
Escalation Management in the Early Stages of International Crisis



19 November 2024

STRATEGIC MULTILAYER ASSESSMENT

Authored by: Dr. Kyle Beardsley, Dr. Jonathan Wilkenfeld, Phuong Pham, Corinne DeFrancisci, Dr. Diana Partridge, Megan Rutter
Series Editor: Sarah Canna, NSI Inc.

This paper was written for Strategic Multilayer Assessment's 21st Century Strategic Deterrence Frameworks project supporting USSTRATCOM.

Biographies

Dr. Kyle Beardsley is a Professor of Political Science at Duke University. He is Co-Director of the International Crisis Behavior data project, and the Director of the Triangle Institute of Security Studies (TISS). His research focuses on the quantitative study of international conflict and peace processes. He is particularly interested in questions related to the role of third parties in shaping conflict dynamics, the interdependence of networks of conflict and cooperation, the links between armed conflict and gender power imbalances, and the impact of nuclear weapons on international crisis behavior.

Beardsley's research appears in leading journals of political science and international relations. His first book, *The Mediation Dilemma*, examines why mediation tends to do well in the short run to abate conflict, but, in the long run, mediated settlements tend to be more fragile. In a book with Sabrina Karim, *Equal Opportunity Peacekeeping*, they examine the consequences of and potential solutions to gender power imbalances in peace operations. Beardsley also co-edited and contributed to a *Research Handbook on Mediating International Crises*. The handbook is an essential reference for postgraduate researchers and practitioners of international politics and conflict negotiation.

Dr. Jonathan Wilkenfeld is a Professor Emeritus and prior Chair of Government and Politics at the University of Maryland and Founding Director of the ICONS simulation project. He is currently a Research Professor at START.

He is a specialist in foreign policy decision-making, crisis behavior, and mediation, as well as in the use of simulation in political science. From 1977 to 2017, Wilkenfeld served as co-Director (with Michael Brecher) of the International Crisis Behavior Project, a cross-national study of international crises in the twentieth century. The project has served as the basis for systematic research into a range of crucial foreign-policy issues, including state motivations during times of crisis, conflict management practices, and protracted conflict trajectories.

Wilkenfeld served as Director of the International Communication and Negotiation Simulations (ICONS) Project until 2017. ICONS provides decision-makers with interactive training experiences in the fields of conflict behavior, negotiation, and crisis management. The development of the ICONS Project grew out of his long-term interest in integrating technology and simulation techniques into the teaching of negotiation and international politics. 2009, Wilkenfeld was designated by the University of Maryland as a Distinguished Scholar-Teacher.

His most recent books include *Myth and Reality in International Politics* (2016); *Research Handbook on Mediating International Crises* (2019 with Kyle Beardsley and David Quinn), and *Escalation Management in International Crises* (2023). Wilkenfeld's current work focuses on escalation and deterrence in international crises.

Escalation Management in the Early Stages of International Crisis

Kyle Beardsley,¹ Jonathan Wilkenfeld, Phuong Pham, Corinne DeFrancisci, Diana Partridge, Megan Rutter

This report evaluates how crisis actors respond to adversary actions in the early stages of an international crisis. We specifically consider the first three stages of a crisis: the **trigger** act from a **challenger**, the **response** act from the **defender** state which has been triggered, and then the **counter-response** act from the **challenger** (see Figure 1). The latter two stages constitute the outcomes of interest: the defender's response to a challenge and the challenger's counter-response to a defender's response. The propensity for defenders to respond with violence or for challengers to counter-respond with violence, we contend, should depend on their adversaries' prior actions—respectively, the challenger's use of a violent or non-violent trigger and the defender's use of a violent or non-violent response. Moreover, we consider how observed characteristics of the adversary related to attributes such as military capabilities, alliance military capacity, and domestic instability condition the relationship between prior actions and actor responses.

¹ Contact Information: kyle.beardsley@duke.edu

Executive Summary

How do states' actions in the early stages of an international crisis shape the potential for violent escalation? The opening moves in an international crisis by the challenger and the defender can be considered tone-setters, rich in signals, including about their respective red lines. There is some indication that the early stages of a crisis are the most volatile and prone to escalation because in many cases the crisis actors do not yet have clear understandings of how their actions will be perceived by the other party and what their respective red lines are. The common theme among states embroiled in an international crisis is a significant rise in the level of stress experienced by decision-makers and its impact on the decisions they take.

When a crisis erupts, general deterrence has failed, and the crisis actors must now use their actions to convince their adversary to back down—a type of deterrence called “immediate deterrence.” Existing scholarship has posited that what makes states successful at general deterrence is different from what makes states successful at immediate deterrence because a challenging state that escalates a crisis will have already accounted for the observed characteristics of the defender before making their challenge. Little is known, however, regarding how a defender's observed characteristics condition the information revealed by the defender's actions. The existing literature also has not well examined how the challenger's observed characteristics condition the information revealed by the challenger's action.

This report evaluates how crisis actors respond to adversary actions in the early stages of an international crisis. We specifically consider the first three stages of a crisis: the **trigger** act from a **challenger**, the **response** act from the **defender** state which has been triggered, and then the **counter-response** act from the **challenger** (see Figure 1). The latter two stages constitute the outcomes of interest: the defender's response to a challenge and the challenger's counter-response to a defender's response. The propensity for defenders to respond with violence or for challengers to counter-respond with violence, we contend, should depend on their adversaries' prior actions—respectively, the challenger's use of a violent or non-violent trigger and the defender's use of a violent or non-violent response. Moreover, we consider how observed characteristics of the adversary related to attributes such as military capabilities, alliance military capacity, and domestic instability condition the relationship between prior actions and actor responses.

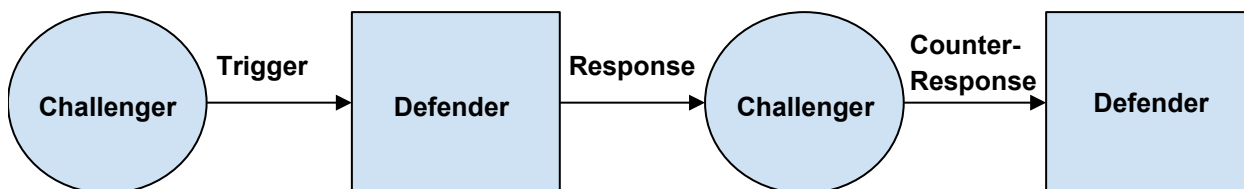


Figure 1. Early stages of an international crisis

We analyzed data on all international crises from 1963 to 2019. During this 57-year period, there were 300 crises involving a myriad of state adversaries, ranging from minor powers involving regional and local interests, to major and superpowers locked in competition over international dominance across vast regions of the globe. Explanatory variables include military spending, number of alliance ties, stakes, ally capacity, level of democracy and regime durability, human rights record, and women's empowerment record. While no two crises are the same, taken in aggregate, a better understanding of the behavior of the leaders of these states, and the choices they made when confronted with crisis conditions, can help us identify general trends in their behavior.

Our findings confirm that deterrence in the earliest stages of an international crisis can be critical to its evolution. To prevail in a crisis without sparking violent escalation, actors must consider the two-edged nature of their escalatory or de-escalatory actions, as they try to walk the fine lines of A) demonstrating strength without provoking their adversary and B) demonstrating commitment credibility without emboldening their adversary.

We highlight three conclusions that bear on optimal immediate deterrence. **A First Deterrence Principle is to exploit your observable strengths:** For challengers, non-violent actions reduce the risk of escalation if they can demonstrate their own strong observable capabilities and resolve. **A Second Deterrence Principle is to exploit your adversary's weakness:** On average, challengers and defenders alike can be more confident that a crisis will stay non-violent if they can play to their adversaries' domestic and international audiences—specifically, domestic constituents and international allies who prefer to avoid disproportionate escalation. **A Third Deterrence Principle contends that escalate-to-de-escalate strategies are unlikely to work when the stakes are low:** Instead, adversaries are more likely to reciprocate escalatory or de-escalatory actions when the stakes are low, such as in the absence of threats of grave damage or threats to the very existence of the adversary. **In sum, crisis actors are best positioned to avoid unwanted escalation when their own actions are non-violent, as well as A) when they can demonstrate relative strength and resolve, B) when they can point to high belligerence costs likely to be incurred by their adversary, and C) when the stakes are relatively low.**

Introduction

How do states' actions in the early stages of an international crisis shape the potential for violent escalation? The opening moves in an international crisis by the challenger and the defender can be considered tone-setters, rich in signals, including about their respective red lines. There is some indication that the early stages of a crisis are the most volatile and prone to escalation because in many cases the crisis actors do not yet have clear understandings of how their actions

will be perceived by the other party and what their respective red lines are (Wright, 2017; Wilkenfeld & Murauskaite, 2023, p. 120).²

When a crisis erupts, general deterrence has failed, and the crisis actors now must choose their actions to convince their adversary to back down—a type of deterrence called “immediate deterrence” (Huth & Russett, 1993). Existing scholarship has posited that what makes states successful at general deterrence is different from what makes states successful at immediate deterrence because a challenging state that escalates a crisis will have already accounted for the observed characteristics—the known capabilities to project military force—of the defender before making their challenge, creating strong selection effects (Fearon, 1994). Little is known, however, regarding how a defender’s observed characteristics condition the information revealed by its actions. The existing literature also has not well examined how the challenger’s observed characteristics condition the information revealed by the challenger’s actions.

Once a crisis has commenced, crisis actors try to take actions that increase their opponent’s utility for backing down. We focus on two different objectives that actors have in calibrating their choices. One objective that crisis actors strive to achieve is to make costly signals that reinforce the credibility of their escalatory *threats* and their de-escalatory *promises*. With regard to threat credibility, crisis actors try to signal higher-than-expected capabilities and resolve, in hopes that their opponent will update their beliefs about how the crisis will play out in a way that is favorable to the signaling state (Fearon, 1997; Sartori, 2013). This signaling is all part of the mutual learning process that is necessary to ultimately end hostilities (Slantchev, 2003). Crisis actors also have an incentive to make costly signals that reinforce the credibility of their commitments to follow through on their promises. A reputation for reneging on the terms of a settlement can prevent de-escalation and agreement even when the terms are mutually beneficial (Crescenzi, 2018; Kydd, 2007). Actors often signal the credibility of their threats and promises through “tying hands.” Generating audience costs for themselves, should they fail to follow through on a threatened or promised response, is a specific type of costly signal (Fearon, 1997; Quek, 2021). In this way, actors benefit if they can “bind themselves” and commit to a threat that deters their adversary (Schelling, 1980) or commit to a promise that enables de-escalation (Maoz & Falsenthal, 1987).

On the flip side of binding oneself, another objective that crisis actors strive to achieve is to “put their opponent into a bind”—to shape the strategic environment such that escalation, or at least

² A number of leading international relations scholars have stressed the critical importance of these early crisis stages, focusing on early diplomacy (George, 1993), early misperceptions and cognitive bias (Holster, 1972; Jervis, 1976), and the early establishment of the trajectory of a crisis (Brecher, 1993). As Brecher notes, the common theme among states embroiled in an international crisis is a significant rise in the level of stress experienced by decision-makers and its impact on the decisions they take (Brecher, 1993, p. 129).

not backing down, appears unpalatable to the adversary. An actor's early crisis decisions can influence the beliefs of the adversary's domestic stakeholders and international allies such that they would support certain actions but not others.

As states strive to signal the credibility of their threats and promises, and as they try to put their adversary into a bind, their actions should shape immediate deterrence differently depending on some observed attributes. That is, states that act with violence or non-violence early in a crisis should be more or less able to get their adversary to back down depending on the actors' characteristics, such as those related to observed capabilities, potential alliance support, or domestic instability. The actions taken in crisis, in combination with what is known about the states, should affect beliefs about what subsequent actions to take.

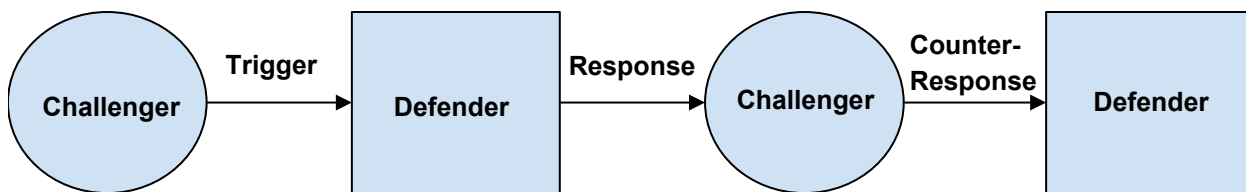


Figure 1: Early stages of an international crisis

This report evaluates how crisis actors respond to adversary actions in the early stages of an international crisis. Depicted in Figure 1, we specifically consider the first three stages of an international crisis: the **trigger** act from a **challenger**, the **response** act from the **defender** state which has been triggered, and then the **counter-response** act from the **challenger**. The latter two stages constitute the outcomes of interest: the defender's response to a challenge and the challenger's counter-response to a defender's response. Note that this language of defender and challenger does not relate to dispositions toward the status quo, as it is sometimes meant in the international politics literature—a defender is simply the crisis actor that is experiencing the foreign policy crisis and making the response, and the challenger is simply the adversary of the defender that is making the trigger and the counter-response moves.

We contend that the propensity for defenders to respond with violence or for challengers to counter-respond with violence should depend on their prior actions in the crisis—respectively, the challenger's use of a violent or non-violent trigger and the defender's use of a violent or non-violent response.³ Moreover, we consider how observed characteristics of the adversary related

³ The defender's response is thus recursive. It is an outcome variable when assessing the relationship between trigger behavior and response behavior, and it is an explanatory variable when assessing the relationship between response behavior and counter-response behavior.

to attributes such as military capabilities, alliance military capacity, and domestic instability condition the relationship between prior actions and actor responses.

Our findings confirm that deterrence in the earliest stages of an international crisis can be critical to its evolution. To prevail in a crisis without sparking violent escalation, actors must consider the two-edged nature of their escalatory or de-escalatory actions, as they try to walk the fine lines of A) demonstrating credible threats to inflict harm without provoking their adversary and B) demonstrating credible promises to pursue de-escalation without emboldening their adversary.

We highlight three conclusions that bear on optimal immediate deterrence. First, for challengers, non-violent actions reduce the risk of escalation if they can demonstrate their own strong observable capabilities and resolve. Second, challengers and defenders alike can be more confident that a crisis will stay non-violent if they can play to their adversaries' domestic and international audiences—specifically, domestic constituents and international allies who prefer to avoid disproportionate escalation. Third, adversaries are more likely to reciprocate escalatory or de-escalatory actions when the stakes are low, exemplified by the absence of grave damage or threat to the very existence of the adversary. **In sum, crisis actors are best positioned to avoid unwanted escalation when their own actions are non-violent; and A) when they can demonstrate relative strength and resolve, B) when they can point to high belligerence costs likely to be incurred by their adversary, and C) when the stakes are low.**

Theoretical Argument

States have at least two objectives when choosing how to respond or counter-respond to an adversary. One objective is to make costly signals about their capability, resolve, and commitment credibility. Another is to put their adversary in a bind.⁴

Signaling Advantageous Information

Starting with the objective to signal advantageous information about themselves, we assume that states' threat credibility and promise credibility shape decisions over whether to escalate or de-escalate. Threat credibility refers to beliefs about a state's **capabilities**—relative advantages in abilities to impose costs on the other side—and **resolve**—relative advantages in abilities to absorb

⁴ These objectives are not exhaustive, especially since psychological mechanisms are likely also at work to explain how decision makers respond to stressful situations with high stakes and time constraints (Holsti, 1972; Brecher, 1993). Moreover, state behavior in international crises can result from other relationships besides the immediate relationship with a given adversary; dyads are almost never isolated (Gartzke & Gleditsch, 2002; Cranmer & Desmarais, 2016; Minhas et al., 2022). For theoretical and empirical traction, we focus on objectives that pertain to state-level characteristics. We do consider extra-dyadic relationships that the crisis actors have with allies and domestic constituents, and the research design accounts for potential confounders and sources of non-independence among the observations.

costs and hold out for better terms. Capabilities and resolve both affect the net benefits (after discounting the costs) of fighting, so we consider them in parallel when developing hypotheses for testing.⁵ Promise credibility, related to trustworthiness, pertains to beliefs about whether a state will follow through on an agreed settlement or other de-escalatory course of action (Cebul et al., 2021; Kydd, 2007).

Putting it all together, a state will be more likely to back down against an adversary that is able to signal an ability to impose more costs than it incurs while fighting, to hold out for a better deal, and to better commit to a path of de-escalation. States will be less likely to back down if the balance of capabilities and the balance of resolve are revealed to be more favorable than previously believed, or if they increasingly suspect that their moves toward de-escalation will be exploited by the other side, for example, by making further demands or advances. Important to this logic is the **trajectory of learning** —as a crisis unfolds, the actors update their beliefs about the costs and benefits of subsequent actions. For a state to be more willing to back down at a later time than it was at the start of the crisis, there must be some new information that the state is responding to.

Zagare and Kilgour (1993; 2000) present formal models of crisis escalation as states strive to signal capabilities and resolve to one another. States are incentivized to signal strength as a means to convince their opponent to back down. Perceptions are key, as there can be mismatches between the perceived capabilities or resolve and the actual capabilities or resolve.

So, states have an incentive to try and convince their adversary that they are more militarily capable, resolved, and trustworthy than was believed at the start of the crisis. However, a dilemma arises because demonstrating self-restraint will help signal promise credibility, but self-restraint may also be interpreted as demonstrating weak capabilities or resolve and thus low threat credibility. Similarly, quick escalation to violence may signal strong threat credibility, but it may also signal low promise credibility by sowing doubts about whether the actor can be a reasonable partner in de-escalating and finding a mutually satisfactory solution.

Moreover, a dilemma arises when states try to signal capabilities and resolve through escalation but actually tilt the balance of resolve in their adversary's favor by endogenously affecting the audience costs that the adversary would face from backing down. By escalating to demonstrate capabilities and resolve, the adversary might become less prone to back down because key domestic and international audiences (allies) would worry that backing down in the face of escalation will hurt the state's reputation, reduce general deterrence efficacy, and invite more

⁵ That is, we assume that increases in both an actor's capabilities and resolve have similar relationships with the crisis behavior outcomes.

challenges in the future. Relatedly, the adversary might be provoked by aggressive action early in a crisis (Dafoe et al., 2021)

The concept of the trajectory of learning in a crisis can help unwind the dilemmas and form expectations about when actors will be more or less willing to back down after observing their adversary's actions. We focus on the cases in which a defender acts "against type" because they are cases in which the challenger is able to learn about the adversary and become more or less willing to back down or escalate. When states are known to have high capabilities and resolve and yet take action to try to de-escalate the crisis, this can signal promise credibility without sacrificing threat credibility, and thereby increase the willingness of a challenger to back down (Cebul et al., 2021). This expectation relates to existing understandings about the importance of negotiating from a position of high resilience (Dudley, 2023; Mastro, 2019; Mastro & Siegel, 2023). It also relates to a finding in which states that face high costs—they have low relative military capability—substitute low-cost actions, thus making it more difficult to reach a settlement (Spaniel & İdrisoğlu, 2023). Conversely, when states are known to have high promise credibility—for example, on the basis of regime type (Weeks, 2012)—yet take actions to try to escalate the crisis, this can signal high threat credibility without being perceived as overly provocative (Debs & Weiss, 2016; McManus, 2014).

The logic thus far yields several hypotheses. First, we expect that when an actor's adversary has high levels of observed capabilities and resolve—the ability to gain much from a conflict and the ability to bear the costs of conflict—the adversary's de-escalatory actions will make it more likely that the actor will be willing to de-escalate as well. Military capability is one observed attribute of an adversary that actors will use when forming their beliefs. Great powers with many "junior" partners also will be perceived as having relatively high resolve, since they have an incentive to show strength as a means to assure their partners and future challengers that they are sufficiently resolved in the event of a crisis situation (Sechser, 2013). Finally, the stakes of the crisis inform the crisis actor's resolve—adversaries fearing a severe threat will be known to have a greater interest in prevailing than adversaries facing lesser threats, with less on the line (Gurantz & Hirsch, 2017).

Alliance military capacity is another attribute that actors assess. States with strong support from allies will tend to be perceived as having high resolve because they can count on an ally to at minimum backstop their efforts to prevent worst-case-scenarios, and also at times to lend assistance in achieving their crisis objectives (Leeds, 2003; Kang, 2017; Johnson et al., 2015; Wolford, 2023). While having allies with high military capacity can enhance perceptions of capabilities and resolve, having allies that are undergoing large changes in their military capacity could attenuate such perceptions. Recent scholarship has studied the conditions in which certain alliances are more credible than others and has pointed to the importance of military power shifts in an alliance as a key determinant of alliance instability (Krainin & Schub, 2021; Johnson & Joiner,

2021). Putting the two pieces together, adversaries that have allies with both high and stable military capacity will be perceived as having high capability and resolve.

From these expectations, a number of testable hypotheses emerge for both the defender's response to a trigger and the challenger's counter-response to a defender's response.

H1: Defenders will be more likely to make a non-violent response when the challenger has made a non-violent trigger and is also perceived as having high capability and resolve. This latter condition will adhere when:

H1a: The challenger has high military spending

H1b: The challenger has many alliance ties

H1c: The challenger has high stakes

H1d: The challenger's allies have high and stable military capacity

H2: Challengers will be more likely to make a non-violent counter-response when the defender has made a non-violent response and is also perceived as having high capability and resolve. This latter condition will adhere when:

H2a: The defender has high military spending

H2b: The defender has many alliance ties

H2c: The defender has high stakes

H2d: The defender's allies have high and stable military capacity

Additional expectations result when considering that adversaries which have high promise credibility will be better able to convince an actor to back down when the adversary shows a willingness to escalate. As one measure of promise credibility, we draw from the literature that has found that democracies tend to be more credible in following through on their commitments because their leaders are more likely to suffer audience costs for backtracking (Leeds & Mattes, 2002; Fearon, 1997; Renshon et al., 2023). Another characteristic that shapes perception of promise credibility in the eyes of an adversary is domestic stability (Di Lonardo & Tyson, 2022). States that face the potential for irregular regime change will be less credible in following through on their state's intentions. States with little likelihood of irregular regime change will be more credible.

H3: Defenders will be more likely to make a non-violent response when the challenger has made a violent trigger and is also perceived as credible in its commitments. This latter condition will adhere when:

H3a: The challenger is a democracy

H3b: The challenger has demonstrated regime stability

H4: Challengers will be more likely to make a non-violent counter-response when the defender has made a violent response and is also perceived as credible in its commitments. This latter condition will adhere when:

H4a: The defender is a democracy

H4b: The defender has demonstrated regime stability

Putting an Adversary into a Bind

The second objective that states strive to achieve through their early crisis behavior is to make it costly for an adversary to continue taking a hardline stance. This means trying to raise an adversary's belligerence costs (Kertzer & Brutger, 2016) when domestic audiences do not support an aggressive foreign policy because they do not deem it called for. Relatedly, states might try to raise an adversary's international audience costs by appealing to patrons of the adversary who fear entrapment—being pulled into a conflict by an ally's recklessness (Snyder, 1984). To raise an adversary's audience costs for escalation, a state might purposefully try to calibrate its own actions to be perceived as non-threatening and measured in its handling of the dispute, at least in the eyes of the adversary's domestic and international audiences (Jervis, 1976; Levin-Banchik, 2021).

One downside of such an approach—in addition to the downside discussed above of potentially signaling weak capabilities or resolve—is that the adversary may perceive little need to resolve the dispute. By trying to appear less threatening, the actor may place their adversary in a bind to not escalate the crisis, but the actor may also reduce the incentive of the adversary to actually back down (Mastro, 2019; Mastro & Siegel, 2023).

In light of this downside, in order to propose observable implications, we posit that attempts to generate high belligerence costs in the adversary can have more traction with some audiences than others. Adversaries that have strong human rights records at home are more likely to have domestic audiences that prefer less unnecessary violence. That is, scholarship has argued that the types of states that have been able to secure human rights protections at home are the types in which the structures and norms favor peaceful dispute resolution (Caprioli & Trumbore, 2006; Sobek et al., 2006). Related, recent scholarship has focused on specific dimensions of human

rights related to women’s empowerment and gender equality that are indicative of norms and structures conducive to peaceful dispute resolution (Cohen & Karim, 2022; Hudson et al., 2020; Sjoberg, 2013). In terms of international audiences, adversaries with allied patrons who have many alliances will be especially sensitive to the problem of entrapment and will urge restraint from a challenger they support when the defender has demonstrated its own restraint.

H5: Defenders will be more likely to make a non-violent response when the challenger has made a non-violent trigger and when the defender is influenced by audiences who have a high value for restraint. This latter condition will adhere when:

H5a: The defender has a strong human rights record

H5b: The defender has a strong women’s empowerment record

H5c: The defender has an alliance with a patron that has many alliance ties

H6: Challengers will be more likely to make a non-violent counter-response when the defender has made a non-violent response and when the challenger is influenced by audiences who have a high value for restraint. This latter condition will adhere when:

H6a: The challenger has a strong human rights record

H6b: The challenger has a strong women’s empowerment record

H6c: The challenger has an alliance with a patron that has many alliance ties

Research Design

To test the hypotheses, we use original data that extends the International Crisis Behavior (ICB) data project.⁶ In the ICB data, an actor experiences a foreign policy crisis when it experiences a threat to basic values, a finite time to respond, and a heightened perceived probability of military hostilities. At the moment when an actor experiences these three criteria, that actor has experienced a crisis trigger. For the purposes of this analysis, the crisis actor under observation is the defender, and the state that is the primary source of threat in the crisis is the challenger. Crises with more than one actor will have multiple observations—each crisis actor will have an

⁶ Original ICB data is available from *International crisis behavior data codebook* (Version 15) [Data Set]. Brecher et al. (2023). <http://sites.duke.edu/icbdata/data-collections/>. Data collection specific for this report was supported by the Minerva Research Initiative and the Asymmetric Threat Analysis Center of the U.S. Department of Defense. Data is available upon request and will be made publicly available upon publication.

observation in which it is the defender, and these actors will usually also appear as challengers for the observations in which their adversary is a defender.

Our dependent variables are two binary variables: violent response and violent counter-response. Whether the defender's major response is violent (V) or nonviolent (NV) constitutes the dependent variable of interest for H1, H3, and H5. Whether the challenger's counter-response is violent or nonviolent constitutes the dependent variable of interest for H2, H4, and H6. To define these variables, information about the tools that the defender employed in its response to a crisis, and information about the tools that the challenger employed in its counter-response to the response was collected. We counted a response or counter-response as violent if violent military action was taken either directly against the crisis adversary or indirectly against an adversary's proxy.

Regarding the explanatory variables, we want to measure a few key characteristics of the actors that condition the relationships between earlier instances of violence or non-violence and subsequent actions of violence or non-violence. Table 1 lists the explanatory variables that pertain to the respective hypotheses. *Military spending* is measured by Barnum et al. (2022), using their imputation of the "NMC" estimator. The *number of alliance ties* is coded based on the number of offensive or defensive formal alliance ties a state has, from version 5.1 of Leeds et al. (2002). For *stakes of the crisis*, we define a dichotomous measure from the ICB actor-level data as whether the gravity threat that an actor experiences is either a "threat of grave damage" or a "threat to existence." *Ally military capacity* is the summed military expenditures of a state's formal offensive and defensive allies. A regime's *level of democracy* is measured by the Polyarchy index from the Varieties of Democracy (V-Dem) data.⁷ *Regime stability* is measured by the number of riots experienced, coded by the Cross National Time Series data.⁸ *Human rights record* is measured by the Physical Violence index from the V-Dem data, and *women's empowerment* is measured by the women's political empowerment index from V-Dem.⁹ To measure the *alliance ties of an ally*, we code the maximum number of offensive or defensive formal allies that a state's offensive or defensive allies have.

⁷ See *V-Dem [Country-Year/Country-Date] Dataset* (Version 13) [Data Set]. Coppedge et al. (2023). <https://www.v-dem.net/data/the-v-dem-dataset/>. We used the "v2x_polyarchy" variable.

⁸ From *V-Dem [Country-Year/Country-Date] Dataset*(Version 13) [Data Set]. Coppedge et al. (2023). <https://www.cntsdata.com/the-data>.

⁹ See *V-Dem [Country-Year/Country-Date] Dataset* (Version 13) [Data Set]. Coppedge et al. (2023). <https://www.v-dem.net/data/the-v-dem-dataset/>. We used the "v2x_clphy" and "v2x_gender" and variables.

Table 1: Explanatory variables

Hypothesis	Symbol	Description
H1a	M _c	Challenger military spending
H2a	M _d	Defender military spending
H1b	A _c	Challenger number of alliance ties
H2b	A _d	Defender number of alliance ties
H1c	S _c	Challenger crisis stakes
H2c	S _d	Defender crisis stakes
H1d	AM _c	Challenger allies' military capacity
H2d	AM _d	Defender allies' military capacity
H3a	D _c	Challenger level of democracy
H3b	D _d	Defender level of democracy
H4a	R _c	Challenger regime durability
H4b	R _d	Defender regime durability
H5a	H _d	Defender human rights index
H6a	H _c	Challenger human rights index

H5b	W_d	Defender women's empowerment index
H6b	W_c	Challenger women's empowerment index
H5c	AA_d	Defender allies' number of alliances
H6c	AA_c	Challenger allies' number of alliances

We use bivariate probit models to test our hypotheses. In a bivariate probit model, the two equations that pertain to the respective dependent variables are estimated simultaneously while controlling for the correlation in the error terms. By controlling for the correlation in the error terms, the model controls for unobservable processes that are at work in shaping both the defender responses and the challenger counter-responses. Moreover, the simultaneous estimation improves the efficiency of the estimation when compared to separate regressions (Amemiya, 1978). We control for both challenger and defender military spending and polyarchy index in all equations since military capacity and regime type are key sources of heterogeneity in crisis escalation dynamics. Robust standard errors, clustered on the crisis number, are estimated because observations from the same crisis are not independent of one another.

Results

With so many hypotheses, we focus our discussion on the hypotheses that were confirmed.¹⁰ Since all of the hypotheses pertain to interactive effects in which the marginal effect of the challenger trigger or defender response is conditioned by the explanatory variables, we present marginal effects plots. When the outcome variable is the defender response, the marginal effects plots display the correlations between A) a (non-)violent challenger trigger and B) a non-violent defender response. When the outcome variable is the challenger counter response, the marginal effects plots display the correlations between A) a (non-)violent defender response and B) a non-violent challenger counter response, while varying the values of the conditioning explanatory variables.¹¹

¹⁰ The full bivariate probit results are available upon request.

¹¹ For visual clarity, we show the marginal effects that pertain to the counter-response outcomes while holding fixed the level of violence in the trigger. That is, the underlying models actually estimate the marginal effects for each different

Table 2: Summary of the Findings

H#	Relevant Marginal Effect	Condition	Result
1a	NV trigger → NV response	High M_c	Confirmed
2a	NV response → NV counter	High M_d	Not confirmed
1b	NV trigger → NV response	High A_c	Confirmed
2b	NV response → NV counter	High A_d	Not confirmed
1c	NV trigger → NV response	High S_c	Not confirmed
2c	NV response → NV counter	High S_d	Not confirmed
1d	NV trigger → NV response	High AM_c , High ΔAM_c	Not confirmed
2d	NV response → NV counter	High AM_d , High ΔAM_d	Not confirmed
3a	V trigger → NV response	High D_c	Not confirmed
4a	V response → NV counter	High D_d	Not confirmed
3b	V trigger → NV response	High R_c	Not confirmed
4b	V response → NV counter	High R_d	Not confirmed

pairing of triggers and responses (i.e., non-violent trigger, non-violent response; non-violent trigger, violent response; violent trigger, non-violent response; violent trigger, violent response). For the graphical representations, we just focus on how movement from one of those pairings to the other pairing with the same trigger is correlated with the counter-response outcome.

5a	NV trigger → NV response	High H _d	Not confirmed
6a	NV response → NV counter	High H _c	Confirmed
5b	NV trigger → NV response	High W _d	Not confirmed
6b	NV response → NV counter	High W _c	Not confirmed
5c	NV trigger → NV response	High AA _d	Confirmed
6c	NV response → NV counter	High AA _c	Confirmed

Table 2 summarizes the expected relationships and the findings. We see that some of the hypotheses are confirmed, while many are not. The lack of support for many of the hypotheses is not surprising, given the data demands of having interaction effects across multiple equations. There are eight possible triples of triggers, responses, and counter-responses—each coded as violent or non-violent—and thus it is challenging to establish relationships between the cells of a 2 x 2 x 2 array of the possible triples while conditioning on the other explanatory variables.

Finding 1: Non-violent triggers are more likely to lead to non-violent responses when the challenger has high levels of observed capabilities and resolve

The findings confirm H1a and H1b. As we observe in Figure 2, defenders are more likely to reciprocate a non-violent trigger with a non-violent response when the challenger has high military capacity or when the challenger has many alliances. In such situations, the challenger’s relatively restrained means of issuing a challenge is paired with observable strength and resolve and thus is not likely taken as a sign of weakness that might embolden the defender to make a violent response. Restraint is met with restraint when the first instance of restraint is not interpreted as a sign of weakness.

On October 4, 2002, U.S. Assistant Secretary of State for East Asia and Pacific Affairs James Kelley led a delegation to Pyongyang and accused North Korea of obtaining technology for the manufacture of nuclear weapons. This verbal act triggered a crisis for North Korea, which responded the following day by admitting that it had indeed acquired the technology, thus triggering a crisis for the United States. In this case, the non-violent nature of the trigger would not have led to much updating of North Korean beliefs about US capabilities and resolve.

Other examples of crises meeting the criteria of Findings 1 include: 1) The United States triggered a crisis for Libya in the Gulf of Syrte Crisis of 1981; 2) the United States triggered a crisis for Iraq

in the Iraq No-Fly Zone crisis of 1992; 3) the IAEA (and indirectly the United States) triggered a crisis for North Korea in the North Korea Nuclear crisis of 1993; 4) the United States triggered a crisis for Iraq in the Iraq Regime Change Crisis of 2002; 5) the United States triggered a crisis for North Korea in North Korea Nuclear II crisis in 2002; and 6) South Ossetia (and indirectly Russia) triggered a crisis for Georgia in the South Ossetia-Abkhazia crisis of 2004.

Several noteworthy recent crises did not fit this pattern and can therefore be designated as failed deterrence. In the Georgia-Abkhazia crisis of 1992, in the midst of armed clashes between Georgian forces and Abkhazian separatists supported by Russia, Russia’s announcement on 25 September of the suspension of arms shipments to Georgia triggered a crisis for Georgia, which ultimately responded with the seizure of a Russian arms depot later that year. In the Russo-Georgian War crisis of 2008, despite Russia’s high observable capabilities, a Russian troop movement inside South Ossetia in support of pro-Russian separatists on 7 August triggered a crisis for Georgia, which responded the next day with a mobilization and a declaration of war. In the Turkey-Russian Jet Incident of 2015 in the midst of the Syrian civil war, a Russian fighter briefly violated Turkish airspace on 24 November and was shot down.

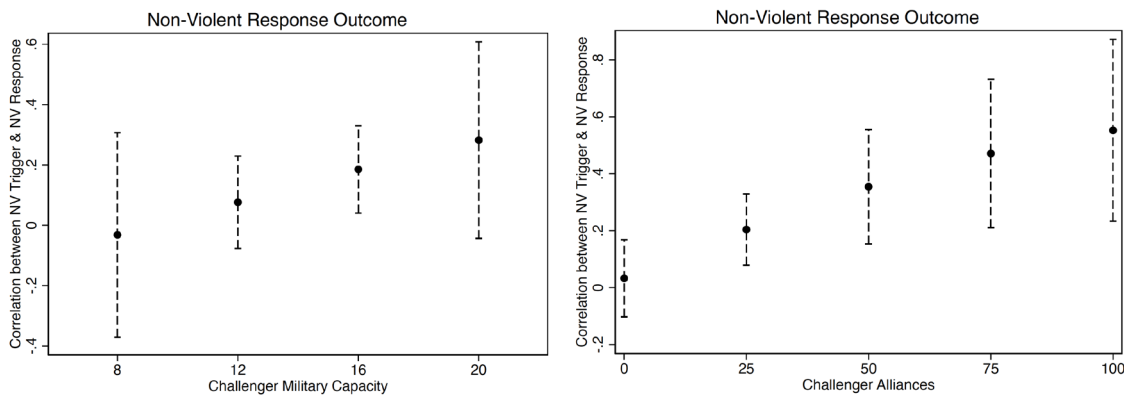


Figure 2: Conditioning effects of challenger military capacity and challenger alliances

Finding 2: Actors are more likely to meet non-violent triggers and responses with non-violence when they are prone to suffer belligerence costs

The findings confirm H5c, H6a and H6c. As we see in Figure 3, a non-violent defender response is more strongly correlated with a non-violent challenger counter-response when the challenger has audiences which have a higher preference for more restrained, less violent behavior. Specifically, when the challenger has strong norms regarding human rights, it is more likely to reciprocate non-violence with non-violence. This comports well with the potential for domestic audiences that are used to a government that is restrained in its use of violence domestically to expect that the state would also only use violence internationally when sufficiently provoked.

When the challenger has an ally with many allies that it is committed to defending, the challenger is more likely to be restrained in its escalation behavior, consistent with the argument that such an ally is likely both a great power with considerable influence and also interested in discouraging reckless behavior by its many allies that could lead to entrapment. We also see in Figure 4 that a defender with similar allies that have many allies is much more likely to respond with non-violence when there is a non-violent trigger. Having strong allies which, in turn, have many commitments appears to be associated with reduced propensity to make disproportionate responses and counter-responses.

A number of crises are characterized by non-violence in triggers, responses, and counter-responses, where belligerency costs in the form of high human rights values for the challenger are present. These cases include the Cod Wars of 1973 (Iceland challenger) and 1975 (UK challenger), Greece and Turkey in the Aegean Sea IV crisis of 1996 (Greece challenger), and the Venezuelan Election crisis of 201s9 (United State challenger). However, two prominent cases run counter to this trend and constitute failed attempts at deterrence. In the Gulf of Syrte crisis of 1981, a dispute over Libyan claims of sovereignty resulted in movement of the Sixth Fleet into the gulf on 12 August to hold maneuvers, triggering a crisis for Libya. Libya responded with a full military alert, and the United States then responded on the 13th by shooting down two Libyan fighters over the gulf. In the Iraq Regime Change crisis of 2002, on 12 September US President Bush announced the conditions Iraq must meet in destroying its nuclear weapons program. Iraq responded by allowing inspectors into the country, resulting in a report indicating that Iraq was not in full compliance with the conditions, and on 20 March the United States launched an invasion with the goal of disarming the country of its weapons of mass destruction.

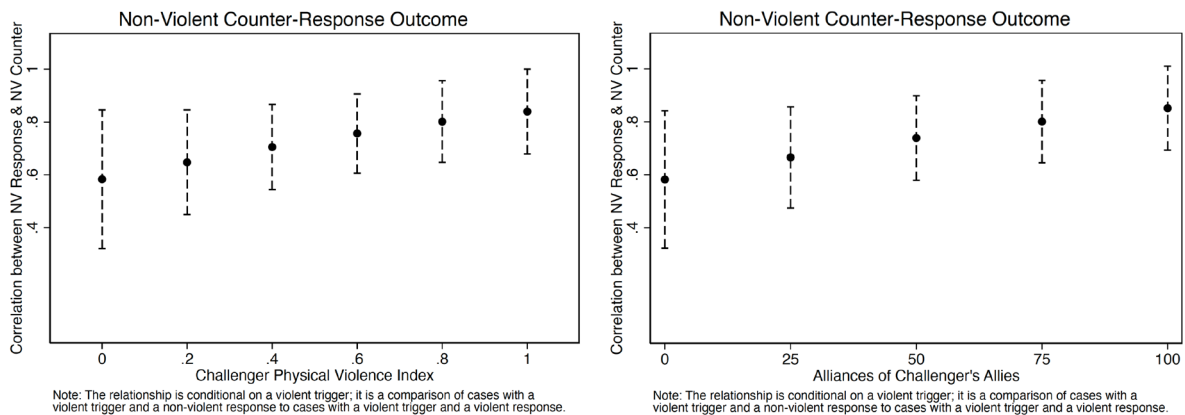


Figure 3: Conditioning effects of challenger human rights performance and challenger allies' alliances

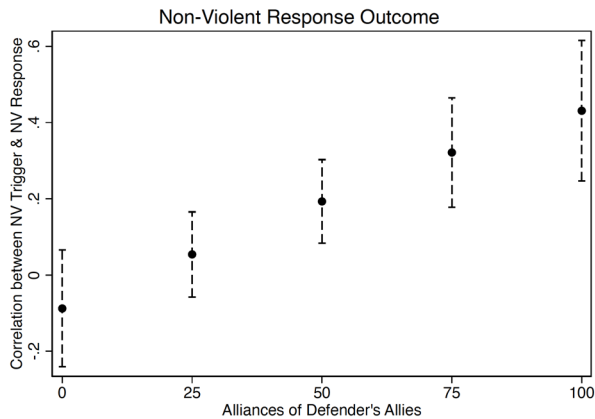


Figure 4: Conditioning effect of defender allies' alliances

Speculative Finding: Actors are more likely to meet violent triggers and responses with violence when the stakes are low.

Interestingly, for H1c and H2c, the findings are opposite to what was hypothesized. As the top two panels in Figure 5 show, when the challenger faces a grave threat and has experienced a non-violent trigger, the defender is *less* likely to respond non-violently. In the same vein, the challenger is also less likely to make a non-violent counter-response when the defender has responded non-violently.

We propose an intuitive explanation for the opposite results. The theory had focused on how a non-violent trigger and response would be interpreted when the stakes are known to be high. But an even stronger mechanism could be in play if we consider how violent triggers and responses are interpreted when the stakes are low. Also consistent with the results is that violent triggers and violent responses are more likely to be met with violence when the stakes are low, as seen in the bottom panels of Figure 5. When the stakes are lower, actors are more willing to reciprocate, perhaps because there is less potential for the violence to extend beyond limited exchanges. The logic here is similar to the “stability-instability paradox” (Snyder & Diesing, 1977) in that stability at higher levels of escalation allows for more aggressive competition at lower levels. But the context driving the aggression at lower levels here is precisely the opposite of the nuclear context—when existential threat takes high escalation off the table—that is typically the locus of the stability-instability paradox. In this case, it is *low* stakes that make it inconceivable for total war to occur and that allow for them to act more aggressively early in a crisis. One implication is that escalate-to-de-escalate strategies appear less likely to work when the stakes are low.

Under conditions in which an opponent faces high stakes, as measured by high threat severity, defenders and challengers are more reluctant to meet violence with violence. That is, there is likely a deterrent effect that leads to less reciprocity: Defenders are less likely to meet a violent

trigger with a violent response, and challengers are less likely to meet a violent response with a violent counter-response.

There were 36 crises between 1963 and 2019 that were typified by low threat perception on the part of the actors involved, but nevertheless showed violence in the trigger, in the response, and in the counter-response. The most recent of these cases were Yemen in the Red Sea Crisis 1995, Côte d'Ivoire in its Presidential Election Crisis 2011, Sudan in the Sudan-South Sudan Crisis 2011, Turkey in the Syria-Turkey Border Incidents 2012, India and Pakistan in India-Pakistan Border Firing 2014, Armenia in the Nagorny-Karabakh April War 2016, and Afghanistan and Pakistan in the Torkham Border Incident 2016. Many of these examples involve territorial disputes in which potential violence is expected to remain confined to border clashes and not extend to the respective capitals.

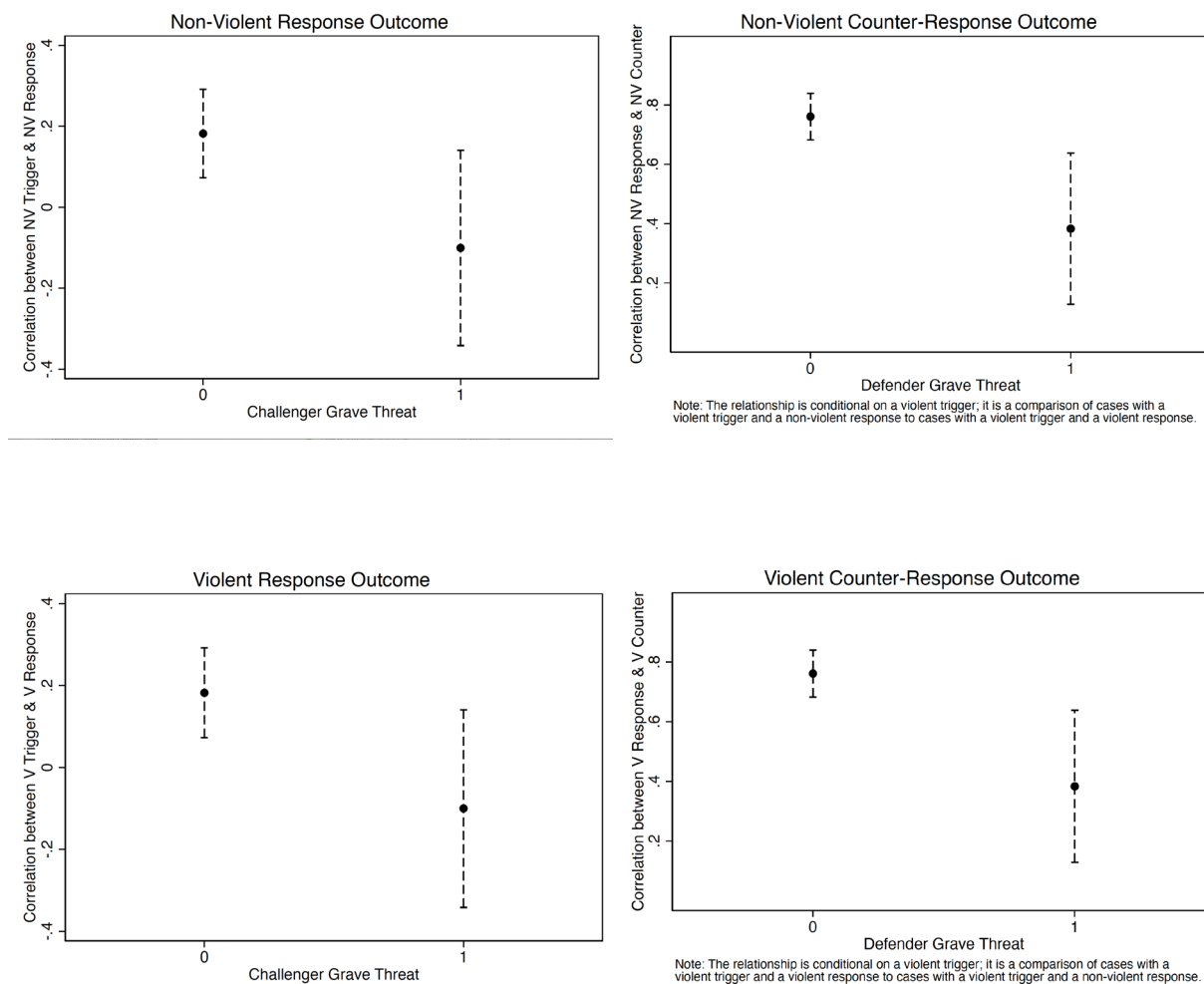


Figure 5: Conditioning effect of challenger and defender severity of threat

Conclusion

This report has focused on the examination of some of the basic premises of deterrence theory, by analyzing data on all international crises from 1963 to 2019. During this 57-year period, there were 300 crises involving a myriad of state adversaries, ranging from minor powers involving regional and local interests to major and superpowers locked in competition over international dominance across vast regions of the globe. While no two crises are the same, taken in aggregate, a better understanding of the behavior of the leaders of these states and the choices they made when confronted with crisis conditions, can help us identify general trends in their behavior. Equipped with the findings generated by this vast dataset, we offer a set of general actionable principles as a summary of the statistical analyses reported above.

1st Deterrence Principle—Exploit your Observable Strengths

Non-violent triggers are more likely to lead to non-violent responses when the challenger has high levels of observed capabilities and resolve. Here we identify factors that will assist a challenger, when contemplating undertaking a non-violent action that is nevertheless likely to be perceived by the defender as sufficiently threatening to trigger a crisis, in evaluating the probability of the defender undertaking a non-violent response. The critical factor here is how capabilities and resolve are conveyed accurately to the defender. An inability to accurately project strong levels of capabilities and resolve could result in a defender's perception of weakness, and therefore a loss of the deterrent effect.

2nd Deterrence Principle – Exploit your Adversary's Weakness

An actor is more likely to meet non-violent triggers and responses with non-violence when it is prone to suffering belligerence costs. Thus, if account is taken of the difficulties facing an adversary due to their publics' or allies' perception of significant costs that would be associated with belligerence, then non-violent actions are likely to be matched by non-violence, rather than with escalation to violence. Some adversaries will have higher potential for "self-deterrence"—they will be reluctant to respond strongly when an actor uses non-violent tactics, and the adversary has domestic and international audiences that are sensitive to disproportionate violence. This dynamic is particularly the case when the adversary has a strong record on human rights, as a manifestation of potentially significant domestic costs of belligerence. An additional related factor is the number of alliance commitments of the allies of the challenger or defender. With high numbers of alliances comes a high sensitivity to entrapment. Consequently, adversaries that are supported by allies who, in turn, have many allies will be under close watch to not behave recklessly.

3rd Deterrence Principle—Escalate-to-De-escalate Strategies are Unlikely to Work when the Stakes are Low

Actors are more likely to meet violent triggers and responses with violence when the stakes are low. Adversaries that know they can use violence without it catalyzing a major war will indeed be more willing to use violence. As a result, uses of violence by an actor that is known to have relatively low stakes—specifically actors that do not perceive a threat of grave damage—are unlikely to deter an adversary from an aggressive response. In contrast, escalate-to-de-escalate strategies—when an actor uses violence in hopes of jolting its adversary into a risk-reduction posture—are more prone to work when it is known that the actor perceives that there is a lot on the line.

The important takeaway here, and implicit throughout our analysis in this report, is that factors seemingly extraneous to the observable dynamics of a crisis among adversaries must be considered when formulating deterrence strategy. Elder and Astorino-Courtois (2023, pp. 263-266) have developed a Challenger-Response Escalation Impact table to help identify the conditions under which there will be mismatches between triggers and responses, and between responses and counter-responses. Decision makers who ignore the level of threat, and the accompanying heightened stress levels experienced by their adversaries, will place their own countries at peril.

References

- Amemiya, T. (1978). The estimation of a simultaneous equation generalized probit model. *Econometrica: Journal of the Econometric Society*, 1193–1205. <https://doi.org/10.2307/1911443>
- Barnum, M., Fariss, C.J., Markowitz, J., & Morales, G.P. (2022). *Measuring arms: Introducing the global military spending dataset* [Data Set]. Open Science Foundation. <https://osf.io/rsxze/wiki/home/>
- Benson, M., & Kugler, J. (1998). Power parity, democracy, and the severity of internal violence. *Journal of Conflict Resolution*, 42(2), 196–209. <http://www.jstor.org/stable/174569>
- Brecher, M. (1993). *Crises in world politics*. Pergamon Press.
- Brecher, M., Wilkenfeld, J., Beardsley, K., James, P., & Quinn, D. (2023). *International crisis behavior data codebook* (Version 15) [Data Set]. <http://sites.duke.edu/icbdata/data-collections/>
- Caprioli, M., & Trumbore, P.F. (2006). Human rights rogues in interstate disputes, 1980–2001. *Journal of Peace Research*, 43(2), 131-148. <https://doi.org/10.1177/0022343306061356>

-
- Cebul, M.D., Dafoe, A., & Monteiro, N.P. (2021). Coercion and the credibility of assurances. *The Journal of Politics*, 83(3), 975-991. <https://doi.org/10.1086/711113>
- Cohen, D.K., & Karim, S.M. (2022). Does more equality for women mean less war? Rethinking sex and gender inequality and political violence. *International Organization*, 76(2), 414-444. <https://doi.org/10.1017/S0020818321000333>
- Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Jan Teorell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Lisa Gastaldi, Haakon Gjerløw, Adam Glynn, Ana Good God, Sandra Grahn, Allen Hicken, Katrin Kinzelbach, Joshua Krusell, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Juraj Medzihorsky, Natalia Natsika, Anja Neundorff, Pamela Paxton, Daniel Pemstein, Josefine Pernes, Oskar Rydén, Johannes von Römer, Brigitte Seim, Rachel Sigman, Svend-Erik Skaaning, Jeffrey Staton, Aksel Sundström, Eitan Tzelgov, Yi-ting Wang, Tore Wig, Steven Wilson and Daniel Ziblatt. (2023). *V-Dem [Country-Year/Country-Date] Dataset (Version 13) [Data Set]*. Varieties of Democracy (V-Dem) Project. <https://doi.org/10.23696/vdemds23>
- Cranmer, S.J., & Desmarais, B.A. (2016). A critique of dyadic design. *International Studies Quarterly*, 60(2), 355-362. <http://www.jstor.org/stable/43869078>
- Crescenzi, M.J.C. (2018). *Of friends and foes: Reputation and learning in international politics*. Oxford University Press.
- Dafoe, A., Hatz, S., & Zhang, B. (2021). Coercion and provocation. *Journal of Conflict Resolution*, 65(2-3), 372-402. <https://doi.org/10.1177/0022002720957078>
- Debs, A., & Weiss, J.C. (2016). Circumstances, domestic audiences, and reputational incentives in international crisis bargaining. *Journal of Conflict Resolution*, 60(3), 403-433. <https://doi.org/10.1177/0022002714542874>
- Di Lonardo, L., & Tyson, S.A. (2022). Political instability and the failure of deterrence. *The Journal of Politics*, 84(1), 180-193. <https://doi.org/10.1086/715258>
- Dixon, W. J. (1994). Democracy and the peaceful settlement of international conflict. *American Political Science Review*, 88(1), 14-32. <https://doi.org/10.2307/2944879>
- Dudley, R. (2023). Turning the tables: Direct military intervention and the onset of negotiations in civil war. *Journal of Conflict Resolution*, 0(0). <https://doi.org/10.1177/00220027231190915>
- Elder, R. J. and Astorino-Courtois, A. 2023. "Toolkit for planners and operators in gray zone crisis environments," in J. Wilkenfeld and E. Murauskaite (eds.), *Escalation Management in*

International Crises: The United States Confronts its Adversaries. Cheltenham, UK: Edward Elgar Publishers.

Fearon, J.D. (1994). Signaling versus the balance of power and interests: An empirical test of a crisis bargaining model. *Journal of Conflict Resolution*, 38(2), 236-269. <http://www.istor.org/stable/174295>

Fearon, J.D. (1997). Signaling foreign policy interests: Tying hands versus sinking costs. *Journal of Conflict Resolution*, 41(1), 68-90. <https://doi.org/10.1177/0022002797041001004>

Gartzke, E.A., & Gleditsch, K.S. (2022). Ties that bias in international conflict: A spatial approach to dyadic dependence from alliance ties and inbetweenness. *International studies quarterly*, 66(1), sqab182. <https://doi.org/10.1093/isq/sqab082>

George, A. (1992). *Forceful persuasion: Coercive diplomacy as an alternative to war*. U.S. Institute of Peace.

Gurantz, R., & Hirsch, A.V. (2017). Fear, appeasement, and the effectiveness of deterrence. *The Journal of Politics*, 79(3), 1041-1056. <https://doi.org/10.1086/691054>

Hegre, H. (2014). Democracy and armed conflict. *Journal of Peace Research*, 51(2), 159–172. <https://doi.org/10.1177/0022343313512852>

Holsti, O.R. (1972). *Crisis, escalation, war*. McGill-Queen's Press-MQUP.

Hudson, V., Bowen, D.L., & Nielsen, P.L. (2020). *The first political order: How sex shapes governance and national security worldwide*. Columbia University Press.

Huth, P., & Russett, B. (2013). General deterrence between enduring rivals: Testing three competing models. *American Political Science Review*, 87(1), 61-73. <https://doi.org/10.2307/2938956>

Jervis, R. (1976). *Perception and misperception in international politics: New edition*. Princeton University Press.

Johnson, J.C., & Joiner, S. (2021). Power changes, alliance credibility, and extended deterrence. *Conflict Management and Peace Science*, 38(2), 178-199. <https://doi.org/10.1177/0738894218824735>

Johnson, J.C., Leeds, B.A., & Wu, A. (2015). Capability, credibility, and extended general deterrence. *International Interactions*, 41(2), 309-336. <https://doi.org/10.1080/03050629.2015.982115>

-
- Kang, C.-N. (2017). Capability revisited: Ally's capability and dispute initiation. *Conflict Management and Peace Science*, 34(5), 546-571.
- Kertzer, J.D., & Brutger, R. (2016). Decomposing audience costs: Bringing the audience back into audience cost theory. *American Journal of Political Science*, 60(1), 234-249.
- Krainin, Colin, & Schub, R. (2021). Alliance dynamics in the shadow of shifting power. *International Studies Quarterly*, 65(4), 905-918.
- Kydd, A.H. (2007). *Trust and mistrust in international relations*. Princeton University Press.
- Leeds, B.A. (2003). Do alliances deter aggression? The influence of military alliances on the initiation of militarized interstate disputes. *American Journal of Political Science*, 47(3), 427-439.
- Leeds, B.A., Long, A.G., & Mitchell, S.M. (2000). Reevaluating alliance reliability: Specific threats, specific promises. *Journal of Conflict Resolution*, 44(5), 686-699.
- Leeds, B.A., & Mattes, M. (2022). *Domestic interests, democracy, and foreign policy change*. Cambridge University Press.
- Leeds, B.A., Ritter, J.M., Mitchell, S.M., & Long, A.G. (2002). Alliance treaty obligations and provisions, 1815-1944. *International Interactions*, <https://www.tandfonline.com/doi/abs/10.1080/03050620213653>
- Levin-Banchik, L. (2021). Precrisis military hostility and escalation in international crises. *Conflict Management and Peace Science* 38(1), 63-86. <https://doi.org/10.1177/0738894220906376>
- Maoz, Z., & Felsenthal, D.S. (1987). Self-binding commitments, the inducement of trust, social choice, and the theory of international cooperation. *International Studies Quarterly*, 31(2), 177-200. <https://doi.org/10.2307/2600452>
- Mastro, O.S. (2019). *The costs of conversation: Obstacles to peace talks in wartime*. Cornell University Press.
- Mastro, O.S., & Siegel, D.S. (2023). Talking to the enemy: Explaining the emergence of peace talks in interstate war. *Journal of Theoretical Politics* 3(5), 182-203. <https://doi.org/10.1177/09516298231185112>
- McManus, R.W. (2014). Fighting words: The effectiveness of statements of resolve in international conflict. *Journal of Peace Research* 51(6), 726-740. <https://doi.org/10.1177/0022343314539826>

-
- Minhas, S., Dorff, C., Gallop, M.B., Foster, M., Liu, H., Tellez, J., & Ward, M.D. (2022). Taking dyads seriously. *Political Science Research and Methods* 10(4), 703-721.
- Morrow, J.D. (1991). Alliances and asymmetry: An alternative to the capability aggregation model of alliances. *American Journal of Political Science*, 35(4), 904–933. <https://doi.org/10.2307/2111499>
- Quek, K. (2021). Four costly signaling mechanisms. *American Political Science Review*, 115(2), 537-549. <https://doi.org/10.1017/S0003055420001094>
- Renshon, J., Yarhi-Milo, K., & Kertzer, J.D. (2023). Democratic reputations in crises and war. *The Journal of Politics*, 85(1), 1-18. <https://doi.org/10.1086/722043>
- Sartori, A.E. (2013). *Deterrence by diplomacy*. Princeton University Press.
- Schelling, T.C. (1980). *The strategy of conflict: With a new preface by the author*. Harvard University Press.
- Sechser, T.S. (2016). Reputations and signaling in coercive bargaining. *Journal of Conflict Resolution*, 62(2), 318-345. <https://doi.org/10.1177/0022002716652687>
- Sjoberg, L. (2013). *Gendering global conflict: Toward a feminist theory of war*. Columbia University Press.
- Slantchev, B.L. (2003). The principle of convergence in wartime negotiations. *American Political Science Review*, 97(4), 621-632. <https://doi.org/10.1017/S0003055403000911>
- Smith, A. (1995). Alliance formation and war. *International Studies Quarterly*, 39(4), 405–425. <https://doi.org/10.2307/2600800>
- Snyder, G.H. (1984). The security dilemma in alliance politics. *World politics*, 36(4), 461-495. <https://doi.org/10.2307/2010183>
- Snyder, G. H. (1991). Alliances, balance, and stability. *International Organization*, 45(1), 121–142. <https://doi.org/10.1017/S0020818300001417>
- Snyder, G.H., & Diesing, P. (1977). *Conflict among nations: Bargaining, decision making, and system structure in international crises*. Princeton University Press.
- Sobek, D., Abouharb, M.R., & Ingram, C.G. (2006). The human rights peace: How the respect for human rights at home leads to peace abroad. *The Journal of Politics*, 68(3), 519-529. <https://doi.org/10.1111/j.1468-2508.2006.00442.x>

Spaniel, W., & İdrisoğlu, I. (2023). Endogenous military strategy and crisis bargaining. *Conflict Management and Peace Science*, 0(0). <https://doi.org/10.1177/07388942231180499>

Weeks, J.L. (2012). Strongmen and straw men: Authoritarian regimes and the initiation of international conflict. *American Political Science Review*, 106(2), 326-347. <https://dx.doi.org/10.2139/ssrn.1748516>

Wolford, S. (2023). Great power intervention and war. *International Studies Quarterly*, 67(1). <https://doi.org/10.1093/isq/sqad003>

Zagare, F.C., & Kilgour, D.M. (1993). Asymmetric deterrence. *International Studies Quarterly*, 37(1), 1-27. <https://doi.org/10.2307/2600829>

Zagare, F.C., & Kilgour, D.M. (2000). *Perfect deterrence*. Vol. 72. Cambridge University Press.

