Iran, compared to 26.1% after attacks by insurgents.

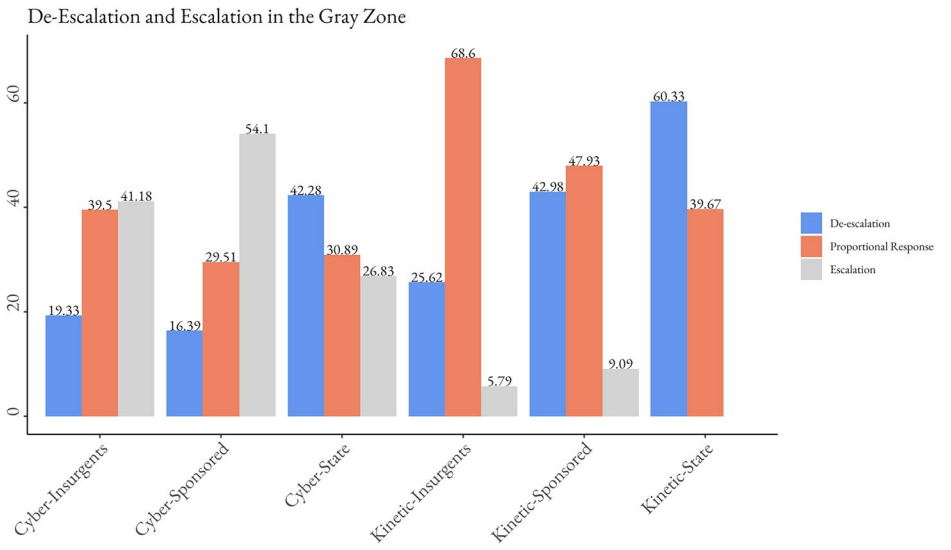

Figure 2. Deniability increases the likelihood of escalation.

of whether a cadet's preferred response to an attack was de-escalatory (−1), neither de-escalatory nor escalatory (0), or escalatory (1). We include a bare-bones model, as well as models that account for cadets' demographics, geopolitical views, and pretreatment perceptions of Iran.[71] We find cyber means increase the likelihood cadets prefer an escalatory response, as does using proxies. Thus, deniability does not reduce the chance of escalation, instead making it more likely that violence continues.

Table 3 shows cadets with globalist attitudes are less likely to support escalation.[72] More vengeful respondents are more likely to support escalation, as are cadets who view Iran less favorably.[73] Respondents' race, political party, and school year appear to have no effect.

We include alternative specifications in Online Supplemental Appendix 2. We use two alternate operationalizations of escalation that focus only on crossing a single threshold, for example, between the cyber and kinetic domains or between targeting a state versus a proxy. Our findings are robust across these alternate specifications.

[71]See Online Supplemental Appendix 6 for balance tables.
[72]61% agree the United States should "play an active role in solving conflicts around the world."
[73]61.5% believe "anyone who kills Americans deserve to be killed"; 47% support the death penalty; and 36% somewhat or strongly agree that the United States should use "enhanced interrogation" techniques, including waterboarding, on terrorists. The vengeance variable combines scores on these questions.

Table 3. Deniability increases the likelihood of escalation.

	Likelihood of escalation		
	Base	Demographic controls	Full model
	(1)	(2)	(3)
Cyber	0.528*** (0.050)	0.524*** (0.051)	0.524***(0.050)
Iranian-Sponsored Insurgents	0.399*** (0.061)	0.382*** (0.062)	0.368*** (0.061)
Insurgents	0.390*** (0.062)	0.411*** (0.063)	0.408*** (0.061)
Female		−0.194*** (0.058)	−0.111(0.059)
Black		0.002 (0.085)	0.047 (0.084)
Republican		0.095 (0.053)	0.036 (0.053)
Seniors		−0.044 (0.067)	−0.045(0.065)
Sophomores		0.095 (0.086)	0.118 (0.084)
Globalism			−0.061*** (0.016)
Vengeance			0.052*** (0.013)
Iran Favorability			−0.089* (0.043)
Constant	−0.643*** (0.050)	−0.632*** (0.063)	−0.301 (0.154)
Observations	727	696	696
R2	0.185	0.205	0.246
Adjusted R2	0.182	0.195	0.234
Residual Std. Error	0.677 (df = 723)	0.672 (df = 687)	0.656 (df = 684)
F Statistic	54.831*** (df = 3; 723)	22.095*** (df = 8; 687)	20.337*** (df = 11; 684)

Note: *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

Deniability Makes Escalation Less Intense

Although deniability generally increases the likelihood of escalation, it also lowers the intensity of retaliation. This is largely because cadets' preferred strategy often matches the means and/or actor of the initial attack. For example, the modal retaliation to a cyber operation is to target a proxy group using cyber means, while the modal retaliation to a kinetic attack is to target a proxy group using kinetic means.

Figure 3 shows the proportion of cadets in each treatment group who chose each possible retaliation.[74] We find illustrative evidence of a lower preferred intensity of retaliation under conditions of deniability, which is supported by inferential analysis in Table 4. Cadets do not appear to be escalating massively or inadvertently but instead choosing calibrated responses.

Table 4 further shows how the treatments affected the intensity of cadets' response preferences. We present four different models focusing on different modes and targets for retaliation. Model 1 assesses cadets' desire to use military—as opposed to economic or diplomatic—means to respond to the initial attack. For each treatment, a majority of cadets prefers a military response.

Model 2 focuses on whether cadets' preferred response used cyber means. Model 3 assesses preferences for retaliation against state actors. Models 2 and 3 show a moderate preference for in-kind responses,

[74]We show all response options, ordered loosely from least-intense (on the left) to most-intense (on the right). However, we remain agnostic to the specific ordering, as keen readers may disagree about when to prioritize certain thresholds over others. Thus, we ensure our interpretations are robust to alternate codings of (de)escalation, including those that focus only on the actor or means dimension.

Table 4. Treatments affect preferences for response mode and target.

	Preferred response type			
	Kinetic or cyber response	Cyber response	Kinetic or cyber against state	Kinetic against state
	(1)	(2)	(3)	(4)
Cyber	0.057 (0.032)	0.323*** (0.032)	0.058 (0.034)	−0.077** (0.027)
Iranian-Sponsored Insurgents	0.049 (0.040)	0.012 (0.039)	−0.316*** (0.038)	−0.208*** (0.030)
Insurgents	0.104** (0.040)	−0.017 (0.039)	−0.341*** (0.045)	−0.243*** (0.036)
Female	0.023 (0.038)	0.054 (0.038)	−0.081 (0.041)	−0.055 (0.032)
Black	0.033 (0.054)	−0.009 (0.054)	0.055 (0.059)	0.065 (0.046)
Republican	0.024 (0.035)	−0.015 (0.034)	0.033 (0.036)	−0.015 (0.028)
Seniors	−0.042 (0.042)	−0.002 (0.042)	−0.018 (0.045)	0.016 (0.035)
Sophomores	−0.019 (0.055)	−0.057 (0.054)	0.120* (0.056)	0.135** (0.044)
Globalism	−0.021* (0.010)	−0.0004 (0.010)	−0.027* (0.011)	−0.023** (0.008)
Vengeance	0.030*** (0.009)	−0.003 (0.008)	0.015 (0.009)	0.015* (0.007)
Iran Favorability	−0.066* (0.028)	0.070* (0.028)	−0.007 (0.030)	−0.046* (0.023)
Constant	0.772*** (0.101)	0.030 (0.099)	0.572*** (0.105)	0.470*** (0.082)
Observations	696	696	602	602
R2	0.055	0.143	0.161	0.156
Adjusted R2	0.039	0.129	0.146	0.141
Residual Std. Error	0.427 (df = 684)	0.421 (df = 684)	0.416 (df = 590)	0.325 (df = 590)
F Statistic	3.586*** (df = 11; 684)	10.339*** (df = 11; 684)	10.329*** (df = 11; 590)	9.945*** (df = 11; 590)

Note: *$p < 0.05$. **$p < 0.01$. ***$p < 0.001$.

stopping short of true firebreaks. Cyber operations were unlikely to be deployed as an off-ramp to kinetic conflict; only 12% chose a cyber response to a kinetic attack. Conversely, one-third of respondents preferred reacting to cyberattacks with kinetic means. This shows a substantial minority were willing to escalate by crossing the purported cyber firebreak.

The preference for in-kind behavior was also evident in the actor dimension. Compared to attacks conducted by the Iranian military, attacks conducted by proxies were less likely to prompt retaliation against states. For example, if the Iranian army conducted the initial attack, 27% of cadets supported a kinetic attack on Iran, but that decreases to 7% if a non-state actor with known ties to Iran conducted the attack. This is the case even though 71% of respondents in this treatment group think ties between Iran and the insurgent group are strong. When cadets read about an insurgent group without clear state ties, they also strongly prefer to retaliate against the insurgent group (72%) than against the state they thought most likely to have sponsored the proxy (8%).

Model 4 narrows in on the highest-intensity response option, a kinetic attack against a state actor. While Table 3 showed escalation was correlated with deniability, Model 4 in Table 4 shows very high-intensity retaliation is negatively correlated with factors that increase deniability.[75] In our experimental setting, we thus link deniable attacks to escalation but show that the preferred type of escalation is restrained.

[75]In Online Supplemental Appendix 2, we provide more detail on this "pure reciprocation" model, which distinguishes between military vs. non-military responses.

In Online Supplemental Appendix 2, we offer a more granular measure of intensity. First, we scaled the responses from 0 to 6, arranged from least-intense (i.e., no response) to most-intense (i.e., a kinetic attack against a state actor) in the order that appears in Figure 3. Second, we include a domain-focused measure of intensity that treats a diplomatic response as the least intense option, followed by economic, cyber, and then kinetic responses. These robustness tests generally show a preference for lower-intensity responses to more-deniable attacks, although we find this is primarily driven by a preference for in-kind responses along both the means and actor dimensions. Online Supplemental Appendix 4.4 examines perceptions of the aggressor's intentions, showing more deniable operations may have greater ethical and legal legitimacy; this could contribute to views that deniable operations signal less-escalatory intentions and demand less-intense responses.

Deterrence as the Framework for Escalation

We find cadets offer strategic, deterrence-based reasons for their choices to escalate or de-escalate, demonstrating an understanding of the strategic dynamics of movement along the escalation ladder.

The cadets echoed the logic of deterrence in their open-ended responses, explaining their preferred choice of retaliation. Table 5 shows some of the most common word stems, alongside their frequency.[76] Respondents appear to consider the escalatory potential of their responses, prioritizing limited options and using logic such as "an eye for an eye." Cadets appear to think about managing risk while signaling to adversaries to deter future aggression.

We find suggestive evidence that cadets consider the strategic dynamics of responding to deniable attacks. For example, cadets explain they seek to deter adversaries, while keeping the pace of their movement up the escalation ladder slow to avoid larger consequences. Cadets who supported targeting insurgents emphasized the need to deter, while taking slow and safe steps. One wrote: "Demonstrating power against the Iranian-backed group ensures that the U.S. is relatively safe from a further show of aggression; getting the point across in the safest manner." Another explained: "I don't believe it would be right to allow an attack like that to go unanswered, and directly attacking Iran … would have much broader

[76]The analysis omits common stop words. Included are items from the 50 most common word stems, leaving out the following descriptors of the scenario and retaliatory options: attack, iran, insurg, physic, cyber, respons, forc, conduct, conflict, american, iranian, action, soldier, u.s., respond, militari, unit, arm, retali, countri, troop, act, sanction, and oper.

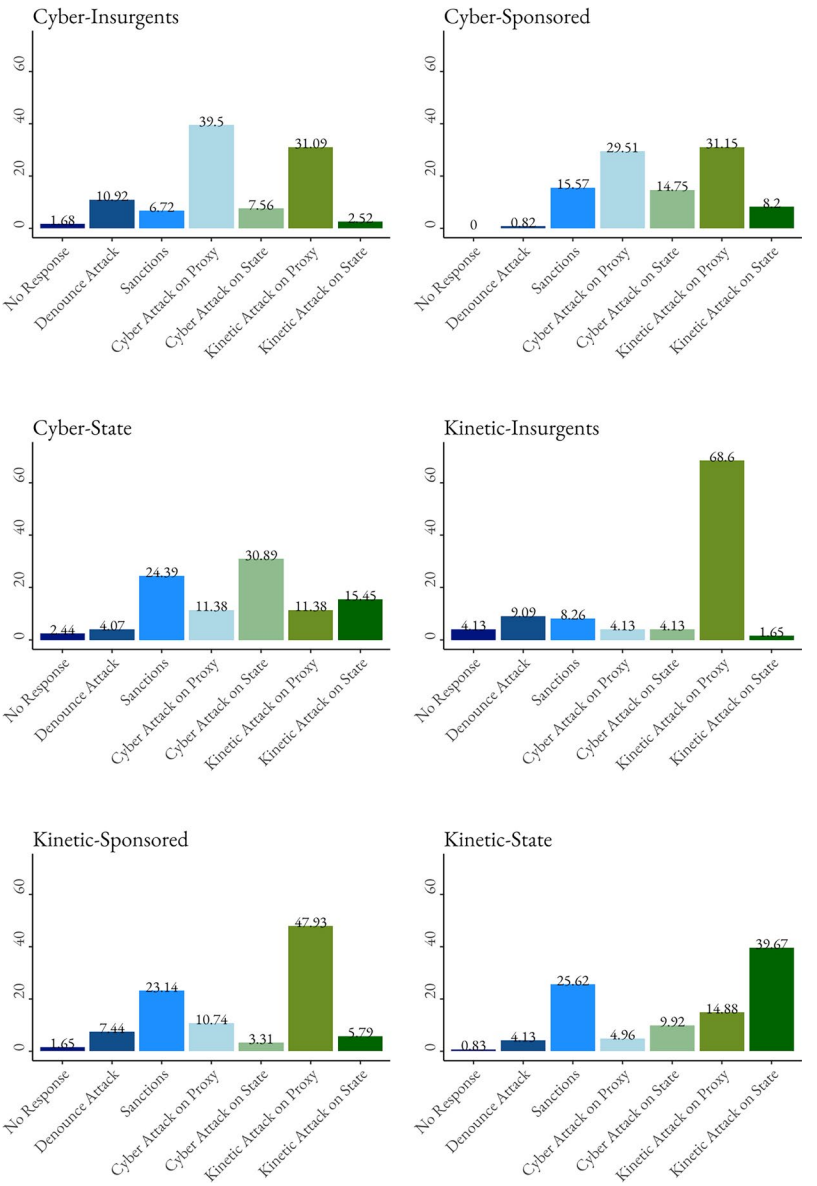

Figure 3. Deniability decreases the intensity of retaliation.

consequences than retaliating against the insurgent group. Attacking the insurgent group would also help to spread that it would be a bad idea to attack U.S. troops." One cadet remarked: "physically attacking Iran would cause way too much conflict in the long run [but] I do think it is important for smaller groups to know that what just happened was not okay." Another wrote: "We should assert dominance against insurgent groups by destroying them physically. Engaging in an attack against [the sponsor] could result in a larger conflict than we anticipated, but conducting an

Table 5. Occurrence of the most common word stems in cadets' explanations for their preferred response.

Word stem	*N*
proport	89
kill	63
direct	52
econom	52
war	50
escal	49
effect	44
civilian	39
harm	32
future	30
result	29
death	28
eye	25
send	23
casualti	22
justifi	22
messag	22
threat	22
live	21
violenc	21
damage	19
involv	19
lead	19
deter	18
feel	18
punish	18

attack on just an insurgent group would keep the conflict between us. My perception is that [the sponsor] does not back them enough to defend them, if we choose only to attack the insurgents." These comments support the logic of our argument, H_3, and directly contradict the logic of H_1, which instead argued that, faced with deniability, deterrence would demand holding the sponsor accountable and would therefore result in escalation.

Respondents offer a similar logic when addressing the tradeoffs between cyber and kinetic attacks, writing, for example: "more intense reactions, such as [a] physical attack, would only bring a harder counterattack" and "to do nothing lets them think it's okay to do it again. A physical attack might escalate the conflict too much, so a cyber attack … would be a preferred response." Other respondents wrote that the use of cyber operations would "provide condemnation without excessive risk of escalation," and that "the appropriate retaliation would be a cyber attack that may be more intense than the one they committed, however, to physically send armed troops would be less beneficial and more risky for the United States." These and other responses highlighted the need to limit the intensity of retaliation, while emphasizing the importance of retaliating against the attack. Striking this balance is at the core of our theory.

While a few respondents pointed to potential audience costs (in line with the de-escalatory hypothesis, H_2), these responses were relatively rare. For example, one cadet wrote: "Since Iran was not directly involved, it

would not be a wise choice to attack Iran directly." Another remarked that "while we know [the insurgents] are Iranian backed, we cannot necessarily prove this and therefore would be seen as 'in the wrong' if we attacked Iran. However, we can attack the insurgent group." Similarly, another cadet wrote: "I feel like everything we do is HIGHLY publicized because we are the U.S. and so to prevent an all out war I do not think we should do something to hold Iran accountable … just yet. If they continue with similar behavior th[e]n I do think we should take stronger actions."

Although these are just a sample of responses, they demonstrate the critical role of deterrence in cadets' understanding of gray-zone conflict. Cadets are careful to balance deterrence and escalation management concerns, and they emphasize that using gray-zone tactics—including targeting proxies and cyber operations—helps them achieve this balance.

Conclusion

As strategic uses of (im)plausible deniability become increasingly prevalent, our understanding of escalation management requires updating. We use a novel source of information—the views of US Military Academy cadets—to explore the strategic dynamics of deniability on escalation. We find that, when faced with more deniable attacks, cadets are more likely to support escalation, but the intensity of their preferred escalation remains limited.

However, our research focuses on just two types of gray-zone strategies. Future research may wish to explore other tactics states use to operate outside of the bounds of traditional warfare. Our study is also, importantly, limited by its geographical focus. Future research may wish to explore whether other state-and-proxy combinations, such as Russian-sponsored operatives working in Eastern Europe and Northern Africa, or Chinese-sponsored organizations in Southeast Asia, might operate differently. Relatedly, future research may wish to examine how the nature of the adversary (e.g., their military capabilities, nuclear status, regime type, etc.) affects escalation dynamics. Future research may also wish to explore the views of other audiences, such as the public or civilian elites, and to explore how various audiences conceptualize the order of the escalation ladder to better understand how intensity is conceived in the real world. Our sample of cadets shares similar demographic characteristics with the officer corps,[77] meaning certain groups are over-represented relative to

[77] See Online Supplemental Appendix 3.

civilian officials. Despite these limitations, our results serve as a theory-building exercise, developing a new perspective on the (de)escalatory potential of deniability and offering an exploratory look into (de)escalation in the emerging gray zone.

Our findings have significant political implications. First, deniability may be an effective shield. States using high-deniability tactics may shift the costs of escalation onto proxies or otherwise keep conflict constrained to the gray zone. While gray-zone operations are likely to escalate, the character of that escalation may be limited. Second, our findings weigh in on an emerging strategic conversation about escalation management. We suggest escalation management is possible in the gray zone, and that the options to target periphery actors, such as proxies, or to use covert means, such as cyber operations, can help limit escalation intensity. Third, our findings raise questions about deterrence in the gray zone. The same deniable strategies that reduce the intensity of competition may also be more difficult to deter. In sum, our research highlights how deniability shapes the strategic landscape of the gray zone by enabling escalation without necessarily inviting high-intensity actions.

Acknowledgments

For comments and suggestions, the authors thank Victor Asal, Janina Dill, Jacquelyn Schneider, Erik Gartzke, and Stacie Goddard as well as the participants at the U.C. San Diego Cyber Escalation in Conflict Workshop, ISA 2023, APSA 2023, and Brown University Coercion Workshop. The authors thank the editorial team and two anonymous reviewers for their helpful feedback.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Funding

For their support, the authors thank the London School of Economics and Political Science, Charles University (UNCE 24/SSH/018, Peace Research Center Prague III), and the Carnegie Corporation of New York (G-PS-24-61252). The statements made and views expressed are solely the responsibility of the authors.

ORCID

Lauren Sukin http://orcid.org/0000-0002-5775-8790
Kathryn Hedgecock http://orcid.org/0000-0002-8667-7439