Cyber Escalation Dynamics: Results from War Game Experiments

International Studies Association, Annual Meeting
Panel: War Gaming and Simulations in International Conflict
March 27, 2019

Benjamin Jensen, Ph.D.
Marine Corps University
American University, School of International Service
&
Brandon Valeriano, Ph.D.
Bren Chair of Armed Politics
Marine Corps University

Introduction

In February 2019, India and Pakistan exchanged artillery and small arms fire while engaging in air-to-air combat and bombing runs after a terrorist attack India blamed on Pakistan. Despite militarized escalation in this key rivalry, there were no observed large-scale cyber actions above the level of website defacements and temporarily taking websites offline (Times of India 2019). Power plants didn't shut down. Dams kept working.

In November 2015, a Turkish Air Force F-16 shot down a Russian Su-24M attack aircraft. Rather than respond with military strikes, Russian reciprocated with a series of overt and covert actions. According to a CrowdStrike report,

> Russian Federal Security Service (FSB) had raided and shut down numerous Turkish bank branches in the Russian Federation, detained Turkish travelers, stopped Turkish vehicles from crossing into Russia, and denied Turkish trade ships from entering Russian ports….Around this time, CrowdStrike observed Distributed Denial of Service (DDoS) attacks targeting Turkish state-owned banks, government sites, and hacking forums. Soon hacktivists operating under Anonymous-style monikers began targeting the Turkish root DNS and threatening to destroy the banking infrastructure claiming that they buy oil from ISIS, amongst other rhetoric (Meyers 2016).

These events present a puzzle: when and why do states use cyber operations to respond during a militarized dispute? Do cyber operations tend to be associated with more or less escalatory responses in crises? How does any new disruptive technology alter the cycle of competitive risk-taking at the core of strategic exchange between rival states? Answering these questions is a core task for IR scholars and policy makers in the digital age.

War games and simulations offer a means to exploring this puzzle. Any crisis is defined by interdependent decision-making in environments characterized by uncertainty, time pressure, and private information. New technologies and tools of coercion compound this uncertainty and create new risk vectors. Therefore, designing exercises that replicate the context and character of national security decision-making environments, draw on players with similar backgrounds to career professionals, and utilize randomization and treatment groups that control for the impact of disruptive technologies offer a viable research path.

To that end, we designed a strategic-level war game that captures decisions about whether or not to escalate in a crisis. The war games controlled for if there was 1) a cyber triggering event to the crisis and 2) whether or not participants could respond with cyber operations alongside more traditional instruments of power (i.e., DIME). After observing 259 crises responses via the war game, two novel insights emerge. First, the presence of cyber operations do not escalate militarized disputes. In experimental settings, participants did not use cyber operations to escalate a hypothetical crisis with a rival state. Second, at the same time when a crisis is triggered by a cyber intrusion and the targeted state lacks effective cyber response options, the crisis tended to escalate. That is, when players lacked the ability to conduct proportional cyber responses they opted to escalate conventionally as a demonstration of resolve.

The paper below unpacks these findings. First, we define escalation and explore major findings in the literature to design a war game that reflects key facets of the literature. Specifically, the war game was primed for escalation, involving rival states engaged in a dispute over territory with uncertain but equal power levels. Second, we detail the war game design and procedures used to administer the treatments. Third, we analyze the findings in relation to two hypotheses grounded in the escalation literature. Fourth, we conclude by discussing additional research needed to unpack cyber escalation dynamics.

Escalation

The modern study of crisis escalation emerged during the Cold War in the examination of strategic competition as a bargaining process (Schelling 1960; 1966).[1] Bargaining involves threats, both implicit and explicit, of violence that seek to achieve the maximum benefit relative to the perceived costs and risks inherent in escalation (Snyder and Diesing 1977, 450). For Herman Kahn (1965), a crisis involves competitive risk-taking. In a crisis two sides bargain via threats, creating multiple pathways from low level conflicts and all-out war (Kahn 1965, 83). As such, crises were framed in terms of rational decision-making given limited information.[2]

For Brecher (1996), escalation is a function of three intervening attributes: the value threatened, time pressure, and the overall likelihood of war. Brecher treats each of these factors as a subjective perception held by decision makers. These perceptions in turn are influenced by a range of larger structural factors, as independent variables, including relative capabilities, regime type, the issue at sake and internal stability of each state in the crisis.

[1] For an overview of recent literature on bargaining, see Ramsay (2018).

[2] This foundational assumption is best captured in the literature on rationalist deterrence, which can be contrasted with psychological perspectives. For an overview of both, see Kurizaki (2016).

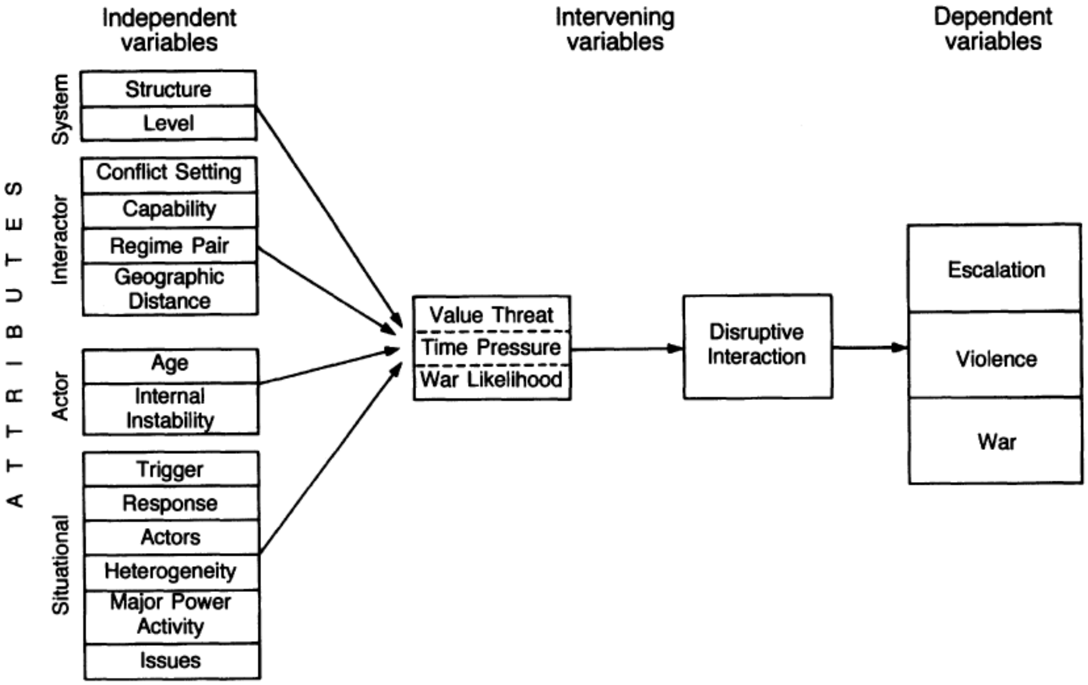


FIGURE 1. *Crisis Escalation Model.*

In early quantitative treatments, researchers measured escalation by the extent to which a party reciprocated a dispute using the militarized interstate dispute (MID) dataset (Braithwaite and Lemke 2011; Ghosn et al. 2004). Early studies using the MID dataset saw a reciprocation rate, where the target responses at the same or a higher level of action, 48% of the time. These studies emphasized key firebreaks between threats and the use of force (Colaresi and Thompson 2002) as well as the role of strategic assets like nuclear weapons (Geller 1990). Most reciprocation studies parallel foundational game theory work (Axelrod 1984), to illustrate the prevailing logic of “tit-for-tat” (Ward 1982, Patchen 1987, Leng and Wheeler 1979). The question then is what causes actors to assume risk and move beyond reciprocation to escalation.

For Brecher (1996), protracted conflicts and rivalry, disputes over territory, sub state crises, and power balances all increase escalation risks. Protracted conflicts and rivalries are more prone to arms races that enhance the value placed on gaining an advantage in a crisis leading to an increased likelihood of escalation (Brecher 1996; 220, Sample 1997). Consistent with the broader literature on the causes of interstate war and rivalry (Vasquez 1993, Thompson 2001), parties in a dispute are more likely to escalate when the crisis involves territorial disputes (Leng and Singer 1988). In particular, whether or not the states involved share a contiguous border and have an active territorial dispute increases the likelihood one side will use force as well as produces more fatalities. (Senese 1996; Toft 2014). Furthermore, these disputes are more likely to escalate into war (Hensel 1996; Vasquez and Henehan 2001).[3] This linkage between territory, rivalry, and escalation parallels major findings in the steps-to-war theory (Senese and Vasquez 2003, 2008; Colaresi and Thompson 2005).

Regime type can influence escalation patterns and pathways. Earlier scholarship in the democratic peace literature found that a dispute between democracies is less likely to escalate or reach higher hostility levels (Maoz and Abdolali 1989; Maoz and Russett 1993). Furthermore, democracies are more likely to reciprocate in a dispute when the initiator is non-democratic (Prins and Sprecher 1999).

There is no consensus regarding how relative power affects crisis dynamics and escalation. Brecher argues that low or no discrepancy in relative power between parties in a crisis “reduces the credibility of a threat to use force” (220). For others relative capabilities, as a measure of power, exhibit a non-monotic relationship (Bueno de Mesquita et al 1997). There is a tipping point after which power preponderance no longer limits escalation. This perspective counters classical and structural realist approaches that hold power balances as decreasing the likelihood of conflict. Alternatively, the literature on power preponderance and power transitions takes the opposite view, holding that power balances actually increase the risk of

[3] Hensel (1996) Vasquez and Henehan (2001)

conflict (Kugler and Lemke, 1996). Other authors hold that power has either no effect or is moderated by intervening factors such as the offense-defense balance (Jervis 1978). For Smith (1999), strategic choices are subject to interdependent decision making and censoring effects that diminish the influence of power imbalances. Similarly, Morrow (1989) finds that relative power should not affect crisis bargaining. The correlation of forces and the beliefs each side holds are not related (Morrow 1989, 958). As such, even when a side has more power, they may lose in a crisis. Maoz (1983) tests claims about relative power finding that minor powers who initiate crises with major powers tend to prevail. Given these dynamics, one would expect that a crisis scenario involving states with a territorial dispute, an enduring rivalry, and similar levels of power will be prone to crisis escalation.

Key capabilities can influence crisis dynamics. In IR, much ink has been shed hunting an elusive offense-defense balance. In empirical tests of extended deterrence, Huth and Russett (1984, 1988) found that neither the military balance nor possession of nuclear weapons produces extended deterrent successes. Rather, deployable conventional forces, as power projection, proved a best predictor of crisis stability. This finding parallels earlier work by Blechman and Kaplan (1978) arguing that conventional military power, not nuclear weapons, determines crisis outcomes.

With respect to cyber capabilities, coercion is possible, but limited and subject to unique effects characteristic of the domain (Borghard and Lonergan 2017).[4] Buchanan (2017) sees a heightened security dilemma owing to their covert character. States have incentives "incentive[s] to break in early either to conduct offense, to set up contingency plans and

[4] For a contrasting perspective on cyber capabilities as a strategic revolution challenging the operative international and even moral order, see Kello (2017). For new approaches to studying cross-domain deterrence, see Gartzke and Lindsay (2019).

deterrents, and to aid their defenses" (Buchanan 2017). This security dilemma is further complicated by the widespread use of proxies in cyberspace (Maurer 2018). The offense-defense balance rests not just on factors like geography and alliances, but human capital and organizational capacity to develop the highly tailored malware required to penetrate classified networks (Slayton 2017).

In fact, certain categories of weapons are so catastrophic that great powers develop a set of strategic norms that place constraints on escalation (Osgood and Tucker 1967, 137). [5] This tacit bargaining limits escalatory upside at the expense of accepting a series of lower-level actions. In Cold War literature, these sub crisis maneuvers (Kahn 1965) and local crises expanded the steps of the escalation ladder as a means of avoiding larger risk taking. For Geller, nuclear states exhibit a greater tendency towards escalation short of war (1990). In the cyber literature, Lindsay and Gartzke (2018) argue that there is a new stability-instability paradox with respect to cyber capabilities. Cyber capabilities are an instrument of political warfare optimized for sub crisis maneuvering (Jensen 2017; Valeriano, Jensen, Maness 2018). There effects are limited (Gartzke 2013) and do not fall neatly into conventional understandings of offense or defense (Gartzke and Lindsay 2015). This line of reasoning produces the first hypothesis:

H1. The presence of cyber response options will not produce crisis escalation[6]

If cyber operations are a form of sub-crisis maneuvering, a combination of ambiguous signals used to manage competition (Carson 2018) or a means of conducting operational preparation of the environment. As an instrument of power and influence cyber capabilities do not cross, at least not yet, the threshold of escalation to armed conflict. In a crisis, when decision

[5] For one of the earliest, comprehensive looks at cyber capabilities, see Rattray (2001). On the integration of cyber capabilities into larger campaigns, see Conti and Raymond (2017).

[6] The null hypothesis is that the presence of cyber response options is associated with episodes of escalation

makers have access to cyber response options they are more likely to use them to conduct reciprocal responses beneath the threshold of armed conflict while taking defensive actions that prepare them for future interactions with an adversary.

Yet, as a new form of strategic competition the use of cyber capabilities can alter how states calculate the risk of escalation. As outlined by Buchanan (2017), cyber capabilities compound uncertainty. Offense and defense blur together. Even intelligence activities can be misconstrued as offensive action.[7] Private information increases. As covert action, cyber intrusions in a network may be defensive counter attacks, espionage, or an attempt to create a window of opportunity for a larger strategic strike. Therefore, the situations in which a state is involved in a dispute with a known cyber intrusion but lacks the ability to retaliate with cyber options, there should be a higher rate of escalation. The demonstration of capabilities makes it difficult for states to respond short of escalating with more conventional options to signal resolve, a concept consistent with earlier studies on capabilities and escalation (Bueno de Mesquita et al 1996). This line of reasoning leads to our second hypothesis:

H2. Crises with cyber triggering events but without cyber response options will be more escalatory[8]

From a broader perspective, both arguments explore how disruptive capabilities alter bargaining processes and crisis dynamics. Beyond reducing capabilities to pre-determined offense or defensive attributes (Jervis 1978) or ascribing independent preferences to actors consistent with the recent turn to Bayesian realism (Kydd 2005), how do new technologies alter information and uncertainty during a crisis? The effects observed in crises involving the use or

[7] On the role of intelligence in bargaining, see Arena and Wolford (2012).

[8] The null hypothesis is that situations with cyber triggering events and no cyber response options are not associated with episodes of escalation.

potential use of cyber capabilities may have similar dynamics to satellites, electronic warfare, and other systems that blur the lines between intelligence, offense, and defense.

Wargaming Cyber Crises

Wargames offer a viable method for evaluating propositions on the character of escalation under the context of cyber operations and disruptive technology more generally. The use of wargames to evaluate interdependent decision-making has a long tradition in the military profession and strategic studies communities (Perla 1990; Van Crevald 2013). Wargames, as a form of simulation, are a useful method for evaluating competing hypotheses, focusing data investigations, and delineating patterns otherwise unobserved (Druckman 1994; Wilkenfeld et al 2003).

Recently, there has been a renaissance not just in war gaming but in the use of war games to evaluate interdependent decision-making in a strategic setting (Barzashka 2018). Pauly (2019) used a sample of war games with strategic elites to examine attitudes towards nuclear weapons, finding restraint based on reputational risks.

There is a special interest in using games to uncover the dynamics of interdependent decision making in cyber exchanges (Krebs and Das 2017; Krebs and Schneider 2018; Gomez and Villar 2018; Gross et al 2017). Jacqueline Schneider used a longitudinal analysis of war games between 2011 and 2016 to study crisis dynamics. Her work revealed that government officials were reluctant to use high-end cyber offensive capabilities (Schneider 2018). Of note, Schneider found that participants only used offensive cyber capabilities after conventional military strikes and expressed concerns that using offensive cyber would increase the risk of nuclear escalation. Jensen and Banks (2018) found similar patterns in a series of war games analyzing how decision maker integrated cyber operations into crises with both great power competitors and non-state actors. Escalation was the exception, not the rule.

Our effort builds on this continuum. We designed the war game to ensure we had the right context, players, and replicated the crisis atmosphere as much as possible outside of actual national security decision-making. Our war game involved two hundred fifty-nine participants including graduate and undergraduate students, government officials, military officers, and private sector employees playing a war game involving two hypothetical nuclear states. This population mix avoids the external validity questions raised by only using college students (Mintz et al 1997; Green and Gerber 2002).[9] The scenario pitted green state vs. purple state, two rival states with power parity. Using hypothetical states helped remove participants from pre-existing biases about current international relations.

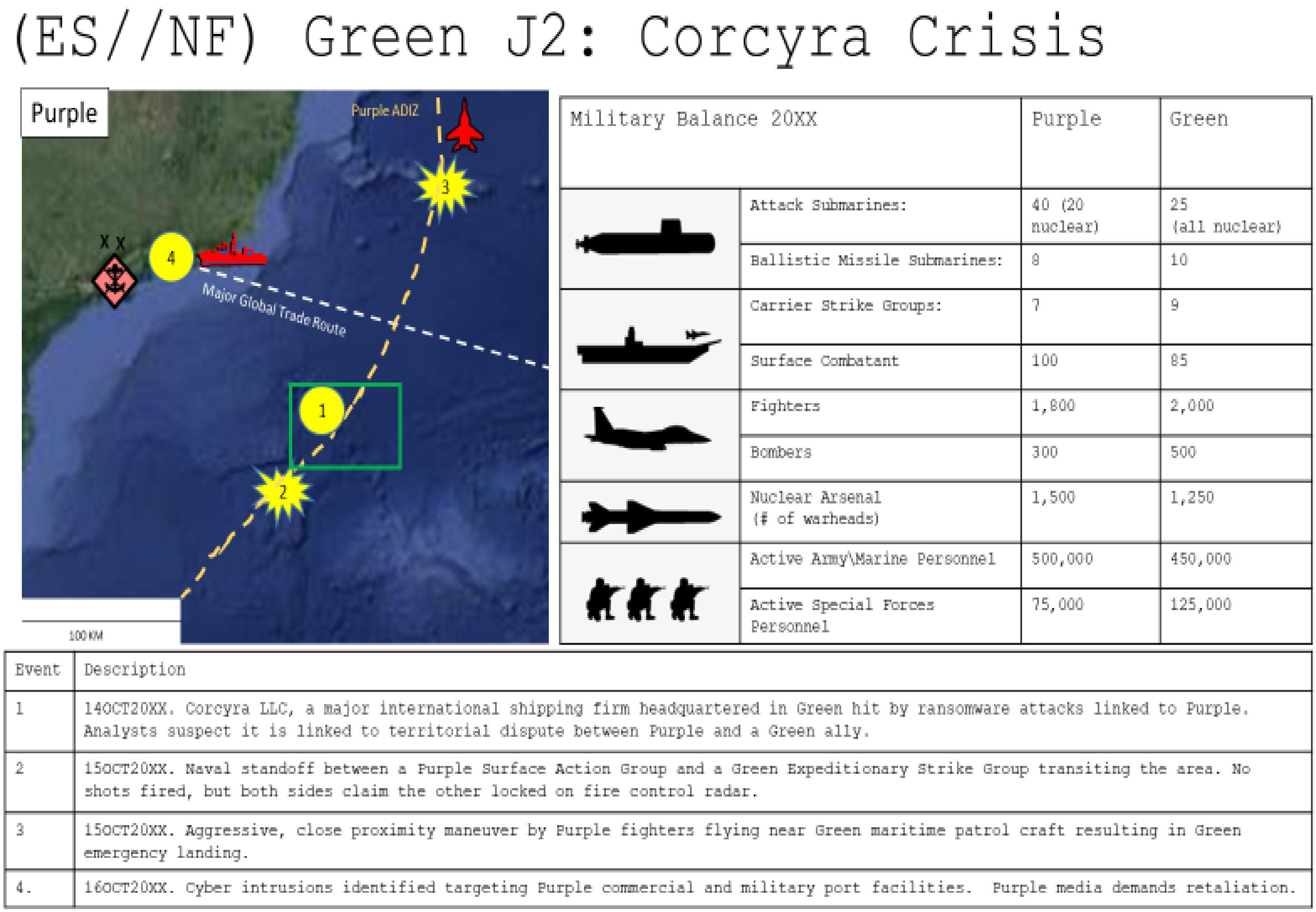
The scenario used findings from the literature on crisis escalation to create an especially volatile situation. Green and purple states are involved in a dispute over territory, have roughly equal power levels, and are rivals. The packets did not address regime type to remove the hypothesized effect of democracy.

Participants were given packets modeled after documents used in the U.S. National Security Council (NSC).[10] The briefing packet provided an overview of the crisis and asked participants to select response options organized in relation to different instruments of power (i.e., diplomatic, information, military and economic). Participants were randomly assigned one of four different packets. Each packet had the following: 1) a tasking memo providing background on the current crisis 2) a J2 update (see figure below) 3) a response packet with pre-approved flexible response options and 4) a post-game survey. The figure below illustrates a

[9] On the experimental method and political science in general, see McDermott (2002).

[10] The game designers reached out to former NSC staffers from two separate administrations.

sample J2 update from one of the treatments.



(ES//NF) Green J2: Corcyra Crisis

Purple
Purple ADIZ
X X
4
3
Major Global Trade Route
1
2
100 KM

Military Balance 20XX		Purple	Green
	Attack Submarines:	40 (20 nuclear)	25 (all nuclear)
	Ballistic Missile Submarines:	8	10
	Carrier Strike Groups:	7	9
	Surface Combatant	100	85
	Fighters	1,800	2,000
	Bombers	300	500
	Nuclear Arsenal (# of warheads)	1,500	1,250
	Active Army\Marine Personnel	500,000	450,000
	Active Special Forces Personnel	75,000	125,000

Event	Description
1	14OCT20XX. Corcyra LLC, a major international shipping firm headquartered in Green hit by ransomware attacks linked to Purple. Analysts suspect it is linked to territorial dispute between Purple and a Green ally.
2	15OCT20XX. Naval standoff between a Purple Surface Action Group and a Green Expeditionary Strike Group transiting the area. No shots fired, but both sides claim the other locked on fire control radar.
3	15OCT20XX. Aggressive, close proximity maneuver by Purple fighters flying near Green maritime patrol craft resulting in Green emergency landing.
4.	16OCT20XX. Cyber intrusions identified targeting Purple commercial and military port facilities. Purple media demands retaliation.

As reflected in Table 1 below, these treatments differentiated between whether or not the players had cyber response options available in each instrument of power and whether or not there was a cyber triggering event involved in the recent crisis. The players played green state while a separate team played purple state. All statistics on crisis response and escalation rates are reported for green.

Table 1
Treatment Groups

Treatment			Number
1.	Cyber Response Options (Yes)	Cyber Triggering Event (Yes)	86
2.	Cyber Response Options (No)	Cyber Triggering Event (Yes)	36
3.	Cyber Response Options (Yes)	Cyber Triggering Event (No)	79
4.	Cyber Response Options (Yes)	Cyber Triggering Event (Yes)	58

N = 259

To assess escalation, the response packet built in three unique measures. First, respondents were asked whether they wanted to escalate, conduct a proportional response or de-escalate. Second, the response options were arrayed in a manner that differentiated between three de-escalatory options and three escalatory options for each instrument of power. The researchers coded an escalatory event as long as any single response option was escalatory. That is, even if the respondent thought they were being proportional and selected three responses – a request for backchannel diplomacy (de-escalatory), public call for calm (de-escalatory) and a threat of economic sanctions (escalatory) – the response was coded as escalatory. The responses were listed moving from least to most escalatory but avoided offering the participants the option of using nuclear weapons. Response options also differentiated between traditional instruments and cyber options. This allowed the researchers to vary the treatments and offer participants cyber response options and no cyber response options. Last, each participant was asked to rate the overall escalation risk at the end of the war game.

The responses reported in the Analysis section below report the second measure of escalation as binary categorical variable (escalation, no escalation) based on players selected response options. Tables 2 and 3 below provide samples of diplomatic and military response options for treatments 1 and 3, where participants had cyber response options. When participants lacked cyber response options only the left column appeared (e.g., D1-D6, M1-M6).

Table 2

Foreign Ministry Flexible Response and Deterrent Options
(Diplomacy)

D1. Publicly call for third-party crisis mediation.	Di1. Publicly call for limiting cyber activities to safeguard a connected world.
D2. Use lobbyists (intermediaries) and tailored messages to target select elites about the risk of escalation.	Di2. Release some known adversary vulnerabilities to cyber security firms as way of signaling risks.
D3. Use a third party to signal a desire to limit escalation.	Di3. Use a third party to signal a desire to limit offensive cyber activities by both sides.
D4. Issue a diplomatic demarche privately threatening escalation if your adversary does not back down.	Di4. Conduct website defacements and circulate false stories on social media with key propaganda themes.
D5. Initiate noncombatant evacuations and withdraw non-essential personal.	Di5. Use spear phishing, waterholing, and other methods to expose sensitive political information.
D6. Expel diplomats and shutdown diplomatic offices while calling for allied support, and requesting UN/regional international organizations to act.	Di6. Shutdown adversary official communication networks through large-scale degradation operations including key satellite interfaces and internet networks.

Table 3
Defense Ministry approved flexible response and deterrent options
(Military)

M1. Pull back forward deployed naval assets and offensive air packages while maintaining ISR collection (Phase I-0 transition).	Mi1. Increase cyber defenses and mobilize additional network defenders as well as run diagnostics on key systems.
M2. Increase active and passive force protection measures.	Mi2. Hold a press conference indicating the risk of future offensive cyber action and a desire to avoid future exchanges.
M3. Increase ISR to provide early warning of adversary military activity.	Mi3. Increase ISR of adversary military networks.
M4. Conduct a public show of force with air and naval assets challenging known defense zones and testing adversary response.	Mi4. Compromise data of individual members of the military to include identify theft, fraud, or direct social media messaging
M5. Conduct a limited strike, targeting a small number of military facilities (1-3) with cruise missiles and threaten overwhelming force if the adversary responds.	Mi5. Use exploit chains to erode rival military navigation, targeting, and C2 CIA triad. This can include periodic disruption, spoofing, or taking networks temporarily offline.
M6. Exercise preventive theater strike options and attack multiple ISR and C2 networks, key airfields and naval assets. Threaten strategic escalation if the enemy responds.	Mi6. Exploit intrusions in adversary strategic weapons systems and national C2 infrastructure to take these weapons offline and blind the enemy.

In each war game, participants were randomly separated into groups and given a crisis brief. The groups reflected the different treatments listed in Table 1 differentiating between cyber response options and cyber triggering events to the recent crisis. Facilitators then let individuals work through the scenario packet. This step usually took thirty to forty-five minutes. The treatment groups then convened to discuss their responses. Facilitators wandered between groups taking notes on how groups discussed various options. Each group then briefed their recommended response options to the rest of the assembled group. At this point, the facilitator would reveal that groups had different response options and transition to a general conversation about strategy, cyber response options, and contemporary crises. The design is thus best characterized as a seminar style, table-top exercise (Pournelle 2017).

To interpret the results, we produce a contingency table and compare the observed frequencies relative to the expected frequencies if there was no association (i.e., the null hypothesis). [11] Specifically, we test whether the decision to escalate is independent of the scenario treatments based on whether the actor has cyber response options and there was a cyber trigger to the crisis. To this end, we first conduct a Chi-Square test of independence. Second, we compare the adjusted residuals to assess the relationship between observed and expected values.[12] Third, we used Cramer's v to assess association between the scenario treatments and the decision to escalate.[13]

[11] On categorical data analysis, see (Agresti 2002; Delucchi 1993).

[12] If the standardized residual is +/- 1.96 the relationship is significant.

[13] A value > .3 indicates a moderately strong relationship. On the use of Cramer's V to test associations between categorical variables, see (Marchant-Shapiro 2015; Howell 2002; Sheskin 2011).

Analysis

Table 4 below reports the number of response options that escalated in relation to each of the four treatments with percentages. When war game participants had cyber response options and the scenario involved a cyber trigger, they escalated only 36% of the time. Of note, this number is below the rate of reciprocated escalation in prior studies of militarized disputes (48%). The largest rate of escalation by percentage was in treatment 2, where respondents were confronted with a cyber triggering event but lacked cyber response options. When a state faced a cyber intrusion alongside more traditional forms of escalation and lacked the ability to respond with cyber options, the participants opted to escalate more frequently (63.9%) than other treatments. The lowest observed escalation level was in treatment 3, when the crisis did not involve a cyber trigger and participants had cyber response options. That is, participants used cyber to respond beneath the threshold of escalation. The distribution of escalation responses across the four treatment groups is statistically significant in terms of the observed differences. Figure 2 illustrates these differences in a bar chart.

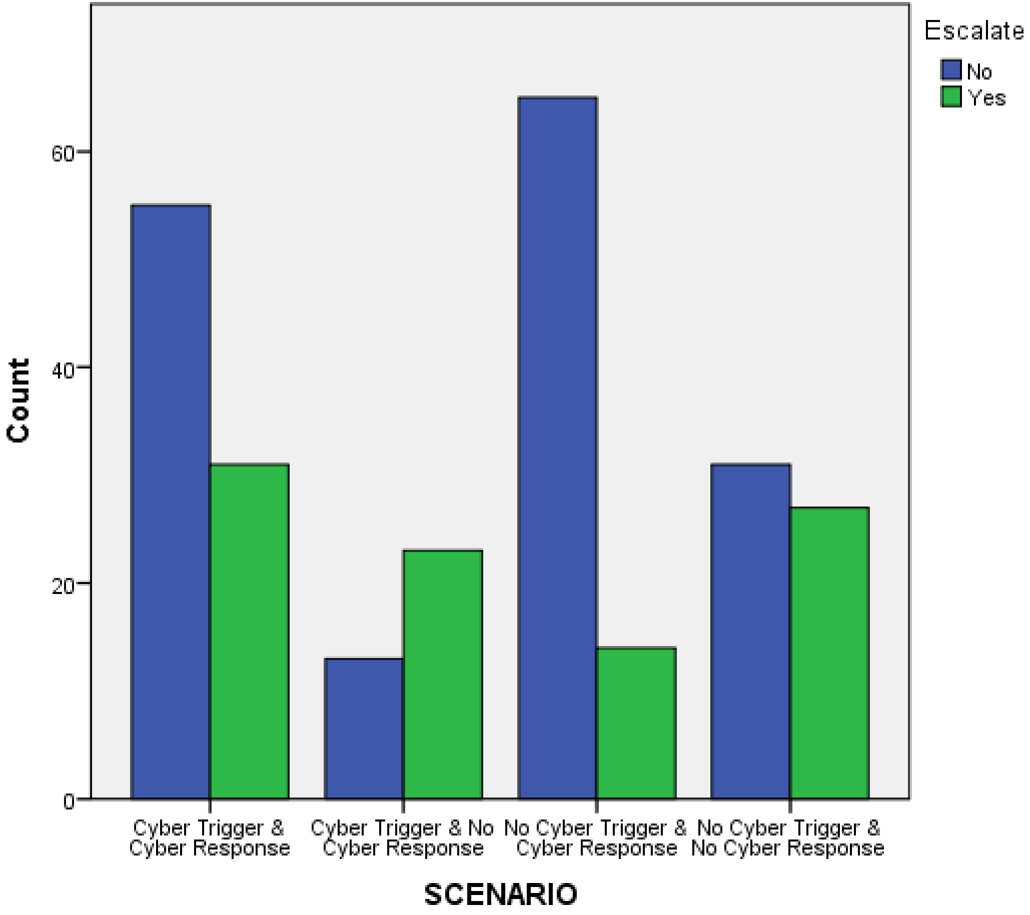
Table 4
Crisis Outcomes

	Cyber Trigger Cyber Response	Cyber Trigger No Cyber Response	No Cyber Trigger Cyber Response	No Cyber Trigger No Cyber Response
Escalation	31 (36%)	23 (63.9%)	14 (17.7%)	27 (46.6%)
No Escalation	55 (64%)	13 (36.1%)	65 (82.3%)	31 (53.4%)
Totals	86	36	79	58

$X^2 = 25.15$
$p < .01$ (two-sided)
n = 259

In Figure 2 below, crisis response are presented as a bar graph. The difference between decisions to escalate relative to no escalation are especially apparent in treatments two and three discussed above. When participants confronted a cyber trigger and lacked cyber response options, they were more escalatory. When they confronted a crisis without cyber attacks and had cyber response options, they were less escalatory.

Figure 2: Escalation Dynamics



Turning to the first hypothesis, it appears the presence of cyber response options does not produce crisis escalation. As seen in table 5 below, there are statistically significant differences across the treatments. Much of this difference comes from the divergence in responses seen between treatment 2 and treatment 3 based on looking at the adjusted residuals. When the participants had cyber response options (treatment 3) they escalated less than would be expected. When participants lacked cyber response options (treatment 2) they escalated more than would be expected. We can reject the null that cyber response options are associated with higher observed frequencies of escalation. The relationship observed is also moderately strong based on Cramer's V. The presence and absence of cyber response options appears to shape how participants respond to a crisis.

Table 5
Crisis Outcomes: Proportional Effects

	Cyber Trigger Cyber Response	Cyber Trigger No Cyber Response	No Cyber Trigger Cyber Response	No Cyber Trigger No Cyber Response
Escalation	31 (-.1)	23 (2.7)	14 (-2.8)	27 (1.2)
No Escalation	55 (.1)	13 (-2.1)	65 (2.1)	31 (-.9)
Totals	86	36	79	58

$X^2 = 25.15$
p < .01 (two-sided)
Standardized Residuals in parentheses
Cramer's V .318
n = 259

These response dynamics were consistent across gender, occupational, citizenship and technological backgrounds. As seen in Table 6 below, the only category that skewed results was age, but this finding is limited. The number of players over 50 years old was too small to make a robust assessment.

Table 6
Crisis Outcomes: Control Groups

	Gender	Occupation	Age	Citizenship	Technology
X^2	4.5	5.31	6.225*	3.15	2.47

Percentage reported
* p < .1, ** p < .05, *** p < .01
n = 259

As expected, cyber response options appear to play a moderating role. Looking at how the presence of cyber response options and cyber triggers affect the odds of escalation and relative risk in Table 7 below, we find odds of escalation with cyber response options are .333.[14] Participants used cyber options as a means of managing crisis and conducting operational preparation of the environment for future moves more than they did immediate escalatory pathways. Cyber responses appear to provide the ability for rival states to respond with a degree of deniability to avoid escalation.

Table 7
Crisis Outcomes: Cyber Response Options and Escalation Risks

	No Escalation	Escalation
No Cyber Response Options	44 (46.8%)	50 (53.2%)
Cyber Response Options	120 (72.7%)	45 (27.3%)

Observed number reported (% within Response)
$X^2 = 17.321$
$p < .01$
$n = 259$

Looking at escalation events, there were 95 episodes out of 259 observations. Of these, only 166 games allowed for selecting between cyber and non-cyber response options (treatments 1 and 3). Specifically, treatment 1, which involved a cyber trigger and offered participants cyber response options, had 27 non-cyber response options and 4 cyber response options (12.9% of escalation) that involved escalation. Treatment 3, which did not have a cyber trigger but offered participations cyber response option, had 7 non-cyber and 7 cyber response options involving escalation (50% of escalation). Combined, participants escalate using cyber only 6.6% of the time. In other words, when actors chose to escalate outside of treatment three, they opted for more conventional forms of statecraft.

Turning to the second hypothesis, crises with cyber triggering events where there were no cyber response options proved to be associated with higher levels of escalation. As seen in table

[14] On calculating odds ratios and relative risk, see Fleiss (1981).

5 above, treatment 2 (cyber trigger, no cyber response) has higher than expected escalation and lower than expected non-escalation based on the standardized residuals. This finding lends credence to the idea of new capabilities complicating response dynamics in a crisis (Jervis 1978; Bueno de Mesquita et al 1997). When the triggering events involved cyber intrusions, the demonstration of cyber capabilities made it difficult for players to respond short of escalating with more conventional options to signal resolve.

Cyber triggers increase the odds of escalation to 1.813. As seen in Table 8 below when there is a cyber trigger escalation occurred 43.9% of the time. When there was no cyber trigger, states opted not to escalate 69.9% of the time.

Table 8
Crisis Outcomes: Cyber Triggers and Escalation Risks

	No Escalation	Escalation
No Cyber Trigger	95 (69.9%)	41 (30.1%)
Cyber Trigger	69 (56.1%)	54 (43.9%)

Observed number reported (% within Trigger)
$X^2 = 5.262$
$p < .05$
n = 259

Reading across the findings an interesting dynamic emerges. When participants had cyber response options, they were significantly less likely to escalate. When there were cyber triggers but no cyber response options it increased the risk of escalation. Lacking the ability to send ambiguous signals, participants opted for a more conventional approach to escalation. There was nothing inherently escalatory about cyber response options, but cyber triggers appear to increase uncertainty and hence higher observed rates of escalation.

Conclusion

This paper provides an overview of a war game designed as an experiment to test how the presence of cyber capabilities alters cyber escalation dynamics. The results indicate that cyber response options can moderate crisis response in scenarios involving rival states with power parity. The results also indicate that new capabilities can create uncertainty. When participants were confronted with situations involving a cyber triggering event but lacked cyber response options, they opted to signal resolve and pursue more conventional responses.

Future research should unpack these results along three axes. First, what would it take to cross the Rubicon? That is, it appears cyber capabilities are subject to a firebreak that limits escalation. Yet, what conditions cause participants to engage in more risk-taking and pursue significant escalation events using cyber capabilities? While to date cyber options appear to reflect a form of sub crisis maneuvering (Jensen 2017; Valeriano, Jensen, Maness 2018), the question remains what sequence of events would cause participants to break the operative norms against escalatory strategic retaliation.

Second, what combinations of cyber capabilities and conventional power projection alongside traditional instruments of statecraft prove the most stable signaling mechanisms in a crisis? This question requires iterating crisis interactions over multiple turns and testing different combinations. This question is especially important to disentangle from the effects of nuclear weapons. Most cyber powers are nuclear powers (Valeriano, Jensen, Maness 2018). As cyber capabilities proliferate and non-nuclear states use them will crisis stability hold?

Third, how do new actors complicate traditional state rivalry in cyberspace? This questions revolves around a new research agenda in cyber civil-military-commercial relations

(Work and Jensen 2018) and exploring how prominent role of the private sector in cyberspace alters bargaining processes at the core of international crises.

References

Agresti, A. (2002). *Categorical Data Analysis (2nd Ed.)*. New York: Wiley

Arena, Philip ad Scott Wolford. 2012. "Arms, Intelligence and War" *International Studies Quarterly* 56: 351-365

Axelrod, Robert. 1984. *The Evolution of Cooperation* (New York: Basic Books)

Barzaskha, Ivanka. "Wargaming: how to turn vogue into science" *Bulletin of the Atomic Scientists* March 15

Borghard, Erica and Shawn Lonergan. 2017. "The Logic of Coercion in Cyberspace" *Security Studies* 26

Blechman, B. M., and S. S. Kaplan. 1978. *Force without war: U.S. armed forces as a political instrument*. Washington, DC: Brookings Institution

Braithwaite, Alex and Dougal Lemke. 2011. "Unpacking Escalation" *Conflict Management and Peace Science* 28(2): 111-123

Brecher, Michael. 1996. "Crisis Escalation: Model and Findings" *International Political Science Review* 17(2): 215-230

Buchanan, Ben. 2017. *The Cyber Security Dilemma: Hacking, Trust, and Fear Between Nations* (New York: Oxford University Press)

Buchanan, Ben. 2017. "The Cybersecurity Dilemma: Where Thucydides Meets Cyberspace" Net Politics 30 January

Bueno de Mesquita, Bruce, James Morrow, and Ethan Zorick. 1997. Capabilities, perception, and escalation. *American Political Science Review* 91(1): 15–27

Carson, Austin. 2018. *Covert Conflict in International Politics* (Princeton: Princeton University Press)

Colaresi, Michael, and William Thompson. 2002. Strategic rivalry, protracted conflict, and crisis escalation. *Journal of Peace Research* 39(3): 263–287

Conti, Gregory and David Raymond. 2017. *On Cyber: Towards an Operational Art for Cyber Conflict* (New York: Kopidion Press)

Delucchi, K. L. (1993). On the use and misuse of chi-square. In G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Statistical issues* (pp. 295-320). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc

Druckman, Daniel. 1994. Tools for discovery: Experimenting with simulations. *Simulation & Gaming* 25:446-55

Fleiss, J.L. 1981. *Statistical Methods for Rates and Proportions* (New York: John Wiley)

Gartzke, Eric. 2013. "The Myth of Cyber War: Bringing War in Cyberspace Back Down to Earth" *International Security* 38: 41-73

Gartzke, Eric and Jon Lindsay. 2015. "Weaving Tangled Webs: Offense, Defense, and Deception in Cyberspace" *Security Studies* 24: 316-348

Gartzke, Erik, and Jon R. Lindsay, eds. Cross-Domain Deterrence: Strategy in an Era of Complexity. Oxford University Press, 2019.
Geller, Daniel. 1990. "Nuclear Weapons, Deterrence, and Crisis Escalation" *The Journal of Conflict Resolution* 34(2): 291-310

Ghosn, Faten, Glenn Palmer, and Stuart Bremer. 2004. The Militarized Interstate Dispute data sets, version 3.0: Procedures, coding rules, and description. *Conflict Management and Peace Science* 21(2): 133–154

Gomez, M. A., & Villar, E. B. (2018). Fear, uncertainty, and dread: Cognitive heuristics and cyber threats. *Politics and Governance, 6*(2), 61-72

Green, Donald P., and Alan S. Gerber. 2002. "Reclaiming the experimental tradition in political science" In *Political science: State of the discipline III*, edited by Helen Milner and Ira Katznelson. Washington, DC: American Political Science Association

Gross, M. L., Canetti, D., & Vashdi, D. R. (2017). Cyberterrorism: its effects on psychological well-being, public confidence and political attitudes. *Journal of Cybersecurity*, *3*(1), 49-58.

Hensel, Paul. 1996. Charting a course to conflict: Territorial issues and interstate conflict, 1816–1990. *Conflict Management and Peace Science* 15(1): 43–74

Howell, D.C. (2002). *Statistical Methods for Psychology* (5th ed.). Pacific Grove CA: Duxbury

Huth, Paul and Bruce Russett. 1984. "What Makes Deterrence Work? Cases from 1900 to 1980" *World Politics* 36: 496-526

Huth, Paul and Bruce Russett. 1988. "Deterrence Failure and Crisis Escalation" *International Studies Quarterly* 32: 29-45

Jensen, Benjamin. 2017. "The Cyber Character of Political Warfare" *Brown Journal of World Affairs* 24: 159-171

Jensen, Benjamin and David Banks. 2018. *Cyber Operations in Conflict: Lessons from Analytical Wargames* (Berkely: Center for Long-Term Cyber Security)

Jervis, Robert. 1978. "Cooperation Under the Security Dilemma" *World Politics* 30(2): 167-214

Khan, Herman. 1965. *On Escalation: Metaphors and Scenario* (New York: Praeger)
Kugler, Jacek, and Douglas Lemke, eds. 1996. *Parity and War*. Ann Arbor, MI: University of Michigan Press

Kello, Lucas. 2017. *The Virtual Weapon and International Order* (New Haven: Yale University Press)

Kreps, S., & Das, D. (2017). Warring from the virtual to the real: Assessing the public's threshold for war over cyber security. *Research & Politics*. https://doi.org/10.1177/2053168017715930

Kreps, Sarah E. and Schneider, Jacquelyn, *Escalation Firebreaks in the Cyber, Conventional, and Nuclear Domains: Moving Beyond Effects-Based Logics* (January 17, 2018). http://dx.doi.org/10.2139/ssrn.3104014

Kurizaki, Shuhei. 2016. "Signaling and perception in international crises: Two Approaches" *Journal of Theoretical Politics* 28(4): 625-654

Kydd, Andrew. 2005. *Trust and Mistrust in International Politics* (Princeton: Princeton University Press)

Leng, Russell, and J. David Singer. 1988. Militarized interstate crises: The BCOW typology and its applications. *International Studies Quarterly* 32(2): 155–173

Leng, Russell, and J. David Singer. 1988. Militarized interstate crises: The BCOW typology and its applications. *International Studies Quarterly* 32(2): 155–173

Lindsay, Jon R., and Erik Gartzke. "Coercion through cyberspace: the stability-instability paradox revisited." The Power to Hurt: Coercion in Theory and in Practice, Greenhill KM and Krause PJP (eds.). New York: Oxford University Press, Forthcoming (2016).

Maoz, Zeev. 1983. "Resolve, Capabilities, and the Outcomes of International Disputes, 1816-1976. " *Journal of Conflict Resolution* 27: 195

Maoz, Zeev, and Nasrin Abdolali. 1989. Regime types and international conflict, 1816–1976.

Journal of Conflict Resolution 33(1): 3–35

Maoz, Zeev, and Bruce Russett. 1993. Normative and structural causes of democratic peace. *American Political Science Review* 87(3): 624–638

Marchant-Shapiro, Theresa. 2015. “Chi-Square and Cramer’s V: What do You Expect?” *Statistics for Political Analysis: Understanding the Numbers* (New York: Sage, 2015)

Mauer, Tim. 2018. *Cyber Mercenaries: The State, Hackers, and Power* (New York: Cambridge University Press)

McDermott, Rose. 2002. “Experimental Methods in Political Science” *Annual Review of Political Science* 2:31-61.

Meyers, Adam. 2016. “Cyber Skirmish: Russia v. Turkey” *Crowdstrike*

Mintz, Alex, Nehemia Geva, Steven B. Redd, and Amy Carnes. 1997. “The effect of dynamic and static choice sets on political decision making: An analysis using the decision board platform.” *American Political Science Review* 91:553-66

Partell, Peter. 1997. “Escalation at the outset: An analysis of targets’ responses in militarized interstate disputes.” *International Interactions* 23(1): 1–35.

Pauly, Reid. 2018. “Would U.S. Leaders Push the Button? Wargames and the Sources of Nuclear Restraint” *International Security* 43: 151-192

Perla, Peter. 1990. *The Art of Wargaming: A Guide for Professionals and Hobbyists* (Annapolis: U.S. Naval Institute Press)

Pournelle, Victor. 2017. “Designing Wargames for the Analytical Purpose” *Phalanx* 50: 48-53

Prins, Brandon. 2001. Domestic politics and interstate disputes. *International Interactions* 26(4): 411–438.

Prins, Brandon, and Christopher Sprecher. 1999. Institutional constraints, political opposition, and interstate dispute escalation. *Journal of Peace Research* 36(3): 259–287.

Rattray, Gregory. 2001. *Strategic Weapons in Cyberspace* (Cambridge: MIT Press)

Sample, Susan G. "Arms races and dispute escalation: Resolving the debate." Journal of Peace Research 34.1 (1997): 7-22.

Ramsay, Kristopher. 2018. “Information, Uncertianty and War” Annual Review of Political Science 20: 505-27

Schelling, Thomas. 1960. *Strategy of Conflict* (Cambridge: Harvard University Press)

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

Schneider, Jacquelyn. 2018. "What War Games Tell Us About the Use of Cyber Weapons in a Crisis" *Net Politics* June 21

Senese, Paul. 1996. Geographical proximity and issue salience: Their effects on the escalation of militarized interstate conflict. *Conflict Management and Peace Science* 15(2): 133–162.

Senese, Paul. 1997. Between dispute and war: The effect of joint democracy on interstate conflict escalation. *Journal of Politics* 59(1): 1–27

Senese, Paul, and John Vasquez. 2003. A unified explanation of territorial conflict. *International Studies Quarterly* 47(2): 275–298

Senese, Paul, and John Vasquez. 2008. *The Steps to War: An Empirical Study*. Princeton, NJ: Princeton University Press

Sheskin, D. (2011). *Handbook of Parametric and Nonparametric Statistical Procedures*. Boca Raton, FL: Chapman & Hall/CRC

Slayton, Rebecca. 20xx. "What is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment" *International Security* 41: 72-91

Snyder, Glen and Paul Diesing. 1977. *Conflict among nations: Bargaining, decision making, and system structure in international crises* (Princeton: Princeton University Press)

"Pulwama attack: Pakistani websites hacked, here's the list" *The Times of India* February 18, 2019

Thompson, William. 2001. Identifying rivals and rivalries in world politics. *International Studies Quarterly* 45(4): 557–586

Toft, Monica. 2014. "Territory and War" *Journal of Peace Research*

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

Van Crevald, Martin. 2013. *Wargaming: From Gladiators to Gigabytes* (New York: Cambridge University Press)

Vasquez, John. 1993. *The War Puzzle*. (New York: Cambridge University Press, 1993)

Vasquez, John, and Marie Henehan. 2001. Territorial disputes and the probability of war, 1819–1992. *Journal of Peace Research* 38(2): 123–138

Ward, Michael. 1982. "Cooperation and Conflict in Foreign Policy Behavior" *International Studies Quarterly* 26: 8-126

Wilkenfeld, Jonatha, Kathleen Young, Victor Asal, David Quinn. 2003. “Mediating International Crises: Cross-National and Experimental Perspectives” *Journal of Conflict Resolution* 471: 279-301

Work, J.D. and Benjamin Jensen. “Cyber-Civil Military Relations: Balancing Interests on the Digital Frontier” *War on the Rocks* September 4, 2018