Wargame of Drones: Remotely Piloted Aircraft and Crisis Escalation

Citation: Lin-Greenberg, Erik. 2022. "Wargame of Drones: Remotely Piloted Aircraft and Crisis Escalation." Journal of Conflict Resolution, 66 (10).

As Published: 10.1177/00220027221106960

Publisher: SAGE Publications

Persistent URL: https://hdl.handle.net/1721.1/148674

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Wargame of Drones:
Remotely Piloted Aircraft and Crisis Escalation

Forthcoming, *Journal of Conflict Resolution*
May 2022

Erik Lin-Greenberg[1]
eriklg@mit.edu

ABSTRACT

How do drones affect escalation dynamics? The emerging consensus from scholarship on drones highlights increased conflict *initiation* when drones allow decisionmakers to avoid the risks of deploying inhabited platforms, but far less attention has been paid to understanding how drones affect conflict *escalation*. Limited theorization and empirical testing have left debates unresolved. I unpack the underlying mechanisms influencing escalation decisions involving drones by proposing a logic of *remote-controlled restraint*: drones limit escalation in ways not possible when inhabited assets are used. To test this logic and explore its instrumental and emotional microfoundations, I field "comparative wargames." I immerse national security professionals in crisis scenarios that vary whether a drone or inhabited aircraft is shot down. I validate wargame findings using a survey experiment. The wargames shed light on the microfoundations of escalation, highlight limits of existing theories, and demonstrate the utility of comparative wargaming as an IR research tool.

Keywords: politics of emerging technologies, escalation, interstate crises, wargame, drones

[1] Nandita Balakrishnan, Richard Betts, Abhit Bhandari, Nick Blanchette, Matthew Caffrey, Matthew Carr, Jonathan Caverley, Jonathan Chu, Jasper Cooper, Matthew Daggett, Alexander de la Paz, Lynn Eden, Tom Fox, Eric Frahm, Melissa Griffith, Kolby Hanson, Ben Harris, Shigeo Hirano, Michael C. Horowitz, Robert Jervis, John Kuconis, Jeffrey Lax, DoYoung Lee, Asfandyar Mir, Theo Milonopoulos, Lama Mourad, Reid Pauly, Lauren Pinson, Trevor Prouty, Tonya Putnam, Steven Saideman, Usha Sahay, Jacquelyn Schneider, Tara Slough, Steven Tate, Rachel Whitlark, Erin York, Fatima tuz Zahra Amjad, and Sherry Zaks generously provided helpful feedback on earlier stages of this project and/or assisted with facilitating the wargames. This project received generous support from the Eisenhower Institute, the Horowitz Foundation for Social Policy Research, and the Smith Richardson Foundation. All errors are my own.

In October 2015, Turkey shot down a Russian drone that had strayed into Turkish airspace. Russia publicly ignored the incident and took no observable retaliatory measures (Coskun 2015). In stark contrast, after Turkey downed an inhabited Russian jet just one month later, Russia launched airstrikes on Turkish-backed rebels and supply convoys in Syria (Bertrand 2015; Lowen 2015).[2] What factors—instrumental and emotional—underpin the radically different responses to the loss of drones and inhabited platforms? More broadly, how do remote warfighting technologies affect decisions on escalation during interstate crises?

Prominent international relations (IR) theories expect belligerents to initiate conflict when technology reduces the costs and risks of military operations (Jervis 1976; Van Evera 2001). Accordingly, the emerging consensus from scholarship on drones highlights increased conflict *initiation* when drones allow decisionmakers to avoid the risks of deploying inhabited platforms. Far less attention has been paid to understanding and empirically testing how drones affect conflict *escalation*. The limited scholarship addressing the link between drones and escalation is largely conceptual. It leaves unexplored the underlying mechanisms, provides limited empirical evidence, and yields mixed results about whether drones contribute to escalation (Horowitz, Kreps, and Fuhrmann 2016; Mahnken, Sharp, and Kim 2020; Lyall 2020).

I intervene in these debates by developing a logic of *remote-controlled restraint* and test it using an innovative methodological approach: comparative wargaming.[3] The logic expects that when used as a substitute for inhabited assets during interstate crises, drones (and other remotely operated weapons) can help prevent crises from escalating into broader conflicts in ways not possible when inhabited platforms are used. Technologies that allow decisionmakers to conduct operations without putting friendly personnel in harm's way reduce the potential human costs of

[2] I use "inhabited" to describe platforms with crewmembers onboard.
[3] Bartels (2018) describes the use of wargames as "structured comparisons."

military action. On one hand, this can mitigate the political obstacles associated with dispatching troops, enabling states to deploy drones more often than inhabited assets (Kaag and Kreps 2014). On the other, drones may ameliorate subsequent escalation spirals. Relative to the loss of an inhabited platform to hostile action, the loss of a drone should be less escalatory for two reasons. First, it is less likely to elicit an instrumental desire to degrade a rival's military capabilities. Second, it is less likely to trigger emotions like anger that contribute to aggressive, risk-acceptant behavior. In turn, leaders should take more restrained retaliatory measures following an attack on a drone. Since a state's response to an attack is a key determinant of whether a crisis escalates, drones can alter the dynamics of crisis escalation.

To test remote-controlled restraint, I field original comparative wargames, which merge a military tool (wargaming) with social science techniques (case study and experimental research design). I recruit national security practitioners to participate in scenario-based exercises that feature the shootdown of a U.S. military aircraft. In the games, I randomly vary whether the aircraft is inhabited and ask participants to formulate a response plan. By holding all other scenario elements constant and repeating the wargames several times with different participants, I create "control" and "treatment" games to qualitatively explore whether and why drones help limit escalation. The wargames, which I validate using survey experiments fielded on military officers, provide support for the logic of remote-controlled restraint.

The article makes three contributions to the study of emerging technologies and IR. First, the findings advance theories that explain the relationship between technology and conflict. Existing scholarship focuses on conflict onset, without fully considering how technology affects escalation once forces are deployed. Moreover, most theories that explicitly link technology and escalation were developed during the Cold War (Schelling 1966; Posen 1991). As a result, they tackle questions related to nuclear escalation or large-scale conventional conflict, but do not assess the role of more recent emerging technologies. Second, the project contributes to the burgeoning research on escalation by exploring the microfoundations that underpin escalation dynamics.[4] Unpacking the instrumental and emotional factors that drive retaliation sharpens our understanding of escalation. Third, the project contributes to IR methodology by advancing the use of wargames to study decision-making in contexts where real-world data are scarce (Pauly 2018; Reddie et al. 2018; Schechter, Schneider, and Shaffer 2021).

TECHNOLOGY AND ESCALATION

Scholars and policymakers argue that the seemingly "costless" nature of drones makes it easier for leaders to resort to military force (Kaag and Kreps 2014). These claims align with prominent realist IR theories that suggest the use of force becomes more likely when technologies are thought to make offensive action less costly (Jervis 1976; Van Evera 2001). They also reflect rational choice logics in which actors initiate and continue conflicts so long as the expected benefits of victory are perceived to exceed the cost of fighting (Mesquita 1983; Fearon 1995). These logics explain the conditions under which technology might lead decisionmakers to initiate the use of force, but yield fewer predictions about subsequent escalation.[5]

Drones and Crisis Escalation

Drones provide an ideal case to explore how emerging technology can affect escalation. These combat and surveillance systems are remotely operated, without crewmembers physically onboard, and are widely deployed during international crises. Indeed, aerial drones are now operated by over 80 countries on missions including close air support and reconnaissance

[4] On microfoundations, see (Kertzer 2017).
[5] Scholarship that explores escalation or the intensity of fighting (Fearon 1994; Weisiger 2016) does not explicitly assess technology's role.

(Horowitz, Schwartz, and Fuhrmann 2020). Although current generation drones are frequently deployed alongside inhabited systems, they are increasingly deployed as substitutes, including on missions against near-peer rivals (Cenciotti 2015). Stealthier and more maneuverable drones under development will allow drones to substitute for inhabited platforms on an expanded set of missions (Clark 2019).

Drones are thought to reduce the potential financial and human costs of military operations. From a financial standpoint, many drones have lower unit costs than inhabited assets. For instance, an inhabited MC-12 reconnaissance plane costs $17 million, compared to just over $4 million for a similarly capable MQ-1 Predator drone (U.S. Air Force 2017; 2010).[6] While treasure is a consideration, decisionmakers likely weigh blood more heavily. Existing scholarship suggests that casualties during military operations are particularly salient to policymakers because attacks on friendly personnel generate more visceral emotional reactions (McDermott, Lopez, and Hatemi 2017) and can reduce public support for leaders who backed military action (Mueller 1973; Reiter and Stam 2002).[7] Indeed, when Iran downed a $130 million U.S. Navy drone, President Trump did not let the drone's price tag drive his decision-making. Instead, he highlighted the lack of casualties: "We had nobody in the drone.…It would have made a big, big difference (Trump 2019)."

Because of the lower costs associated with drone operations, scholars commonly argue that drones expand the range of issues for which states are willing to deploy force (Kaag and Kreps 2014; Boyle 2020). Scholars find U.S. presidents (Macdonald and Schneider 2017), national security bureaucrats (Schulman 2018), and the American public are more likely to support the

---

[6] Gilli and Gilli (2016) argue drone operation costs include more than a drone's unit cost due to support personnel and infrastructure. When shot down, however, only the unit cost should matter since only the drone is lost.
[7] Some studies suggest policy objectives and likelihood of battlefield success matter more to the public than casualties (Gelpi, Feaver, and Reifler 2009).

deployment of drones than inhabited platforms (Kreps 2014; Walsh and Schulzke 2018).[8] Focusing on how drones affect the initial use of force is important, but overlooks whether drone use influences the subsequent arc of a crisis.

Yet, just as technology affects initial decisions on the use of force, it should also influence escalation–the patterns by which crises expand in intensity or scope (Kahn 1965; Morgan et al. 2008). Decisionmakers can climb the "escalation ladder" by increasing the intensity of actions vis-à-vis a rival (vertical escalation), by broadening the affected geographic area (horizontal escalation), or by doing both simultaneously (Kahn 1965, 3–5). According to Thomas Schelling, each of these actions represents a crossing of thresholds that "distinguish new activity from more of the same activity (Schelling 1966, 135)." The potential for escalation is particularly acute during interstate crises, situations that lie at "the nexus between peace and war" (Snyder and Diesing 1977, 3) where a change in the type and intensity of interaction between states can dictate whether tensions de-escalate or spiral into conflict (Brecher 1993, 3).

While technology is an important factor underlying how crises unfold (Jervis 1976; Creveld 1991; Cohen 1996; Horowitz 2020), "weapons don't make war (Gray 1993)," and technology's effect on escalation is not deterministic. Instead, technology is an instrument of policy that "will be deployed to different effects in different cultural and organizational settings (Gusterson 2016, 92)." Put differently, decisionmakers mediate the effect of technology on conflict and escalation; how they use and react to new technologies can shape a crisis's trajectory (Biddle 2004; Lieber 2008; Talmadge 2019).

[8] Beyond conflict initiation, drone scholarship explores ethics (Kaag and Kreps 2014), counterterrorism and counterinsurgency operations (Mir 2018; Mir and Moore 2019; Macdonald and Schneider 2019), and proliferation (Horowitz, Schwartz, and Fuhrmann 2020).

Escalation should therefore be conceptualized as an action-reaction process in which perceptions of one state's moves trigger the other state's response (Smoke 1977, 36; Morgan et al. 2008, 11). Any response depends largely on decisionmakers' perception of what thresholds the other actor crossed (Smoke 1977, chap. 2). For instance, using certain weapons may be seen as crossing a more escalatory threshold (Tannenwald 1999). Or, attacks on certain targets might be perceived as falling at a higher threshold, triggering a more significant response (Posen 1991).

Escalation is controlled when crisis behavior remains at low-rungs on the ladder (Smoke 1977, chap. 3; Kahn 1965, 3–9). Escalation control, however, does not require a complete absence of conflict. Actors might engage in low-scale skirmishes or covert actions, but these interactions generally involve lower costs than more escalatory, conventional conflict (Carson 2018).

Investigating whether emerging technologies affect escalation is critical to understanding today's security environment. Despite widespread drone proliferation, most existing scholarship linking technology and escalation focuses on nuclear crises (Talmadge 2017; Lieber and Press 2017; Acton 2018) or cyber operations (Valeriano, Jensen, and Maness 2018; Kreps and Schneider 2019). The more limited research on drones and escalation has been largely conceptual (Horowitz, Kreps, and Fuhrmann 2016; Boyle 2020) or focuses on drones as tools of crisis signaling (Zegart 2020).

This limited attention means that debates about the impact of drones on escalation remain unresolved. Some scholars suggest drones make it easier to escalate frozen conflicts (Lyall 2020). Under this logic, low-cost drones operating in support of traditional assets help states threaten punishment, creating incentives to break stalemates.

Other analyses suggest drones limit crisis escalation. Mahnken, Sharp, and Kim (2020) propose a "deterrence by detection" model in which drones stabilize crises by minimizing the information asymmetries that often lead to conflict (Fearon 1995). Specifically, drones identify information about adversary mobilizations that allow states to better posture forces and deter rivals. This logic, however, focuses on deterring rival action without fully theorizing about how drone losses during these operations might influence escalation. Horowitz, Kreps, and Fuhrmann (2016) explicitly address drone losses, but their analysis yields mixed findings and stops short of exploring underlying mechanisms. On one hand, they explain that ambiguous policies governing responses to rival drone operations could "lead to mutual misunderstandings and further escalation (Horowitz, Kreps, and Fuhrmann 2016, 28)." Yet they also acknowledge that drone losses need not result in escalation. Given the increasing role of remote warfighting technologies in modern conflict, more nuanced theorization and empirical analysis are needed.

Remote-Controlled Restraint

During a crisis, decisionmakers can take several actions involving drones that result in escalation. Leaders, can intensify a crisis by using drones to strike new targets. Or, they might need to decide whether to aggressively retaliate or exercise restraint after rival attacks. I focus on retaliation because states are increasingly taking hostile action against their rivals' drones, and a state's response to shootdowns ultimately shapes whether an incident de-escalates or escalates.

Scholars argue that instrumental and emotional microfoundations inform decisions on retaliation. If decisionmakers perceive a drone loss as less costly than an inhabited asset loss, they should feel less need to degrade a rival's warfighting capabilities and less visceral emotional reactions. Drawing from the language of Richard Smoke, an early escalation scholar, attacks on drones may not meet a "saliency criterion," the threshold that separates actions that warrant escalatory retaliation from those that do not (Smoke 1977, 32–33).

According to instrumental logics, decisionmakers respond to attacks for utilitarian reasons. These instrumental drivers are *future*-oriented, rational efforts to prevent further aggression (Schelling 1966; Liberman 2006, 696; Löwenheim and Heimann 2008, 686). To prevent future harm, decisionmakers attempt to alter a rival's behavior by threatening future damage and signaling the "power to hurt (Schelling 1966)," or by bolstering defenses that deny a rival the ability to conduct future attacks (Snyder 1959). These actions often follow a tit-for-tat logic that inflicts a similar amount of harm to that of the initial transgression (Schelling 1966; Axelrod 2006).

Under this tit-for-tat framework, decisionmakers may view some attacks as so inconsequential that a forward-looking preventative response is not seen as warranted. As the perceived harm increases, decisionmakers should be more willing to ratchet up the intensity of their response to prevent future harm. Indeed, Schelling suggested retaliation be unambiguously connected to and proportional to the rival's initial actions to be effective (Schelling 1966, 146–47). Decisionmakers acting instrumentally should be reluctant to exceed tit-for-tat measures as doing so might elicit a larger counter-reaction that risks unwanted escalation.

Beyond the instrumental goal of preventing future harm, decisionmakers also confront potential political and reputational costs for an insufficient response to an attack. In the face of losses that are perceived as more costly, the public may call for more assertive retaliation (Jervis 1995, 24). Policymakers may respond to these demands to prevent the political repercussions of appearing weak or incompetent (Gelpi and Grieco 2015; Tomz, Weeks, and Yarhi-Milo 2020). While domestic political concerns are more likely to be observed among elected officials, national security practitioners may worry that letting an attack go unpunished will sully a state's reputation for resolve and weaken its bargaining position during future crises (Huth 1997; Weisiger and Yarhi-Milo 2015). These instrumental foundations lead to the first testable hypothesis.

> ***H*$_1$*:* National security decisionmakers are less likely to describe a need to significantly degrade a rival's warfighting capability or to protect a state's resolve after losing a drone to enemy activity than after losing an inhabited asset.*

Emotions can also shape retaliatory behavior (Crawford 2000). These emotional drivers can operate alongside instrumental ones by coloring the lens through which actors interpret events, or by acting as a "switch" that influences decisionmakers' actions (Mercer 2010; Petersen 2002, 3; Lerner et al. 2015). Unlike forward-looking instrumental drivers, emotional factors are typically retrospective. They tend to be aimed at avenging a wrong, rather than preventing future harm (Löwenheim and Heimann 2008).

Decisionmakers' actions depend in part, on the discrete emotions they experience. The type and intensity of emotions is shaped by the events that elicit them (Frijda 1986, 34). Acts of aggression, for example, often generate feelings of anger toward the wrongdoer, and this anger is often more intense when the transgression triggers higher levels of moral outrage or humiliation (Löwenheim and Heimann 2008; McDermott, Lopez, and Hatemi 2017).[9] In contrast, an event that is perceived as less of an affront is likely to generate less anger.

Anger is thought to generate desires for revenge and retribution against the offending actor (Liberman 2006; Löwenheim and Heimann 2008). A harmed actor launches punitive actions to teach its rival a lesson and to correct past harms by inflicting suffering (Löwenheim and Heimann 2008). This quest for vengeance can lead decisionmakers to be less sensitive to the costs they incur when carrying out reprisals, potentially risking further escalation (Löwenheim and Heimann 2008, 691–92). Indeed, experimental research has found that anger leads to more risk-acceptant behavior when avenging past harms (Lerner et al. 2003).

---

[9] Severe acts of aggression (e.g., attacks on the homeland) may elicit hatred (McDermott, Lopez, and Hatemi 2017, 83). I focus on anger since the wargame features attacks on deployed aircraft.

H*$_2$*: *National security decisionmakers are less likely to describe a need to punish a rival, seek revenge, or engage in risk-acceptant behavior after losing a drone to enemy activity than after losing an inhabited asset.*

The lower costs and weaker emotional reactions associated with a drone loss should leave decisionmakers less compelled to launch escalatory responses. In line with Carson's (2018) argument that escalation can be controlled by keeping activities out of public view, decisionmakers might simply disavow drone losses in a way that is harder to do when inhabited assets are attacked. Since friendly personnel are not captured or killed when a drone is downed, their loss is easier to keep hidden. As one analysis of Cold War drone operations noted, "A name plate [on a drone] doesn't receive the same public attention as an American pilot in prison (Wagner 1982, 78)." Dead or imprisoned personnel preclude this deniability, ratcheting up tension and calls for action, like the massive U.S. military mobilization after North Korea killed 31 Americans in the 1969 shootdown of a U.S. reconnaissance plane (National Security Agency 1989).

Yet even if a drone shootdown is in full public view, decisionmakers can publicly acknowledge the loss without launching an escalatory response. This sort of restrained retaliation unfolded when the United States responded to Iran's 2019 shootdown of a U.S. Navy drone with a cyber, rather than kinetic, attack (Barnes 2019). Decisionmakers may believe, as President Trump reportedly did, that retaliatory action that risks major escalation is not proportionate to the loss of a machine.[10] If a response does occur, it may be more muted than those following the loss of inhabited assets. These limited or non-responses can control escalation or provide off-ramps that deescalate crises. Replacing friendly personnel with machines can therefore shift escalation dynamics, leading to the *remote-controlled restraint* hypothesis:

H*$_3$*: *National security decisionmakers will take less escalatory retaliation measures after losing a drone to enemy activity during an interstate crisis than after losing an inhabited asset.*

[10] Twitter, @realDonaldTrump, 21 June 2019.

To be sure, there are limits to remote-controlled restraint. Repeated attacks on drones could lead to more escalatory responses, particularly if mounting losses significantly degrade a state's ability to conduct military operations. The willingness of a state to endure losses will vary, but states with smaller drone fleets may reach this point more quickly than states with larger arsenals. All else equal, however, attacks on drones are less likely to lead to escalation than attacks on inhabited aircraft.

Skeptics might also suggest that leaders "select into" non-escalatory situations by limiting drone deployments to low-stakes missions where escalation is unlikely. The empirical record suggests otherwise, with drones deployed on operations involving peer competitors and critical national issues. China uses drones to monitor contested territory in the East and South China Seas, and the United States has deployed drones to the Baltics and Central Europe (Panda 2015; Rempfer 2018). Moreover, the interactive nature of escalation makes it difficult for leaders to select into non-escalatory situations. Even if one actor seeks to avoid escalation, a rival might take actions that threaten escalation.

METHODOLOGY

Wargaming as a Research Tool

To test remote-controlled restraint, I create comparative crisis scenarios by embedding manipulations in wargames played by national security professionals. Drawing insights from wargames—"simulation[s] of…a conflict situation in accordance with predetermined rules, data and procedures [that] provide decision-making information (UK MoD 2017, 10)"—helps overcome many obstacles associated with studying the strategic implications of emerging military technologies. The relatively short operational history of new systems provides few cases to analyze, and data that exist are often classified and unavailable to researchers. Even when data are made public, researchers confront the "fundamental problem of causal inference": they cannot see how a specific event plays out both with and without the involvement of new technology (Holland 1986).

Because wargames are interactive events that simulate realistic scenarios, they offer a reasonable proxy to study real-world military operations (Lin-Greenberg et al. 2022). Just as wargames provide governments with a tool to prepare decisionmakers for crises, they also provide researchers with a unique data source. Wargames ask expert participants to collaborate and employ the same expertise and thought processes they would apply during actual contingencies.[11] Even though there are no real-world consequences when wargames go awry, elite participants understand they might one day conduct operations that resemble those they previously wargamed, and generally take wargames seriously (Perla 1990, 7). Most importantly for researchers, wargames generate two types of analyzable data: outcome data that identify how leaders might act in a situation and deliberative data, a rich narrative of player interactions that reveals the assumptions, logics, and thought processes that shape decisions (UK MoD 2017, 5).

Since wargames can approximate actual crisis decision-making settings, IR scholars have increasingly leveraged wargames for research. Some projects have used reports from historical wargames as archival sources (Pauly 2018; Schneider 2017). Other researchers have fielded their own wargames to generate data on nuclear and cyber warfare (Reddie et al. 2018; Jensen and Valeriano 2019; Schechter, Schneider, and Shaffer 2021).

Like any research approach, wargames have limitations. First, wargames are not predictive. They can reveal that a certain action is plausible, but cannot definitively predict the course of an

[11] Some studies suggest that non-elites behave differently than elites (Mintz, Redd, and Vedlitz 2006). Others suggest the behavioral gap between elites and non-elites is overstated (Kertzer 2020).

actual crisis given the role of chance and the decisions of specific participants (UK MoD 2017, 12–13; Bartels et al. 2019). Second, wargames are simplifications of reality (Sabin 2014, 19–30). They feature hypothetical situations, low stakes, and compressed timelines, which might lead participants to be more risk-acceptant during games than during real-world contingencies (McHugh 1966; Sabin 2014, 5). To lessen the limitations and generalizability concerns of drawing conclusions from a single game (Hermann 1967), fielding multiple games generates a spread of outcomes that reveals trends in decision-making (UK MoD 2017, 55).

The most significant drawback of traditional wargames for scholars is that wargames are generally not designed to assess how variation in a specific factor—or variable—of interest affects decision-making. Most wargames present all participants with the same information at the same time, preventing researchers from isolating effects of specific variables.[12] When wargames do include variation, manipulations are often introduced to all players in subsequent rounds of a game, or across different games held at different times.[13] This approach limits the ability to hold constant potentially confounding factors, like real-world events outside the wargame, that might influence decision-making.

To more systematically control for potential confounders, I draw from experimental and case study research design to field *comparative wargames*. I randomly assign participants with similar backgrounds into parallel wargames that are identical across all dimensions, except for a specific variable of interest (i.e., the involvement of drones or inhabited aircraft). By manipulating only the variable of interest, I generate "control" and "treatment" games to study how the presence or absence of drones affects escalation in a hypothetical crisis. To identify trends, I field several game iterations in close temporal proximity. I consider each game a unit of observation. Treating

[12] Bartels (2018, 6) describes exceptions. A notable exception is (Schechter, Schneider, and Shaffer 2021)
[13] Schneider (2017) introduces variation across rounds.

individual players as units of observation would generate more data, but would be less meaningful for hypothesis testing since crisis decision-making is a collaborative process.

Comparative wargaming follows elements of the experimental turn in IR research. On one hand, wargames expose participants to controlled environments that generate internal validity. These environments also more deeply immerse subjects than the survey experiments increasingly used for IR research and better mirror actual crisis decision-making settings, potentially bolstering external validity.[14] Instead of completing surveys individually, wargame participants have face-to-face interactions, debate decisions, and face stresses induced by time constraints, creating what McDermott (2002) terms "experimental realism." Indeed, Thomas Schelling, who planned many government wargames, explained that it is difficult for participants to be immersed in a wargame scenario "without its beginning to seem either real or as one that could be real (Department of Defense 1966, D3)."

On the other hand, wargames typically lack the large sample size of experimental work, preventing statistical inference. Ideally, a researcher would field hundreds of games to ensure that differences in outcomes are the result of experimental manipulations rather than chance or participant-specific factors. Recruiting for large-scale fielding, however, is challenging given the demanding schedules of national security practitioners. The wargames presented here involved 28 expert participants playing seven games. Recent projects have generated larger samples using students and online wargamers, enabling statistical analysis, but typically without the benefits of in-person expert interaction (Goldblum, Reddie, and Reinhardt 2019).

Comparative wargaming compensates for the limitations of small samples by producing qualitative insights on mechanisms like those associated with comparative case studies (George

[14] Scholars find that immersive scenarios can trigger cognitive processes that mirror those in real world settings (Daniel and Musgrave 2017).

and Bennett 2005; Bartels 2018). Like case study research, comparative wargaming allows for cross-case analysis that identifies trends across small numbers of observations. It also permits within-case process tracing by analyzing participant discussions in a single game. Unlike traditional case studies where researchers generally cannot control for factors other than an independent variable of interest, the experimentally-inspired manipulations in comparative wargames more systematically control for potential confounders.

Research Design

To test remote-controlled restraint and to explore the microfoundations of escalation, I embed manipulations into a series of one-sided seminar wargames that vary whether a drone or inhabited aircraft is shot down during a hypothetical interstate crisis. Seminar wargames are simulations where participants respond to pre-scripted scenarios rather than playing against an opposing team (UK MoD 2017, 39). The one-sidedness allows researchers to avoid adjudicating outcomes between competing teams, a process that can introduce researcher bias.[15]

My wargame features a crisis between two fictional states, Dakastan and Katunia. I use fictional, rather than actual, states to ensure military officers can participate without fear of divulging classified information. To be sure, critics may worry the scenario's hypothetical nature could lead participants to take the game less seriously, propose unrealistic actions, or make assumptions about the rival's identity and capability. To counteract this risk, I enhanced realism and decrease abstraction by providing players significant background information about Dakastan and Katunia's military capabilities, political situation, and diplomatic partners prior to the game.

Dakastan is a U.S. ally engaged in a territorial dispute with neighboring Katunia. Dakastan is increasingly subjected to attacks by Katunian-backed forces—including an incident in which

[15] This does, however, reduce the realism of interaction between forces.

Katunia shoots down a U.S. reconnaissance aircraft.[16] Control teams received a vignette that features the loss of an inhabited aircraft, while treatment teams received an identical vignette, except a drone is lost. Teams represented U.S. military planning cells at a regional air operations center; they were asked to develop a response plan for a senior military commander. Although the vignette focused on a seemingly tactical decision, tactical actions can escalate into broader conflicts. During the Cuban Missile Crisis, for instance, Soviet officers in Cuba shot down a U.S. aircraft without authorization from Moscow, nearly triggering U.S. strikes on air defense sites on the island.[17]

Participants – individuals with military experience – were simultaneously randomly assigned to three to five member teams and to either a control or treatment game.[18] The team size was selected to maximize the number of teams, while ensuring enough participants on each team to develop and debate potential responses. The teams capture several characteristics of military organizations. First, participants represented multiple ranks and, reflecting military hierarchy, teams generally included fewer senior than junior personnel. Second, participants' military experience provided an understanding of military capabilities and of principles like operational risk that non-experts (e.g., student samples) likely do not have (Mintz, Redd, and Vedlitz 2006; Oberholtzer et al. 2019). Finally, half of participants had air defense or drone operations experience. These participants acted like subject matter experts on real-world planning teams, answering questions about tactics and capabilities.

[16] Vignettes in Online Supplement A.

[17] President Kennedy stated that if additional planes were fired on, "we should take out the SAM sites in Cuba by air action." See Minutes of the Ninth Meeting of the Executive Committee of the National Security Council; 27 October 1962; JFK Presidential Library and Museum, https://microsites.jfklibrary.org/cmc/oct27/doc3.html. Robert Kennedy reportedly used a veiled threat of retaliatory strikes to drive negotiations that ended the crisis. See Dobrynin Cable to the USSR Foreign Ministry; 27 October 1962; National Security Archives; https://nsarchive2.gwu.edu/nsa/cuba_mis_cri/621027%20Dobrynin%20Cable%20to%20USSR.pdf, 1

[18] Online Supplement A describes randomization procedures and demographics.

To be sure, the military personnel in my games might have different preferences than the civilians who often oversee decisions on the use of force (Betts 1991). Elected officials might, for instance, weigh factors like public opinion more than military officers. Yet military decisionmakers provide advice on contingency plans that informs civilian policymakers' decisions. They also make choices that can constrain or drive policy. For instance, a military planner's decision to substitute an inhabited aircraft with a drone might shape crisis outcomes.

Teams had 30 minutes to work through the scenario, helping to simulate time-constraints of real-world crises. Rapporteurs assigned to each team documented deliberations and logged each team's final plan.[19] To be sure, notetaking involves selection and interpretation, with observers recording certain observations and omitting others (Emerson, Fretz, and Shaw 2011, 13). To account for this, rapporteurs also led post-game group interviews that asked all teams the same questions. Recording direct interview responses requires less interpretative discretion than collecting deliberative data.

To test hypotheses, I analyze each team's response plan and the deliberations that led to decisions. I identify similarities within the treatment and control games and look for differences across them. Specifically, I assess whether the type of platform shot down leads to variation in the concepts discussed during the decision-making process and in recommended response measures. Do players discuss escalation concerns, emotions, or desires to prevent future harm? Table 1 outlines expected speech evidence. I treat speech data in much the same way that scholars use archival materials in case studies: I review game transcripts for evidence that either supports hypotheses or disproves them. Moreover, as with most case study research, space constraints limit

[19] Online Supplement A describes data collection procedures.

me to presenting only information directly related to hypothesis testing. The full game transcript, however, is available in the replication files.

To explore the microfoundations that underpin decisions on escalation, I assess *why* participants make the recommendations they do. What factors do they consider when developing response plans? Discussion about the need to degrade a rival's ability to attack additional aircraft provides evidence that instrumental factors are at work, while an expressed desire for revenge is suggestive of emotional factors. To be sure, multiple factors can inform policy choices, making it challenging to determine which specific factors are most influential (Jervis 2006, 645–46). Although disentangling instrumental and emotional foundations and identifying precise "causal weights" is difficult, the deliberative data can still reveal which factors are salient as leaders make decisions. The absence of discussion associated with a theorized logic suggests it is less important than a factor that is repeatedly discussed. Remote-controlled restraint yields several expectations about the factors decisionmakers should consider when responding to attacks (Table 1)

Table 1. Expected Evidence from Wargame Discussions

	Expected Content of Discussions	
Underlying Foundations	***When drone is shot down:***	***When inhabited aircraft is shot down:***
Instrumental Foundations (H_1)		
Desire to degrade/deter rival capabilities (Tit-for-tat response logic)	- Less desire to degrade/deter - Attack viewed as insignificant - Limited discussion of cost	- Significant desire to degrade/deter - Attack viewed as significant - Significant discussion of cost (focused on blood cost)
Concern that retaliation will trigger excessive escalation	- Significant concern	- Less concern since escalatory retaliation is seen as justified
Concern over reputation for weak response	- Limited discussion	- Significant discussion
Emotional Foundations (H_2)		
Desire for punishment (due to anger)	- Less/No desire to punish rival - Limited expression of anger - More risk averse with retaliation	- Significant desire to punish rival - More significant expression of anger - Greater risk acceptance with retaliation

When assessing outcomes, I view escalation as a continuum. I consider an event that involves kinetic strikes as more escalatory than a show of force, which in turn, is more escalatory than a diplomatic rebuke.[20] While other approaches, like survey experiments, often use ordinal multiple-choice questions to offer quantitative measures, these unrealistically constrain outcomes in a way that my wargames do not. The remote-controlled restraint hypothesis yields several expectations about the less escalatory courses of action decisionmakers should pursue when a drone is shot down (Table 2).

Table 2. Expected Wargame Outcomes

Expected Course of Action	
When drone is shot down:	***When inhabited aircraft is shot down:***
- Lower likelihood of retaliation/recovery operations - Non-military response preferred - Limited/No escalation (e.g., non-kinetic response)	- Higher likelihood of retaliation/recovery operations - Military response preferred - More significant escalation (e.g., kinetic strikes)

FINDINGS

In my wargames, Katunian forces used a surface-to-air missile (SAM) to down an unarmed U.S. Air Force reconnaissance plane flying in Dakastani airspace. The aircraft crashes, with wreckage falling into Katunian territory. All teams were exposed to the identical scenario, but the four treatment games were told the downed aircraft was an unarmed MQ-1 Predator drone, while the three control games experienced the loss of an inhabited MC-12 Liberty intelligence aircraft whose four-member crew was killed.

If remote-controlled restraint explains behavior, military decisionmakers should recommend a less escalatory response to an attack on a drone than an attack on an inhabited aircraft. Participants should view a drone loss as less costly in blood and treasure. Subsequently, they will worry that significant retaliation will trigger unwanted escalation and have fewer concerns about the reputational consequences of a weak response. Their discussions should also

[20] This aligns with the Militarized Interstate Disputes dataset that classifies actions by intensity.

be less laden with emotional language about anger than teams exposed to the shootdown of an inhabited aircraft. Table 3 summarizes expectations and findings. Shading represents theoretical expectations, and a check indicates observed actions or discussion. Each column represents a wargame team.

Table 3. Summary of Findings

Recommended Actions	**Expected & Observed Course of Action**						
	When drone is shot down:				***When inhabited aircraft is shot down:***		
	Team A	Team B	Team C	Team D	Team E	Team F	Team G
Military strikes in retaliation					√	√	√
Launch recovery operations (Demonstrates risk-acceptant behavior)			√*		√	√	
Non-kinetic response (diplomacy, posturing, etc)	√	√	√	√			
Underlying Foundations	**Expected & Observed Content of Discussions**						
	When drone is shot down:				***When inhabited aircraft is shot down:***		
	Team A	Team B	Team C	Team D	Team E	Team F	Team G
Instrumental Foundations							
Desire to degrade rival capabilities					√	√	√
Concern that retaliation will trigger unwanted escalation	√	√	√	√	√		√
Concern over reputation for weak response		√	√	√		√	√
Emotional Foundations							
Desire for punishment (due to anger)					√	√	

Note: * Recovery only launched "if easy."

Outcomes: Evidence for Remote-Controlled Restraint

The wargames provide support for remote-controlled restraint. As Table 3 shows, none of the four teams that lost drones recommended military strikes against Katunia. In contrast, all three teams that lost an inhabited aircraft proposed retaliatory strikes. Teams that lost drones typically coupled diplomatic efforts with shifts in military operations in attempts to prevent future attacks. One team moved future reconnaissance flights further from the border and recommended that diplomats publicly condemn the shootdown. Another team proposed condemning Katunia at the United Nations and threatened economic sanctions if Katunia failed to withdraw anti-aircraft systems from the contested border region. This team also planned to escort future drone missions with fighter jets and threatened kinetic retaliation for future attacks. A third team proposed issuing a diplomatic demarche, while simultaneously increasing the alert posture of U.S. forces in the region and flying another reconnaissance mission to demonstrate resolve. A fourth launched a show of force mission near the contested region. In addition to restraining their retaliatory measures, teams that lost drones generally avoided escalation in their handling of wreckage. Two teams (50%) opted to leave wreckage in Katunia, one team (25%) planned to ask Katunia to return it, and one team (25%) considered operations to destroy the wreckage.[21] These findings align with the expectation of limited, non-kinetic reactions to drone losses.

In contrast to the tempered response to a drone shootdown, all three teams that lost an inhabited aircraft recommended kinetic military strikes against Katunian targets. Two of these teams recommended targeting Katunian military facilities to degrade the country's air defense capabilities, while the third team planned to attack Katunian forces impeding a recovery operation. These actions would violate Katunian sovereignty, risked additional loss of life, and involved greater risk of escalation. Teams that lost inhabited aircraft were also more adamant about the need to recover the downed crewmembers' remains—even though doing so involved risk and was escalatory. Two teams deployed military forces into Katunia to retrieve remains and destroy sensitive wreckage. The team that did not deploy a military recovery operation recommended sending a diplomatic request for the remains.

Emotional Drivers

[21] See Online Supplement A.4 for team recommendations.

As expected, the lack of killed or captured personnel underpinned the muted reactions on teams that lost a drone. Participants often joked about the lack of humans onboard. "Where do you bury the survivors?" quipped one participant.[22] He noted, "The good thing is that there is no pilot being dragged along."[23] Another participant suggested Katunia was "helping us out" in expediting the retirement of older drones.[24] "Send them a bill for the MQ-1," another jokingly suggested.[25] The lighthearted attitude suggests attacks on drones do not elicit the type of emotional reactions that can cue more escalatory responses. Indeed, there was no explicit discussion of a need to punish Katunia and little evidence of risk-acceptant behavior, suggesting the incident triggered little anger.

When exposed to the downing of an inhabited aircraft, however, participants' language took on a more emotive tone.[26] An Army colonel signaled less restraint and more risk acceptance in the measures he was willing to take: "The gloves are off."[27] An Air Force officer voiced a similar sentiment, proclaiming, "This represents a conflict. If they shot down our service members, we retaliate."[28] Emotional factors also informed deliberations on whether to retrieve the downed airmen's remains. Participants invoked the hallowed military principle of never abandoning fallen personnel. "Get[ting] the bodies is an automatic choice. [It's] just what you do," remarked one Air Force officer.[29] The participants understood such retaliatory and retrieval actions incurred risk, but felt obliged to act. In sum, the lack of blood costs with drone losses triggered minimal emotional reactions, allowing a more tempered response than the loss of an inhabited asset.

---

[22] Army Officer (A-4). Participants cited using a descriptor and ID (Team-Participant #).
[23] Army Officer (A-4).
[24] Department of Defense Civilian (D-4).
[25] Army Enlisted (A-1).
[26] There was no explicit use of terms like "reprisal" or "vengeance," perhaps because Law of Armed Conflict training instills norms against the wanton use of force.
[27] Army Officer (E-4).
[28] Air Force Officer (F-4).
[29] Air Force Officer (F-4).

Instrumental Drivers

Although the wargames provide some evidence that emotions influence decisions on escalation, team deliberations centered on forward-looking instrumental factors informed by the shootdown's perceived costs, including concerns about reputation, escalation, and preventing future attacks. Teams facing drone shootdowns generally believed losses incurred relatively low human and material costs. One participant recalled, "Predators have been lost before. It's a mostly disposable asset," suggesting that the downing of a drone does not meet the saliency criterion that justifies a significant response.[30]

Although drone losses were perceived as low cost, participants still sought to prevent future attacks. Indeed, participants were more concerned about both the reputational and operational consequences of inaction following a drone shootdown than expected. Teams worried that entirely ignoring a drone loss could open the door to future aggression, perhaps against inhabited assets. Participants feared that failing to take action could put U.S. resolve "in question."[31] As one participant noted, the "adversary will take advantage if we don't do anything."[32] On another team, a participant explained, "Yes, they destroyed only material, but we need to deter future activity."[33] An Army officer agreed, "We must let them know future activity will be responded to."[34]

Most teams, however, believed kinetic responses to a drone loss were too escalatory and could prompt an unwanted Katunian reaction. A fighter pilot expressed concerns that retaliatory airstrikes incurred a "high risk of casualties" among the Katunian population, potentially triggering undesired escalation.[35] A team member agreed, noting that Katunia would "definitely respond to

[30] Army Enlisted (A-1).
[31] Army Enlisted (C-3).
[32] Army Enlisted (C-3).
[33] Army Enlisted (D-3).
[34] Army Officer (D-2).
[35] Air Force Officer (C-2).

any targeted strike."[36] On another team, a participant explicitly asked whether a drone shootdown justified escalatory retaliation: "What's the line between the value of low-stakes and high stakes assets?…Do we want to escalate to deter?"[37] He was quickly met with pushback. An Army officer on his team responded, "We want to avoid escalation – we want to control the escalation ladder.…We don't want to start a war."[38] Participants on another team held similar preferences, ultimately coming to the consensus that "We are not going to war over a Predator."[39]

This desire to deter future shootdowns while minimizing escalation risk led teams to propose diplomatic responses coupled with military threats or posturing. On one team, participants laid out potential responses ranging from halting reconnaissance missions to a large-scale invasion of Katunia. The team believed striking the offending SAM site and doing nothing were realistic, but determined they "[couldn't] just pack up and leave"[40] as "do[ing] nothing gives them a win."[41] One team member suggested, "We need to demonstrate air superiority."[42] His teammates agreed, noting that "Air superiority combines the best of both worlds. Don't destroy the SAM, but signal that we can."[43] Consistent with remote-controlled restraint, the team settled on this non-kinetic approach as "a good way to do something but refuse to escalate."[44] Another team recommended issuing a diplomatic demarche and bolstering the alert status of U.S. forces in the region, suggesting participants viewed non-kinetic actions as a sufficient, but non-escalatory means, of responding to a drone shootdown. In short, teams viewed an attack on a drone as falling below the threshold that triggered a significant response.

[36] Army Officer (C-1).
[37] Army Enlisted (D-3).
[38] Army Officer (D-2).
[39] Army Officer (A-4).
[40] Army Officer (B-3).
[41] Navy Petty Officer (B-4).
[42] Army Non-Commissioned Officer (B-1).
[43] Army Officer (B-3).
[44] Army Non-Commissioned Officer (B-1).

The perception of drones as a low-cost asset also dominated discussions on what to do with wreckage. Although three of the teams believed Katunia could gain sensitive technology, the Predator was typically not viewed as sufficiently sensitive to justify the escalation that could result from deploying a recovery team or destroying the wreckage. Teams recognized that operations in Katunian territory violated international law and risked casualties.[45] Although one team recommended a military operation to recover or destroy the drone "if easy" to accomplish,[46] the other three teams believed recovery was not worth the risk of escalation. One officer suggested that, "If we lost an extremely sensitive aircraft, we could send in [special operations forces] or carry out a strike. This is high risk in terms of escalation, but we could take this action….[But] by and large the technical loss is minimal."[47] His teammates agreed, noting that there was "no return on investment" in taking action to recover or destroy the drone.[48] Another commented that dealing with the wreckage was "best handled politically. If we did carry out an airstrike, we could kill their troops. A strike would escalate the situation."[49]

In contrast, teams responding to an inhabited aircraft shootdown perceived a significant loss that warranted greater risks and escalation to prevent future attacks. Discussions about losses focused on the American crew, without mentioning financial costs. Teams generally weighed the risks of various actions, before settling on a tit-for-tat response intended to degrade adversary capabilities. On one team, an Air Force aviator proposed a "do nothing option or a measured response that is either kinetic or non-kinetic."[50] He ruled out doing nothing because of the implications of inaction. One team member pushed back, concerned that any response might

[45] Army Officer (C-1).and Air Force Officer (C-2).
[46] Army Enlisted (C-3). The team provided no details on what constitutes an "easy" mission.
[47] Army Officer (A-4).
[48] Army Enlisted (A-1).
[49] Army Officer (A-4).
[50] Air Force Officer (G-2).

"tak[e] the U.S. into potential war with Katunia." This member was outnumbered by teammates, suggesting decisionmakers are willing to risk escalation to preserve reputation and prevent future attacks on inhabited assets.[51] An Air Force cyber operations officer on the team recommended a tit-for-tat action, suggesting, "maybe cyber with a kinetic response that doesn't need to be disproportionate."[52] An Army officer agreed, explaining the need to demonstrate resolve and "message that you attack our people, we at least respond in kind, even if this risks escalating conflict."[53]

This team then turned to the timing and type of kinetic response, shedding additional light on the factors that inform responses to attacks on inhabited aircraft. The cyber operations officer argued, "It's not the time to drop weapons. Posture forces first and if [Katunian] behavior does not change, then use force."[54] The aviator pushed back, "I totally disagree. Tit-for-tat strike on the SA-8…If we don't respond, it suggests they can shoot down Americans who are not violating [their] airspace."[55] An Army officer agreed with the need to deter Katunian hostilities, but urged a stronger response that involved targeting an "entire company of SAM sites or destroy[ing] command and control." He believed a simple tit-for-tat response was "not enough to shift behavior or deter" and that a response should "mak[e] it more costly to shoot down Americans."[56] The cyber operations officer continued to resist the immediate use of force, explaining, "We can deter further aggression without taking kinetic action now."[57] His teammates overruled these preferences. At the end of the gaming period, the team recommended a moderate response that fell between the posturing of forces and the destruction of Katunia's air defense network: striking a single SA-8

[51] Army Enlisted (G-5).
[52] Air Force Officer (G-4).
[53] Army Officer (G-1).
[54] Air Force Officer (G-4).
[55] Air Force Officer (G-2).
[56] Army Officer (G-1).
[57] Air Force Officer (G-4).

SAM site using a U.S. asset operating in friendly airspace over Dakastan to avoid putting additional U.S. forces at risk and "flaring things up" further.[58]

The team also debated how to handle the crew remains and aircraft wreckage. They considered "going in with a recovery team to get remains and sensitive equipment and intelligence[,]" or bombing the wreckage, but believed operations in Katunian territory would be an "act of escalation."[59] The team assumed Katunian authorities had already removed the remains which would complicate recovery operations, and worried striking the wreckage could lead to civilian Katunian casualties. Because of the complexity of recovery operations and the risk that civilian casualties could adversely affect world opinion toward the United States, the team recommended sending a diplomatic request for the remains.

Other teams exposed to an inhabited shootdown followed a similar decision-making logic. Participants believed responses should be forceful enough to significantly degrade Katunia's ability to conduct future attacks and demonstrate American resolve. Non-kinetic means like sanctions, embargos, or cyber operations were considered, but were generally perceived as an insufficient tit-for tat for the death of American servicemembers. One participant, for instance, lamented, "an embargo makes us look weak, we must degrade [Katunia's] capabilities to wage war."[60] These teams also had concerns about the risks of destroying wreckage and recovering bodies, but decided to recover the downed crew's remains—even though doing so was escalatory and violated international law. Participants on one team were driven by both military norms and concerns about America's reputation. They agreed that recovering remains was "everything we

[58] Air Force Officer (G-2).
[59] Air Force Officer (G-2).
[60] Marine Corps Enlisted (F-3).

stand for."[61] Another team member suggested "we look weak if we don't get bodies. It says something if we leave them."[62]

While responses to inhabited losses were more escalatory and risk-acceptant than responses to drone shootdowns, participants still sought to avoid disproportionate escalation. This aligns with recent scholarship that suggests that even in high-stakes confrontations, states seek to avoid destabilizing escalation (Carson 2018). To minimize the likelihood their actions would trigger a significant counter-reaction from Katunia, teams carefully selected targets they believed would reduce the risk of Katunian casualties. One team, for example, recommended striking Katunian military hangars as a means of degrading capabilities, but minimize risk of civilian or military casualties.[63] These strikes were designed to "roll back options" that Katunia could use in future hostilities. Another team opted to strike any Katunian forces and military infrastructure that would hamper efforts to recover crewmember remains.

In short, participants demonstrated a greater willingness to escalate and accept risk to prevent future attacks after the loss of an inhabited aircraft than after the loss of a drone. Yet, drone losses were not seen as "costless" as conventional wisdom might suggest. Accordingly, decisionmakers took non-kinetic steps to deter future hostility and to demonstrate U.S. resolve.

Validating Findings: Survey Experiment Insights

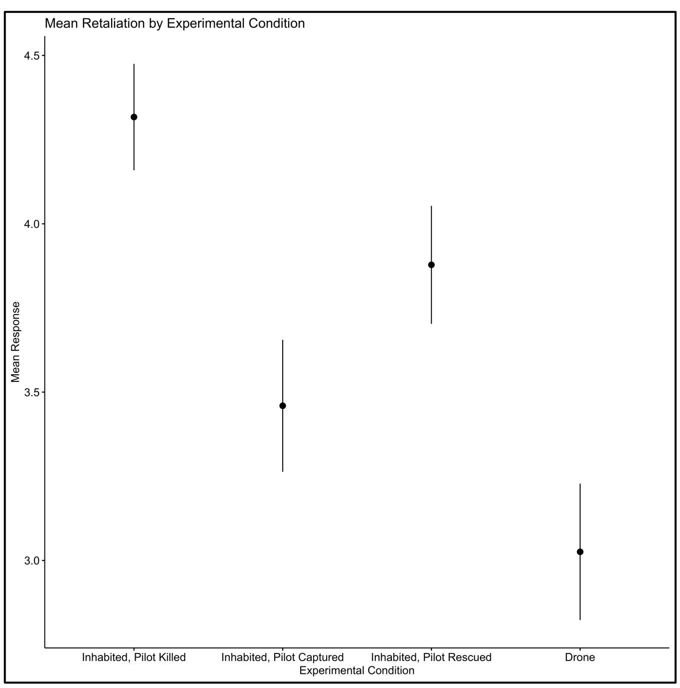
As an additional test that helps compensate for the small wargame sample size, I field a survey experiment with a similar vignette on an expert sample of 158 U.S. military officers.[64] As in the wargame, a rival downs a U.S. military aircraft. The survey instrument varies whether the

[61] Marine Corps Enlisted (F-3).
[62] Army Enlisted (F-2).
[63] Air Force Officer (E-3).
[64] See Online Supplement B.

aircraft was inhabited or remotely piloted. It also varies whether the pilot of an inhabited aircraft was killed, rescued, or captured. Respondents are then asked to select the most appropriate response. Unlike the wargames in which players formulated response options, survey respondents chose from a list of five actions coded on an ordinal 5-point scale: 1) no action, 2) formal diplomatic protest, 3) economic sanctions, 4) show of military force, or 5) limited airstrike on the missile site that downed the aircraft.

Consistent with the remote-controlled restraint logic and the wargame findings, respondents, on average, call for less escalatory retaliation following an attack on a drone than on an inhabited aircraft (Figure 1). Even when the pilot is not killed (i.e., captured or rescued), respondents still support more escalatory responses when an inhabited aircraft is attacked than when a drone is. This suggests military officers respond not only to the death of servicemembers, but treat drones and inhabited platforms as qualitatively different.

Figure 1.



The survey experiment's larger sample allows for precise causal inference and additional manipulations that validate wargame findings and probe their generalizability. While survey experiments allow researchers to efficiently study outcomes, they are typically unable to provide the type of insights on decision-making processes that are revealed through participant interaction during immersive wargames. This suggests scholars may gain analytic leverage by employing multi-method approaches that layer the qualitative richness of wargames with quantitative analysis of survey experiments featuring larger samples.

CONCLUSION

This project advances research on the security implications of emerging military technologies by using a novel comparative wargaming approach to assess whether and why drones help limit crisis escalation. The deliberations of actual national security practitioners immersed in realistic crisis scenarios reveals that escalatory retaliation can be avoided after drone losses primarily because decisionmakers see little need to degrade a rival's ability to conduct future attacks and because the loss of a machine does not trigger strong emotional reactions that can elicit more risk-acceptant behavior. To be sure, if leaders believe that attacking a drone is unlikely to trigger retaliation, they might be more prone to down a rival's drones – something I demonstrate using another wargame.[65] The muted response to drone losses, however, still helps control escalation.

The findings extend drone scholarship beyond studies on conflict initiation and provide empirical support for existing conceptual work examining drones and escalation. The findings also contribute to studies that link drones with crisis signaling (Zegart 2020) by explaining why leaders

[65] See Online Supplement C.

continue deploying drones in the face of hostile actions. More broadly, the finding that drones can help control escalation suggests there are limits to prominent realist theories that posit that the world becomes less stable when technologies lower the costs of warfighting. Indeed, drones appear to limit destabilizing escalation precisely because their use incurs low costs and risks.

By shedding light on the microfoundations that underpin decisions on retaliation, the project also contributes to the resurgence of scholarly attention on escalation (Acton 2018; Carson 2018). The wargames demonstrate that instrumental desires to degrade a rival's warfighting capability inform escalation decisions and that emotions can amplify decisionmakers' risk acceptance. Further, the findings suggest measurements for escalation in many academic studies require greater nuance, with distinctions between actions involving drones and those involving inhabited systems.

Finally, comparative wargaming advances IR methodology by blending analytically desirable features of case study and experimental research. Immersing security practitioners in scenarios that manipulate a variable of interest and control for confounders offers significant promise to researchers studying topics where real-world data are scant.

The analysis suggests several avenues for future research. First, fielding additional wargames with non-U.S. participants or using samples that include civilian officials might test the generalizability of remote-controlled restraint across different contexts and identify variation in the decision-making logics of different types of actors. Second, future studies could explore factors which might influence decisions on escalation such as relative military capability vis-à-vis an adversary. Third, scholars might field wargames that generate data to inform ongoing political science debates.[66] For instance, games featuring expert and non-expert players could help study

[66] Lin-Greenberg, Pauly, and Schneider (2022) identifies potential research areas.

whether elite and mass subjects behave differently during simulated crises or to study how group composition affects decision-making. While many important questions remain, the comparative wargames presented here deepen our understanding of the strategic implications of drones—a topic that will become increasingly important as these systems proliferate globally.

References

Acton, James. 2018. "Escalation through Entanglement: How the Vulnerability of Command-and-Control Systems Raises the Risks of an Inadvertent Nuclear War." *International Security* 43 (1): 56–99.

Axelrod, Robert. 2006. *The Evolution of Cooperation*. New York: Basic Books.

Barnes, Julian. 2019. "U.S. Cyberattack Hurt Iran's Ability to Target Oil Tankers, Officials Say." *The New York Times*, August 28, 2019. https://www.nytimes.com/2019/08/28/us/politics/us-iran-cyber-attack.html.

Bartels, Elizabeth. 2018. "Games as Structured Comparisons: A Discussion of Methods." In . San Francisco, CA.

Bartels, Elizabeth, Igor Mikolic-Torreira, Steven Popper, and Joel Predd. 2019. *Do Differing Analyses Change the Decision?: Using a Game to Assess Whether Differing Analytic Approaches Improve Decisionmaking*. Santa Monica, CA: RAND Corporation.

Bertrand, Natasha. 2015. "Russia Is Already Exacting Its Revenge on Turkey for Downing a Russian Warplane." *Business Insider*, November 26, 2015. https://www.businessinsider.com/russia-turkey-downed-jet-2015-11.

Betts, Richard. 1991. *Soldiers, Statesmen, and Cold War Crises*. Second Edition. New York: Columbia University Press.

Biddle, Stephen. 2004. *Military Power: Explaining Victory and Defeat in Modern Battle*. Princeton, NJ: Princeton University Press.

Boyle, Michael. 2020. *The Drone Age: How Drone Technology Will Change War and Peace*. New York, NY: Oxford University Press.

Brecher, Michael. 1993. *Crises in World Politics: Theory and Reality*. Oxford: Pergamon.

Carson, Austin. 2018. *Secret Wars: Covert Conflict in International Politics*. Princeton, NJ: Princeton University Press.

Cenciotti, David. 2015. "Here's the Route of a US RQ-4 Global Hawk Drone on a Surveillance Mission over Ukraine and Eastern Europe." Business Insider. July 20, 2015. https://www.businessinsider.com/route-of-us-rq-4-global-hawk-drone-surveillance-mission-over-ukraine-2017-7.

Clark, Colin. 2019. "US 'Loyal Wingman' Takes Flight: AFRL & Kratos XQ-58A Valkyrie." Breaking Defense. March 7, 2019. https://breakingdefense.com/2019/03/us-loyal-wingman-takes-flight-afrl-kratos-xq-58a-valkyrie/.

Cohen, Eliot. 1996. "A Revolution in Warfare." *Foreign Affairs* 75 (2): 37–54.

Coskun, Orhan. 2015. "Turkey Shoots down Drone near Syria, U.S. Suspects Russian Origin." *Reuters*, October 16, 2015. https://www.reuters.com/article/us-mideast-crisis-turkey-warplane/turkey-says-its-warplanes-shot-down-unidentified-aircraft-near-syria-idUSKCN0SA15K20151016.

Crawford, Neta. 2000. "The Passion of World Politics: Propositions on Emotion and Emotional Relationships." *International Security* 24 (4): 116–56.

Creveld, Martin Van. 1991. *Technology and War: From 2000 B.C. to the Present*. New York: Touchstone.

Daniel, J Furman, and Paul Musgrave. 2017. "Synthetic Experiences: How Popular Culture Matters for Images of International Relations." *International Studies Quarterly* 61 (3): 503–16.

Emerson, Robert, Rachel Fretz, and Linda Shaw. 2011. *Writing Ethnographic Fieldnotes, Second Edition*. Chicago: University of Chicago Press.

Fearon, James. 1994. "Domestic Political Audiences and the Escalation of International Disputes." *The American Political Science Review* 88 (3): 577–92.

———. 1995. "Rationalist Explanations for War." *International Organization* 49 (3): 379–414.

Frijda, Nico. 1986. *The Emotions*. New York: Cambridge University Press.

Gelpi, Christopher, Peter Feaver, and Jason Reifler. 2009. *Paying the Human Costs of War: American Public Opinion and Casualties in Military Conflicts*. Princeton, NJ: Princeton University Press.

Gelpi, Christopher, and Joseph Grieco. 2015. "Competency Costs in Foreign Affairs: Presidential Performance in International Conflicts and Domestic Legislative Success, 1953–2001." *American Journal of Political Science* 59 (2): 440–56.

George, Alexander, and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: The MIT Press.

Gilli, Andrea, and Mauro Gilli. 2016. "The Diffusion of Drone Warfare? Industrial, Organizational, and Infrastructural Constraints." *Security Studies* 25 (1): 50–84.

Goldblum, Bethany, Andrew Reddie, and Jason Reinhardt. 2019. "Wargames as Experiments: The Project on Nuclear Gaming's SIGNAL Framework." Bulletin of the Atomic Scientists. May 29, 2019. https://thebulletin.org/2019/05/wargames-as-experiments-the-project-on-nuclear-gamings-signal-framework/.

Gray, Colin. 1993. *Weapons Don't Make War: Policy, Strategy, and Military Technology*. Lawrence, KS: University Press of Kansas.

Gusterson, Hugh. 2016. *Drone: Remote Control Warfare*. Cambridge, MA: The MIT Press.

Hermann, Charles. 1967. "Validation Problems in Games and Simulations with Special Reference to Models of International Politics." *Behavioral Science* 12 (3): 216–31.

Holland, Paul. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945–60.

Horowitz, Michael. 2020. "Do Emerging Military Technologies Matter for International Politics?" *Annual Review of Political Science* 23 (1): 385–400.

Horowitz, Michael, Sarah Kreps, and Matthew Fuhrmann. 2016. "Separating Fact from Fiction in the Debate over Drone Proliferation." *International Security* 41 (2): 7–42.

Horowitz, Michael, Joshua Schwartz, and Matthew Fuhrmann. 2020. "Who's Prone to Drone? A Global Time-Series Analysis of Armed Uninhabited Aerial Vehicle Proliferation." *Conflict Management and Peace Science*.

Huth, Paul. 1997. "Reputations and Deterrence: A Theoretical and Empirical Assessment." *Security Studies* 7 (1): 72–99.

Jensen, Benjamin, and Brandon Valeriano. 2019. "Cyber Escalation Dynamics: Results from War Game Experiments." In . Toronto, Canada.

Jervis, Robert. 1976. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.

———. 1995. "Political Implications of Loss Aversion." In *Avoiding Losses/Taking Risks: Prospect Theory and International Conflict*, edited by Barbara Farnham, 23–40. Ann Arbor: University of Michigan Press.

———. 2006. "Understanding Beliefs." *Political Psychology* 27 (5): 641–63.

Kaag, John, and Sarah Kreps. 2014. *Drone Warfare*. Cambridge: Polity.

Kahn, Herman. 1965. *On Escalation: Metaphors and Scenarios*. New York: Fredrick A. Praeger.

Kertzer, Joshua D. 2017. "Microfoundations in International Relations." *Conflict Management and Peace Science* 34 (1): 81–97.

———. 2021. "Re-Assessing Elite-Public Gaps in Political Behavior." *American Journal of Political Science*.

Kreps, Sarah. 2014. "Flying under the Radar: A Study of Public Attitudes towards Unmanned Aerial Vehicles." *Research & Politics* 1 (1): 4–7.

Kreps, Sarah, and Jacquelyn Schneider. 2019. "Escalation Firebreaks in the Cyber, Conventional, and Nuclear Domains: Moving beyond Effects-Based Logics." *Journal of Cybersecurity* 5 (1).

Lerner, Jennifer, Roxana Gonzalez, Deborah Small, and Baruch Fischhoff. 2003. "Effects of Fear and Anger on Perceived Risks of Terrorism: A National Field Experiment." *Psychological Science* 14 (2): 144–50.

Lerner, Jennifer, Ye Li, Piercarlo Valdesolo, and Karim Kassam. 2015. "Emotion and Decision Making." *Annual Review of Psychology* 66 (1): 799–823.

Liberman, Peter. 2006. "An Eye for an Eye: Public Support for War Against Evildoers." *International Organization* 60 (3): 687–722.

Lieber, Keir. 2008. *War and the Engineers: The Primacy of Politics over Technology*. Ithaca, NY: Cornell University Press.

Lieber, Keir, and Daryl Press. 2017. "The New Era of Counterforce: Technological Change and the Future of Nuclear Deterrence." *International Security* 41 (4): 9–49.

Lowen, Mark. 2015. "Missing Russian Pilot 'Alive and Well.'" *BBC News*, November 25, 2015. https://www.bbc.co.uk/news/world-middle-east-34917485.

Löwenheim, Oded, and Gadi Heimann. 2008. "Revenge in International Politics." *Security Studies* 17 (4): 685–724.

Lyall, Jason. 2020. "Drones Are Destabilizing Global Politics." Foreign Affairs. December 16, 2020. https://www.foreignaffairs.com/articles/middle-east/2020-12-16/drones-are-destabilizing-global-politics.

Macdonald, Julia, and Jacquelyn Schneider. 2017. "Presidential Risk Orientation and Force Employment Decisions: The Case of Unmanned Weaponry." *Journal of Conflict Resolution* 61 (3): 511–36.

———. 2019. "Battlefield Responses to New Technologies: Views from the Ground on Unmanned Aircraft." *Security Studies*, 1–34.

Mahnken, Thomas, Travis Sharp, and Grace Kim. 2020. *Deterrence by Detection: A Key Role for Unmanned Aircraft Systems in Great Power Competition*. Washington, DC: Center for Strategic and Budgetary Assessments.

McDermott, Rose. 2002. "Experimental Methodology in Political Science." *Political Analysis* 10 (4): 325–42.

McDermott, Rose, Anthony Lopez, and Peter Hatemi. 2017. "'Blunt Not the Heart, Enrage It': The Psychology of Revenge and Deterrence." *Texas National Security Review* 1 (1): 68–88.

McHugh, Francis. 1966. *Fundamentals of War Gaming*. 3rd ed. Newport, RI: US Naval War College.

Mercer, Jonathan. 2010. "Emotional Beliefs." *International Organization* 64 (1): 1–31.

Mesquita, Bruce Bueno de. 1983. *The War Trap*. New Haven: Yale University Press.

Miller, Andrew. 2020. “The Information Game: Police-Citizen Cooperation in Communities with Criminal Groups.” Thesis, Massachusetts Institute of Technology. https://dspace.mit.edu/handle/1721.1/128634.

Mintz, Alex, Steven Redd, and Arnold Vedlitz. 2006. “Can We Generalize from Student Experiments to the Real World in Political Science, Military Affairs, and International Relations?” *The Journal of Conflict Resolution* 50 (5): 757–76.

Mir, Asfandyar. 2018. “What Explains Counterterrorism Effectiveness? Evidence from the U.S. Drone War in Pakistan.” *International Security* 43 (2): 45–83.

Mir, Asfandyar, and Dylan Moore. 2019. “Drones, Surveillance, and Violence: Theory and Evidence from a US Drone Program.” *International Studies Quarterly*.

Morgan, Forrest, Karl Mueller, Evan Medeiros, Kevin Pollpeter, and Roger Cliff. 2008. *Dangerous Thresholds: Managing Escalation in the 21st Century*. Santa Monica, CA: RAND Corporation.

Mueller, John. 1973. *War, Presidents, and Public Opinion*. New York: John Wiley & Sons.

National Security Agency. 1989. *The National Security Agency and the EC-121 Shootdown (Top Secret)*. Fort Meade, Maryland: Office of Archives and History, National Security Agency. https://www.nsa.gov/Portals/70/documents/news-features/declassified-documents/cryptologic-histories/EC-121.pdf.

Oberholtzer, Jenny, Abby Doll, David Frelinger, Karl Mueller, and Stacie Pettyjohn. 2019. “Applying Wargames to Real-World Policies.” *Science* 363 (6434): 1406–1406.

Panda, Ankit. 2015. “Meet China’s East China Sea Drones.” *The Diplomat*, June 30, 2015. https://thediplomat.com/2015/06/meet-chinas-east-china-sea-drones/.

Pauly, Reid. 2018. “Would U.S. Leaders Push the Button? Wargames and the Sources of Nuclear Restraint.” *International Security* 43 (2): 151–92.

Perla, Peter. 1990. *The Art of Wargaming: A Guide for Professionals and Hobbyists*. Annapolis, MD: US Naval Institute Press.

Petersen, Roger. 2002. *Understanding Ethnic Violence: Fear, Hatred, and Resentment in Twentieth-Century Eastern Europe*. New York: Cambridge University Press.

Posen, Barry. 1991. *Inadvertent Escalation: Conventional War and Nuclear Risks*. Ithaca, NY: Cornell University Press.

Reddie, Andrew, Bethany Goldblum, Kiran Lakkaraju, Jason Reinhardt, Michael Nacht, and Laura Epifanovskaya. 2018. “Next-Generation Wargames.” *Science* 362 (6421): 1362–64.

Reiter, Dan, and Allan Stam. 2002. *Democracies at War*. Princeton, NJ: Princeton University Press.

Rempfer, Kyle. 2018. “Air Force Reapers Are Now Flying ISR Missions from Poland.” *Air Force Times*, June 1, 2018. https://www.airforcetimes.com/news/2018/06/01/air-force-reapers-are-now-flying-isr-missions-from-poland/.

Sabin, Philip. 2014. *Simulating War: Studying Conflict through Simulation Games*. London: Bloomsbury.

Schechter, Benjamin, Jacquelyn Schneider, and Rachael Shaffer. 2021. “Wargaming as a Methodology: The International Crisis Wargame and Experimental Wargaming.” *Simulation & Gaming*, January.

Schelling, Thomas. 1966. *Arms and Influence*. New Haven, CT: Yale University Press.

Schneider, Jacquelyn. 2017. "Cyber Attacks on Critical Infrastructure: Insights from War Gaming." War on the Rocks. July 26, 2017. https://warontherocks.com/2017/07/cyber-attacks-on-critical-infrastructure-insights-from-war-gaming/.

Schulman, Loren DeJonge. 2018. *Weird Birds: Working Paper on Policymaker Perspectives on Unmanned Aerial Vehicles and Their Impact on National Security Decision-Making*. Washington, D.C.: Center for New American Security.

Smoke, Richard. 1977. *War: Controlling Escalation*. Cambridge, MA: Harvard University Press.

Snyder, Glenn. 1959. "Deterrence by Denial and Punishment." Research Monograph No. 1. Princeton, N.J.: Woodrow Wilson School of Public and International Affairs.

Snyder, Glenn, and Paul Diesing. 1977. *Conflict Among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton, NJ: Princeton University Press.

Talmadge, Caitlin. 2017. "Would China Go Nuclear? Assessing the Risk of Chinese Nuclear Escalation in a Conventional War with the United States." *International Security* 41 (4): 50–92.

———. 2019. "Emerging Technology and Intra-War Escalation Risks: Evidence from the Cold War, Implications for Today." *Journal of Strategic Studies* 42 (6): 864–87.

Tannenwald, Nina. 1999. "The Nuclear Taboo: The United States and the Normative Basis of Nuclear Non-Use." *International Organization* 53 (3): 433–68.

Tomz, Michael, Jessica Weeks, and Keren Yarhi-Milo. 2020. "Public Opinion and Decisions About Military Force in Democracies." *International Organization* 74 (1): 119–43.

Trump, Donald. 2019. "Remarks by President Trump and Prime Minister Trudeau of Canada Before Bilateral Meeting." The White House. June 20, 2019. https://www.whitehouse.gov/briefings-statements/remarks-president-trump-prime-minister-trudeau-canada-bilateral-meeting-2/.

UK MoD. 2017. *Wargaming Handbook*. Swindon, UK: Development, Concepts, and Doctrine Centre.

U.S. Air Force. 2010. "FY2011 Budget Estimates, Aircraft Procurement, Air Force, Volume 1." https://web.archive.org/web/20120304052331/http://www.saffm.hq.af.mil/shared/media/document/AFD-100128-072.pdf.

———. 2017. "MC-12 Factsheet." https://www.af.mil/About-Us/Fact-Sheets/Display/Article/104497/mc-12/.

Valeriano, Brandon, Benjamin Jensen, and Ryan Maness. 2018. *Cyber Strategy: The Evolving Character of Power and Coercion*. New York: Oxford University Press.

Van Evera, Stephen. 2001. *Causes of War*. Ithaca, NY: Cornell University Press.

Wagner, William. 1982. *Lightning Bugs and Other Reconnaissance Drones*. Fallbrook, CA: Aero Publishers.

Walsh, James Igoe, and Marcus Schulzke. 2018. *Drones and Support for the Use of Force*. Ann Arbor: University of Michigan Press.

Weisiger, Alex. 2016. "Learning from the Battlefield: Information, Domestic Politics, and Interstate War Duration." *International Organization* 70 (2): 347–75.

Weisiger, Alex, and Keren Yarhi-Milo. 2015. "Revisiting Reputation: How Past Actions Matter in International Politics." *International Organization* 69 (2): 473–95.

Zegart, Amy. 2020. "Cheap Fights, Credible Threats." *Journal of Strategic Studies* 43 (1): 6–46.