JOURNAL OF CYBERSECURITY

Journal of Cybersecurity, 2021, 1–12
doi: 10.1093/cybsec/tyab001
Research paper

Research paper

Risk and uncertainty can be analyzed in cyberspace

Aaron F. Brantly *

Department of Political Science, Virginia Tech, Blacksburg, Virginia, USA

*Correspondence address. Department of Political Science, Virginia Tech, Blacksburg, Virginia, USA. Tel: +20-27-25-34-02; E-mail: abrantly@vt.edu,

Received 6 February 2020; revised 29 December 2020; accepted 21 January 2021

Abstract

Perceptions of risk and uncertainty are pervasive in all international interactions. How states perceive risk and uncertainty and how they respond to these conditions impacts their policies and diplomatic behaviors. Despite a robust literature encompassing of risk and uncertainty within conventional state to state interactions including conflict, state interactions in cyberspace have received less attention. How states perceive and interpret risk and uncertainty in cyberspace varies widely by state. Very often, these perceptions are mutually incompatible and lead to a sub-optimal status quo that fosters increased risk and uncertainty. While the prospects of uncontrolled escalation or worries about a "Cyber Pearl Harbor" might be hyperbole, the reality remains that for decision-makers within states assessing the conditions of and the actions undertaken in cyberspace at present foster instability and encourages risk-seeking behaviors. This work analyzes the formulation of state perceptions of risk and uncertainty and seeks to establish a heuristic within which risk and uncertainty can be analyzed.

Key words: cybersecurity, risk, uncertainty

Perceptions of risk and uncertainty form part of the framework for decisions on state policy development and actions. How states perceive one another and the actions they undertake in any domain whether conventional or digital is important for understanding and predicting their future actions. How states formulate perceptions of risk and uncertainty within international interactions has been examined previously by multiple scholars across the conflict studies literature [1–3], yet the overwhelming majority of prior analyses have privileged conventional state-to-state interactions transpiring outside of cyberspace. Whether conventional decision-models pertaining to conventional conflict hold within cyberspace has received less attention. This work examines the formulation of state perceptions of risk and uncertainty in cyberspace. Ultimately, this article finds that states are deciding to engage in different behaviors as a result of divergent interpretations of shared experiences.

Perceptions of risk and uncertainty in cyberspace differ in their construction than in other domains for a number of reasons. Among these differences are anonymity, order of effects, complexity, uncertainty of adversary capabilities, and field specific jargon and technical concepts. The ability to remain anonymous over critical periods of time shields both the violator and the violated from crucial information flows that inform and help to clarify decisions [4]. Likewise, actions and their effects in cyberspace are not straightforward and can be obfuscated as states do not seek to cause first-order effects, but rather effects that extend from second- and third-order consequences [5]. These effects foster confusion related to action and effects pairs as the complexity of networked environments makes the unintended spread of cyber capabilities beyond intended targets far more possible than the comparative use of conventional capabilities with the exception of biological weapons and perhaps nuclear fallout (both of which far exceed the presently demonstrated destructive power of cyber capabilities). Moreover, just as uncertainty about conventional weapons changes perceptions of risk, uncertainty surrounding adversary cyber capabilities elevates the perceptions of risk associated with adversary cyber capabilities. Finally, as is often the case with new technologies in conflict—concepts of risk and uncertainty associated with cyber capabilities befuddle decision-makers with an overwhelming amount of jargon and complex concepts.

Placing the risk and uncertainty associated with state interactions in cyberspace within the broader context of international relations and security studies research makes it possible to examine the question of whether risk and uncertainty can be examined in cyberspace. At its most basic the article asks how states perceive risk and uncertainty in cyberspace. Assessing how states perceive risk and uncertainty increases the accuracy of decision-modeling for states acting in and through cyberspace. This article proceeds in four sections. It begins by briefly examining the literature on uncertainty and risk to contextualize state behaviors prior to examining their actions relating to cyberspace. After examining the literature on uncertainty and risk, the attributes of cyberspace are examined and contrasted with attributes often found in more conventional conflict dynamics. Next, a brief discussion of two cases are used to illustrate the formulation of risk by both the Russian Federation and the USA. Lastly, state learning in cyberspace is examined.

From unquantifiable to quantifiable uncertainties

To understand why states behave differently from one another in cyberspace requires briefly examining how decisions are made within a rational choice framework and beyond. Specifically, this work extends beyond a simplified rational choice framework predicated on five core axioms: completeness, transitivity, continuity, monotonicity, and substitution [6]. This extension is necessary because states are learning how to "bound" decisions pertaining to actions in cyberspace. It is the act of bounding decisions for future rational choice frameworks that differentiates the decisions in cyberspace from those in conventional interactions. Bounding is the result of a learning process, a heuristic. Von Neuman and Morgenstern identify that the "bounds" embedded within decisions leave considerable room for the interpretation of the games being played [6].

As noted by Jack Levy: "The concept of learning is difficult to define, isolate, measure, and apply empirically. . ." [7].[1] For the purposes of this article learning is an iterative process of knowledge acquisition and shapes the bounds within which decisions under conditions of uncertainty are undertaken. The development of an expected utility for international conflict, whether conventional or transpiring within cyberspace requires parameterization. Parametrization converts Knightian unquantifiable uncertainty to quantifiable decisions with uncertain outcomes. Several millennia of experiences have helped shape the subjective bounds of expected utilities for conventional conflict. Below I lay out two layers of the decision-model starting with that which is unquantifiable and moving into that, which is.

Knightian uncertainty

Knightian uncertainty constitutes the nonquantifiable conditions within which risky (quantifiable) decisions are made. Forming the foundation of nearly all decision models containing calculi on risk is the notion that the underlying conditions are quantifiable. Rational choice and most bounded cognition models are predicated on the quantification of decision. As a result, most formal analyses avoid the challenge of Knightian uncertainty, or unpredictable events that are beyond quantification [8]. Most analyses of decisions in international relations are bounded either with a priori assumptions based on experience or within the limits of available time or resources to seek out new information [9].

Experience reduces Knightian uncertainty through *a priori* information that can be ascertained in multiple ways. Experience can be constitutive of biological/genetic characteristics rooted within the behavioral patterns of living beings. Experience can be abstractly learned via historical records, cultural or social norms. It can be lived, such as learning that touching a hot stove results in pain or witnessing others experience something. Experience is endemic to the human existence and by its nature reduces Knightian uncertainty. Most expected utility models build on experience in some fashion to "bound" and thereby quantify risk.

The assumption of boundedness pervades both cognitive and rational models. The tighter the presumed knowledge boundaries of cognitive or rational frameworks, the greater the Knightian uncertainty the less accurately decisions can be quantified. As risk assessments become more parsimonious, they are increasingly divorced from reality. Parsimony, in and of itself is not necessarily a bad thing. The presumption of simplicity facilitates decisions [10]. There is, however, a tradeoff between parsimony and reality. As parsimony simplifies the decision process, events that fall beyond those predictable within cognitive or rational models increasingly arise. When an outcome occurs that falls outside of the bounds that experience would quantify as possible, the events is best referred to as a Black Swan [11].

At its most basic, risk is the probability and impact of losing something, a tradeoff between choices resulting in a potential loss. Or as defined by Huth *et al.*: "The concept of risk propensity is most clearly revealed by comparing patterns of individual choice between options that have similar expected values but vary in their probabilities"[3]. The overwhelming majority of analyses of decision-making in international relations examine state actions in the context of risk [2, 12–15]. The exact outcome is still uncertain, but it is quantifiable. Of considerable importance to the study of international relations and predictions related to state behavior is the distinction between cognitive and rational approaches to decision-making [1, 16, 17]. However, both cognitive and rational models seek to understand or provide predictions based on state perceptions of risk [18]. While cognitive models lean on a variety of analyses, several have emerged as dominant explanations, in particular, prospect theory [19]. Rational modeling of states can be divided along utility modeling and more generalized rational models [17, 20].

The more bounded a state is in its information environment, the less likely its perceptions of risk or its assessments of utility are to match reality.[2] Larger volumes of information increase modeling fidelity and reduce uncertainty by adding to the decision matrix qualities which are quantifiable. In contrast, the less information a state has relative to the environment (international system) or domain in which it operates, the more uncertainty will be present. Similarly, if too much information is available, and the boundaries of information within a decision-model are too large and the time to process and generate probabilities or utilities for various decisions is constrained temporally, decisions can be plagued by what is known as polythink [21]. Both too little and too much information can result in inaccurate utility constructions, thereby increasing uncertainty.

1 For a full discussion of learning in international relations see: Levy JS. Learning and foreign policy: sweeping a conceptual minefield. *Int Organ* 1994;**48**:279–312.

2 Bounding and the reduction of fidelity to reality are particularly acute in nonfinite environments. As the number of variables and their interactions increase, predicting outcomes based on highly constrained deterministic models can lead to increasingly inaccurate outcomes. For an example of this, see: Lorenz Edward, Deterministic nonperiodic flow. *J Atmospheric Sci* 1963:**20**.

The latter's inability to foster decisions is a fundamentally different problem than the former and often results from a lack of processing power and/or the ability to fully comprehend or assess complex probabilities or utilities. Often the latter case is overcome through sustained analysis, but as Kissinger writes, the diplomat has substantially less time than the analyst [22].

Yet, to get to the utility model itself, states must make assumptions based on the experience. To do this, they must move beyond Knightian uncertainty to quantifiable uncertainties. Here in lies the core issue and challenge presented in this article. To get to quantifiable states, i.e. the utility model, states must bound their decisions, this bounding is accomplished principally through lived or learned experience. Absent of a framework, within which to develop quantifiable risks, the uncertainty of states increases. Given a tabula rasa, states are unable to formulate an expected utility model due to the unquantifiable conditions of Knightian uncertainty. As they engage in activities within cyberspace, they begin to establish the contours of the domain and generate the experiences[3] needed for quantification and thereby utility modeling [10].

Measurements of state behavior over time, whether the detailed analysis of diplomacy [22], the quantification of wars and the relative power of the states that wage them [23], or analyses that parse out how different options are formulated within a government [24], these and many more seek to divine the means by which states identify and develop the information associated with any given decision. Jervis's analysis of perception and misperception highlights problems within historical analyses as well as the challenges in the quantification of previous events, as he correctly indicates, very often the wrong interpretation of events is dependent on the state, or individuals involved [25]. While perception and misperception are rampant within international politics, each knowledge point acquired potentially further adds to the construction of utility models associated with risk and reward for various decisions.

War, as so aptly described by James Fearon, is the ultimate model within which states identify and assess the information they had antecedent to a given conflict [20]. As a war transpires, the bounds of information tend to be shed and each side is more accurately able to assess the probability of outcomes based on decisions and action. If states had and comprehended perfect information on the probable outcomes of any given conflict in advance, they might have been able to forego war through negotiated settlements. However, as is poignantly examined by Erik Gartzke, War is the "error term" [26]. Specifically, Gartzke notes: "We cannot predict in individual cases whether states will go to war, because war is typically the consequence of variables that are unobservable ex ante, both to us as researchers and to the participants" [26]. These unobservable variables constitute the Knightian uncertainty that underpins decisions and falls outside the scope of quantifiable risk. Yet within conventional conflict, despite Gartzke's notion, states do attempt to predict the likelihood of war by bounding their decision models based on experience. Yet, even with abundant historical and lived experiences, the resultant war or conflict often is perceived to be a "black swan" event.

Moreover, despite Gartzke's analysis, prior to conventional conflict the information available to most states is substantial. Strategic intelligence seeks to provide as much information as possible regarding the resource capabilities and the decision-making logic of adversaries [27]. National Security intelligence collection and analysis serves to reduce Knightian uncertainty and more accurately bound utility models. While much of the intelligence derived is delivered in the form of assessments, often probabilistic in nature, these products improve the quantification of risks and thereby reduce the uncertainty of outcomes.

Even with the abundance of intelligence, the formulation of accurate assessments in international relations is consistently plagued by failures of one kind or another [28, 29]. This is despite efforts by decision-makers to utilize information to make optimal decisions. Mistakes are not necessarily caused by information provision errors that lead to faulty decision-making, but rather issues outside of the information used in assessing the risks and rewards associated with any decision related to a given action. Thus, even if a state knows how many tanks and soldiers another state has, and even if it knows that the adversary state is willing to fight to the last man or woman, there are factors that extend beyond an analysis of risks and rewards based on simple probabilities of success or failure. These intangibles challenge even the most rigorous of analyses on decision-making in international relations [30]. Intangible attributes within the broader decision-matrix makes the analysis of perceptions of risk at the state level unstable.

In formulating choices, through both the framing of the information being signaled and received, and the information (intelligence) available to the state it is apparent that at the state level, leaders are privy to a large amount of information. While this information may be insufficient to eliminate all error, it is hopefully enough to minimize egregious errors. The attributes of conventional conflict are multifaceted. The generic features of states in advance of conventional conflict range from weapon systems and troop numbers to their economic and political attributes. While states try to deliberately obfuscate, minimize, or elevate certain attributes, the ability to hide many of them in totality is difficult. Despite the volume of readily available information in advance of conventional conflict wars still take place. States still perceive the risks are worth taking and that the utility being sought is positive [17].

To summarize the current state of affairs regarding state perceptions of risk and uncertainty, within international relations, there are two principal decision-making frameworks rational choice and cognitive models which lead the way in providing readily accessible and mathematically rigorous means to predict state decisions. While their predictive qualities vary over time and scope, both are generally well-regarded means of assessing when and how states will decide between difficult choices. Yet, as the literature and history demonstrate both are susceptible to substantial error in predictive quality. Both are also only as strong as the known unknowns and known knowns upon which they are based (those things which are quantifiable). As the level of uncertainty increases, the quantifiable nature of the world in which events are transpiring results in less accurate predictions. When there are unknown, unknowns states are confronted with uncertainty categorized as black swans or unknown/unforeseen events.

Examining the attributes of risk and uncertainty in cyberspace

The short examination of a broad swath of literature above contextualizes the pre-cyber world. A world in which the quantification of risk and the comprehension of uncertainty provide decision-makers a framework within which to assess outcomes. Whether these assessments are accurate or not is the matter of discussion. Cyberspace

3 No qualitative or empirical position is taken here regarding whether learning occurs best from positive or negative outcomes. Both are likely, but with different effects, rather the intent is to highlight the establishment of reference points.

challenges the conventional formulations of state behavior for a number of reasons. Unlike in the conventional domains of land, sea, air, and space, assessing the relative power of states is highly subjective and subject to changes that have less to do with material or financial capabilities than educational, infrastructural, or other attributes often ignored in conventional analyses of power. The term domain itself implies the importance and frames the metaphorical space in which a new typology of activities involving code and internetworked computers that facilitate military and civilian activities take place [31]. Defining the space (cyberspace) of operation is useful insofar as it helps delineate and differentiate it from activities occurring in other domains. Several scholars have attempted to develop parallel understandings of state power in cyberspace as a means to extend analysis beyond conventional capabilities [4, 32, 33]. There is a great deal of skepticism about the construction of analyses on state power in cyberspace [34] as it comprised a large number of intangibles not necessarily visible to outside observers. Some scholars even contend that the risk of conflict within or emanating from cyberspace is minimal or will simply not take place [35].

At the outset, it should be stated that all concepts of risk and utility associated with conventional conflict are applicable to cyber conflict. Evidence from repeated cyber-attacks [36], espionage campaigns [37], subterfuge [38], and more [39] indicate a willingness and desire on the part of states to engage in conflict within cyberspace. Whether this conflict is of significance or not to the core interests of states and whether it changes state behavior remains a topic of considerable debate [40]. Many states do consider cyber conflict to be of significance and have created national strategies to reflect this reality [41]. Yet, despite many of the similarities to conventional conflict in terms of the creation of national strategies and rhetoric in the literature [42], the reality is that cyber conflict differs substantively from conventional conflict across many dimensions that impact the assessment of risk and utility functions for both conventional rational and cognitive models [43, 44].

Many of the attributes of conventional conflict that facilitate the creation and framing of probabilities for risk are difficult to discern in cyberspace. While the assessment of passive and active weapon systems can and does pose problems in conventional conflict [45], the subsequent impact of such weapon systems is often clearer than equivalent distinction between capabilities in cyberspace [46]. Gartzke and Lindsay [47] note that differentiating the values of offensive and defensive capabilities in cyberspace is difficult and has led to a presumption that offense is dominant. As they note, this might not be the case due to the inherently deceptive nature of interactions in cyberspace that are themselves elevating perceptions of risk and thereby fostering strategic interactions that might not otherwise occur. They are supported in their assessment and perceptions by Ben Buchanan who argues that the nature of the capabilities themselves fosters uncertainty as to adversary capabilities and intentions [48].

Because the code used to exploit and steal information from a system can closely resemble that used to disrupt, degrade, or deny access, the ability to generate a risk assessment tends to err on the side risk aversion[4] for the defender and risk acceptance[5] for the attacker [48–50]. The defender assumes the worst while the attacker assumes better. It is not that an attacker assumes that their espionage, subterfuge, cyber-attack, or other activity transpiring within cyberspace is somehow without risk, but rather that they (the attacker) assume that the signaled intent of the activity is more readily obvious than it often is [51–54]. Signaling in cyberspace is particularly difficult as many of the activities undertaken to signal an adversary in cyberspace occur clandestinely or covertly [4]. The attacker is not overly optimistic, nor is the defender overly pessimistic. Both are operating under the bounds of information available to them. Whereas the attacker knows the intent of the malicious code and presumably its limits, the defender does not know the intent or the limits until substantial forensic analysis has been completed [4]. These are oversimplifications, but generally, they align with the behavior of states and actors as well as the rhetoric of journalists and others who cover cyber incidents [55]. Perceptions of risk in cyberspace as in conventional domains are shaped by facts interpreted via a variety of mediums.

Rose McDermott [56] in writing on risk taking in international politics notes that people are susceptible to judgment errors often derived from three basic heuristic biases: representativeness, availability, and anchoring. The first judgment error, representativeness, refers to the proper categorization of an event or object. Whereas the implications for an object such as a bomb, a new fighter aircraft, or another conventional weapon system are often fairly evident, there remains a sufficient amount of ambiguity to lessen the probability of accurate categorization. Examples of failures of accurate object or event categorization are common (happen regularly, although not necessarily at statistically significant levels), and examples abound in which US intelligence agencies have misinterpreted various weapons platforms or events based on mirror-imaging or other biases unknowingly introduced [57]. Despite a wealth of information available on the object, often obtained through intelligence collection methods such as imagery analysis (IMINT), objects are frequently miscategorized as was the case of the Soviet Backfire Bomber (Tuplolev Tu-22m) which was initially categorized as a long-range bomber only to be later reclassified by US intelligence as an "over-engineered peripheral bomber" [57]. Representativeness is informed by information or intelligence about an object or event. The less information available, the more likely miscategorization is to occur.

Time and distance are critical components of accurate categorization. Time and distance are intimately related within the concept of risk analysis in international relations. Thomas Schelling [45] identifies the relationship in discussing the use of weapons for signaling in his analysis of perceptions associated with various types of mobilization. Perceptions of risk for a tank at depot differ significantly from a tank loaded onto a rail car and further differ from a tank on the border. The tank is the object capable of engaging in violence, but its location enabling timely use either escalates or de-escalates the perceptions of risk of a potential target nation. Although the assessment of the weapon as offensive or defensive, active defensive or passive defensive, all alter the representativeness of the object, these work in tandem with the time a state has to categorize and respond to a weapon based on its proximate location in relation to its interests. As the time to use decreases, the likelihood of accurate categorization of both an object and an event similarly decreases. The time from launch to impact of an ICBM is ~15 min and SLBMs are potentially even shorter. The necessary identification of mobilization and utilization of a weapon occur within a short timeframe; however, the result of an ICBM's subsequent impact is

4 Here risk aversion is defined as the behavior of humans (or in this case states as unitary actors), who when operating under conditions uncertainty, attempt to lower that uncertainty.

5 Here risk acceptance is defined as the behavior of humans (or in this case states as unitary actors) who when operating under conditions of uncertainty are willing to accept risk and its potential impact.

well-known and reasonably predictable. Thus, categorization focuses not on the impact of the weapon, but whether it is truly being used and how to appropriately respond.

In contrast, the time it takes to engage in forensic analysis of malware and its proximity within the infrastructure of the defender elevates alarm and hampers effective assessments relative to impact. Here, the nuclear argument is almost inverted, while the time available to a defender is limited or nonexistent, as most cyber incidents are discovered after or during an attack, the proper categorization of a cyber-incident remains extremely difficult. While it is unlikely, the severity of a cyber-incident will rise to a level comparable with nuclear weapons, its impact can, or is often thought to be substantial [58]. The assessment of actions undertaken within cyberspace is plagued by problems of proper categorization in the short-term. This is due both to problems associated with the intentions of the actor and the qualities of the cyber capabilities themselves.

Tim Stevens [59], Eric Gartzke [60], and Max Smeets [61] note a further problem plaguing the proper categorization of cyber capabilities lies in their transitory nature (also known as the use it or lose it problem). Although Thomas Rid and Peter McBurney [62], and Rebecca Slayton [63] note that cyber weapons have increasing costs with complexity, Steven's, Gartzke's, and Smeet's argument on the transitory nature of the capabilities themselves has particular bearing on proper categorization.[6] Assessing new models of tanks, planes, ICBMs, or other conventional weapon systems while laborious and continually changing as technologies evolve, occurs far more slowly than comparable change occurring relative to the development of malware. Moreover, the increasingly digitized nature of conventional military and civilian hardware combined with rapid development timelines of malicious cyber capabilities challenges individual capability categorization.

A conventional attacker can facilitate the accurate categorization of the use of a given weapon system by slowing mobilization, moving civilians into shelters, and making increasingly bellicose displays of force. Each shift away from a status quo can begin the process of signaling to the potential defender what to expect. Such signals can reduce, if desired, errors of representation. Such displays both help to frame risks from a cognitive approach as well as foster more accurate utility calculations inclusive of resolve. Because shifts in conventional arms are increasingly difficult to hide at scale due to advances in intelligence collection, states are incentivized to facilitate the accurate framing of risk.

Accurate representation of capabilities and intent in cyberspace while potentially possible is not understood. If a state overtly broadcasts its movement through networked infrastructure toward an enemy's position, networks, or critical infrastructure within their homeland, the defender is likely to counter such moves through a variety of defenses including but not limited to patching, changing firewall, and IDPS configurations, or in dire circumstances cutting off global Internet access or turning systems off altogether to prevent damage. Progressing slowly and deliberately from one system to another does occur and is often a hallmark of advanced persistent threats (APTs), but this progression is nearly always clandestine and very often also covert. Signals in cyberspace to alter the representation of a cyber capability are impractical at best and damaging or undermining to the capability at worst.

Accurate representation of objects or events occurring in cyberspace differs substantially from representations in conventional domains. At the present time, representation of objects and events in cyberspace appears to be substantially more difficult based on the technical and temporal realities of cyberspace. This might not always remain true, but for the time being, framing risk using this first point of analysis appears to confound prior attempts at categorizing objects and events in a way that facilitates accurate probabilities. Actions undertaken in and through cyberspace, at least initially, often foster a great deal of uncertainty that undermines risk assessments. Very often the recipient of an attack is uncertain as to whether they are under attack, and if they are whether such an attack is espionage, subterfuge, attack, criminal behavior, or other; moreover, the intent of the attack, which often frames response options is often unclear at its outset.

Moving beyond challenges of representation, "availability refers to inferences about the frequency of events, where such frequency is judged according to the associations triggered in memory or imagination." [56]. As noted above, war and conflict are difficult to predict and constitute the "error term" [26], thus a lack of frequency impairs accurate judgment. Additionally, individuals are notoriously bad at assessing probabilities and are likely to overweight certain events or objects as demonstrated through repeated experimentation by Kahneman and others in psychology and economics [56, 64]. Whereas in conventional forms of conflict, there are often a dearth of cases upon which to formulate an analysis and a substantial overestimation of frequency, within cyberspace there are an overwhelming number of cases. Despite a large frequency of events occurring within cyberspace, these events are not normally distributed across types of events. Data from a variety of sources indicate that the actual volume of severe events that degrade, disrupt, destroy, or deny is limited [40]. Despite the lack of frequent severe events, the perception that severe events are frequent appears to be quite evident within the existing literature [65, 66].

Unlike availability problems in conventional conflict in which the severity of the rare events lead to an overestimation of their frequency; in cyberspace, it is the overwhelming frequency of all forms of events that obfuscates and confuses the frequency of rare events. It is not that individuals misinterpret that a large number of events are occurring, only the severity of those events. The likelihood of any given cyber-incident resulting in substantial physical damage due to the degradation, destruction, disruption, or denial of a system is low. The perception of high frequencies of severe attacks is exacerbated by news coverage that tends to elevate the severity of cyber incidents and conflate all forms of cyber incidents with cyber-attacks rather than distinguishing between those incidents, which are best described as theft or espionage from those better categorized as attacks.

Because cyber incidents of all kinds are conflated with significant attacks in cyberspace, there is little incentive for hostile parties to hold back on those activities short of causing levels of violence that might lead to war. Whether an attacker engages in a severe attack or espionage, perceptions toward the adversary remains generally the same. Understanding risks calculations requires more than raw numbers of attacks, and instead requires context, intent, insights into adversary, and defender capabilities. Additionally, because few cyber-incidents receive any form of response relative to the number of incidents undertaken and in tandem with the frequency and misidentification of incidents, attackers are prone to either over-estimate or discount the potential for being found out [67]. The assumption for

6 It is true that all weapon systems are transitory; however, the time horizon of most weapon systems is more than a couple of months or years as is the case for many cyber capabilities.

the attacker becomes one of risk seeking because the value of risk to reward is perceived as high relative to the value risk of loss as was clearly demonstrated in comments made by GEN. Nakasone, Commander of US Cyber Command in March 2018 when he stated: "I think that our adversaries have not seen our response in sufficient detail to change their behavior" [68].

Building on the first two concepts, anchoring[7] is predicated on the impact of initial values, regardless of their relevance to the assessment of risk in any given situation [56]. Anchoring is particularly a difficult bias to overcome in cyberspace as most individuals outside of security specialists do not understand how malware affects computers. Often the numbers used to represent the severity of various cyber incidents fail to offer decision-makers with an understanding of the severity of each system impacted. Numbers of daily attacks, accountings of aggregate numbers computer infected, and similar metrics obscure accurate notions of risk [69, 70]. Whereas the impact of a bomb on a building is relatively straightforward or nuclear weapon's blast impact on a city understood as total devastation, the relative impact of malware on a system is poorly understood and differs from incident to incident. Malware does not function like a bomb or a bullet and its impact differs markedly across individual computer systems [69]. Therefore, estimates of number of computers impacted needs more nuance in analysis. The number of people who catch a cold in a given year and the number of people who become sick in a given year are not equivalent. Being sick can range from extremely severe diseases to the common cold. There is a false logic in assuming equivalency [71]. Another similar example of anchoring can be derived from an analysis of perceptions of Cyberterrorism conducted by Gross, Canetti, and Vashdi. These authors found that it was the perception of the threat, rather than the actual reality of the threat that resulted in the most militant response [72]. The anchored perception thereby informs the calculation of risk.

Each of these forms of judgment errors informs the framing of the risk as understood by decision-makers. The framing of events and objects alters the willingness or individuals to undertake or assume risk. The manner in which these judgment errors are understood in cyberspace frequently diverges from the way in which they are understood in conventional domains of interaction. This divergence in understanding is aggravated by a lack of understanding on how computers and networks function, what types of objects (malware) impact them, the timing of the object utilization relative to the time a defender has to adapt and respond, the frequency of events, and the anchoring of information associated with events antecedent to an event or object creation and or utilization. It is also unclear how well states are able to frame their use of cyber capabilities in ways to manage perceptions of risk.

Cyber-attacks with the exception of public attacks meant to disrupt or deny access such as DDoS are covert or clandestine in their implementation for fear of patching, alterations of the target uptime, remote accessibility, or a variety of manipulable features that distinguish cyberspace from comparable risk framing in conventional state interactions [4]. While covert implies the deliberate obfuscation of both the attack and its sponsor, clandestine emphasizes the former. Whether covert or clandestine, the nature of cyber incidents, and the capabilities involved in undertaking them do not lend themselves toward the formulation of rational utility analyses. Moreover, the effects of cyber-attacks on adversaries often have mixed results that do not result in easily definable utility. Nadiya Kostyuk and Yuri Zhukov attempted to assess the impact of sustained cyber operations on an adversary in a single case, Ukraine. They found no significant coercive impact associated with low to moderate scale attacks [73]. Their study is not alone, Brandon Valeriano, Ryan Maness, and Benjamin Jensen were likewise unable to discern with a robust dataset any meaningful coercive impact derived from the overwhelming majority cyber incidents [34, 40]. The ability to win a war through cyber means alone outside of science fiction, remains unlikely and the therefore ascertaining the utility of attacks, i.e. assessing the positive yield derived from attacks beyond intelligence or theft remains difficult to assess.

Whether it is the framing of risk in cyberspace within international interactions or the codification of accurate utilities based on bounded assessments of capabilities and potential rewards, understanding state risk and uncertainty in cyberspace remains a difficult task. This difficulty impairs decisions and increases the challenges of states operating within cyberspace.

Perceptions of the potential for loss are difficult to assess in cyberspace as illustrated by the challenges of framing and errors in judgment. Among the challenges faced, the ability to remain anonymous can mean that losses can occur months or years before they are detected. The average discovery time for victims of data breaches in 2017 was 191 days [74]. This finding poses significant challenges to scholars of international relations. If states do not know they have been attacked, or that they have suffered a loss, or they only find out far after the fact how do we interpret their perceptions of risk associated with that loss. The short answer from econometrics is that assessing distant future losses leads individuals to discount that loss [75]. The discounting of risk over time shields both the violator and the violated. The violator perceives low risk for the engagement in actions. Even small actions can result in disproportionate gains with little to no consequence, or consequences so far in the future that they are not included within present risk framing or rational choice. The defender not knowing or being unable to assess where risks are located must either allocate resources blindly or elevate the costs associated with violations to such a degree that it mitigates discounting. The allocation of resources is not truly blind, as all states and individuals have certain items, which they hold more dearly than others, and it is toward these privileged items that they devote more time and energy on protection. The inability to close the temporal divide between a violation and discovery undermines accurate risk and utility calculations.

Beyond the temporal challenges of framing risks accurately within cyberspace are the relationship of actions and effects. A bombing campaign such as those conducted against Serbian forces in the 1990s used kinetic munitions as a first-order action to achieve effects. The bomb landing on a target results in an immediate effect, i.e. explosion. Yet, in cyberspace as noted by Herbert Lin [5], many cyber effects occur as second- or third-order consequences of actions. The third-order effect is often difficult to directly connect to an initial action. Although frequently used as an example, Stuxnet illustrates the third-order effect principle nicely. Kim Zetter in writing about Stuxnet noted that beyond simply manipulating the control terminals overseeing centrifuges and the PLCs themselves, the imbalanced cycling of Uranium Gas caused the U-235 gas to be deemed in impure and resulted in the dumping (wasting) of a precious commodity into cooling tanks [38]. The wasting of U-235 gas limited the production potential and capacity of Iran. To make the leap from the entry point of the attack through to the manipulation

7 Rose McDermott defines anchoring as: "Anchoring relates to predictions that are based on initial values, or anchors, that may or may not be adequately adjusted before a judgment is made of a second, possibly unrelated, object or event."

of various stages in the process to the ultimate effect of wasting a rare mineral demonstrates a third order effect. Framing such risks is difficult whether one is an attacker or a defender. Developing rational decisions based on a third order effect requires a wholly different level of utility calculation than is commonly present in most models on international politics that see the effects as cause-and-effect pairs.

Yet, more than a challenge with time and cause and effect pairs, the complexity of networked environments makes the unintended spread of cyber capabilities beyond intended targets far more possible than the comparative use of conventional capabilities with the exception of biological weapons and perhaps nuclear fallout (both of which far exceed the presently demonstrated destructive power of cyber capabilities). The notpetya malware illustrates the complexity of the global digital environment [58]. Andy Greenberg in writing on the effects of notPetya on the global shipping firm Maersk noted that it was not for an unexpected power outage in Ghana, the company would not have been able to recover its operational documents and would have suffered far greater than it did [76]. Assessing risk within highly complex environments in which systems geographically separated by thousands of miles can inadvertently impact one another challenges how states plan for and assess risk as well as the potential benefits of actions both malicious and defensive in nature. While the impact of NotPetya likely far exceeded its developer's wildest dreams, its inability to "finish the job" was hampered by the complex environment it was itself operating in. Likewise, Maersk was saved by that same complexity and randomness within it. Introducing malicious software into networked environments absent full considerations of the complexity of those environments can compound risks in ways that an anchored starting point might not have accounted for.

A failure to fully understand the functional limits and interactions of cyberspace exacerbates perceptions of risk; uncertainty surrounding adversary cyber capabilities elevates the perceptions of risk associated with adversary cyber capabilities. This failure is both a result of technical challenges, but it should not overlook the fact that concepts of risk and uncertainty associated with cyber capabilities befuddle decision-makers with an overwhelming amount of jargon and complex concepts. Moreover, because effects are often not direct, but rather second or their order in nature decisions on their use are mired in abstraction.

All the issues listed above and more impede the development of accurate risk assessments in cyberspace and make predicting and planning actions difficult. Yet, aside from those issues addressed here and modeled on concepts first developed by McDermott is yet another substantial problem that undermines the ability of states to formulate assessments of risk and uncertainty. Where William Gibson once defined Cyberspace as a "Consensual hallucination," [77] and Ernest Cline prognosticated on the dystopian future of a world fully enraptured and melded with the digital [78], the fact remains that despite similarities digital and physical actions result in differing cognitive responses. The next section turns to a few brief case examples in which risk and uncertainty in offensive and defensive actions in cyberspace are examined.

Exploring the cyber decision-making heuristic

As I have developed above, decisions are complex. Whether decisions deal with conflict in conventional domains or virtual ones, how states bound and shape their decision matrices often determines their perceptions of risk. While every decision is unique, each decision is informed by those decisions, which preceded it. The preceding decisions, how they were constructed, and their results constitute the basis upon which future decisions are made. Very often decisions appear to be emotional in nature, yet, as examined by Anthony Damasio, emotion, is the product of learning [10], it is a functional shortcutting of rigor and robust thinking that forces us into what Daniel Kahneman calls "system 1" thinking [79]. While this article emphasizes rational and cognitive thought, these structures of decision are bounded by the learning frameworks within which states and their principal decision-makers decide. While the cold hard exposition of rational choice is seemingly devoid of emotion, the reality is that as clinical studies time and again demonstrate, rationality is necessarily shaped by what could best be mathematically referred to as an informed prior. Absent this informed prior, human minds tend to cycle endlessly between probabilities unable to rationally make complex optimizing decisions [10]. To assess decisions, to understand them, we must understand the process by which states arrived at their logic and upon what data they built their informed prior. We must understand the learning framework, the heuristic that defines how they interpret what we later build into rational or cognitive models. To do this, I highlight two cases below. These cases are vignettes that highlight the way in which two actors, the USA and the Russian Federation, have learned from their actions in cyberspace. First, I examine the actions of the Russian Federation in 2014–16 and the USA response in 2016–17 and again in 2018. Second, I examine the well-worn and familiar path of Stuxnet and the divergent takeaways between the USA and the Islamic Republic of Iran. These cases are meant to highlight the rough contours of the development of perceptions of risk and uncertainty by different actors.

Examples of bears and eagles learning

Below is a comparison of actions taken by the Russian Federation and the USA in cyberspace. The intent of these short cases is to highlight how states might be translating experiences into perceptions about how to operate and engage in activities in and through cyberspace. As noted above, learning is a complex process and direct causality should not be inferred; however, there are strong correlations in actions by actors over time to their likely perceptions of operating within cyberspace. The state behaviors in these cases contrast substantially and therefore indicate divergence in perceptions arising from similar activities.

There is little doubt that the Russian Federation through the use of APT-28 and APT-29 both hacked and engaged in substantial targeted information operations against the 2016 US Presidential Elections [80]. The origins of the hacking began well-prior to the 2016 election. There are indications that the FBI was aware of Russian cyber activities to undermine the 2016 elections dating back to at least 2014 and had alerted the Democratic National Committee (DNC) to intrusions by suspected Russian hackers as early as September 2015 [81]. Nor was it a secret that Russia had a long and sordid history with information manipulations and hacking efforts. Ofer Fridman in tracing the lineage of modern Russian "hybrid" efforts identifies a robust discourse dating back to the height of the Cold War in which Russian military leaders and civilian strategists both at home and in the diaspora wrote about and conceptualized many of the concepts associated with what has become known as active measures [82]. Nor were Russian cyber capabilities entirely a mystery, senior leaders in the Department of Justice and the Department of Defense were aware of growing cyber

capabilities of the Russian Federations both in zones of conflict such as Ukraine [83, 84] and Syria, but also in criminal endeavors as well [85]. Further knowledge of Russian cyber activities in the run up to the 2016 elections were discovered by private security firms and academic researchers [58].

The Russian efforts leading up to the 2016 election appears to have been fragmented and disjointed. The nature of the attacks and subsequent information operations spanned multiple agencies within the Russian Federation and seemingly lacked formal unitary organizational direction and decision-making processes [80]. The coordination of activities within individual agencies such as the GRU (Russian Military Intelligence), SVR (Russian Foreign Intelligence), and privately funded, but state-sanctioned organizations such as the Internet Research Agency (IRA) was unitary and might have possibly had central government coordination, yet the fragmented nature of their disparate efforts led to operational overlap and the obscuring of signaling if any was formally intended [80].

Russian activities in Ukraine from 2013 onward went largely unanswered with the exception of rhetoric out of Washington and Brussels. Its efforts in Syria sailed largely under the radar. However, its efforts to impugn the 2016 election and similar elections in the UK, France, and Germany received substantial attention. With relatively limited resource allocations and common exploits, the Russian attack against the USA began with simple phishing attacks designed to steal the credentials of DNC or members of the 2016 Clinton campaign. These phishing attempts yielded the motherload when Clinton campaign Chairman John Podesta, following what he thought was the advice of the IT team, clicked on a link and divulged his credentials to the Russians. The Podesta leaks combined with similar leaks out of the DNC fed a relatively inexpensive ($1.25 Million) [86], yet sizeable (>10.6 million pieces of content) [86] disinformation program that sought to undermine the credibility of both the campaign and the democratic party. These leaks were combined with a steady drumbeat of dis/misinformation seeking to sow discord within the US electorate.

While drawing a causal chain from Russian activities to election outcomes is not possible or advisable, Russian efforts had an impact on the perceptions and discourse surrounding the election, the newly elected president, and the stability/resilience of American democracy [87]. Yet, the US response to these activities was anemic at best. A warning by President Obama, a minor uptick in economic sanctions, the closure of several diplomatic facilities (which were ostensibly being used for espionage efforts against the USA), and the expulsions of diplomatic staff provided a clear signal to the Russian federation that it could wreak havoc and suffer few costs in response.

The paltry response of the Obama administration failed to impose costs that might change the decision framing of the Russian Federation. To correct this, the National Security Agency (NSA) and US Cyber Command's (USCYBERCOM) actions against the IRA in 2018 were an attempt to impose costs to change the risk framing of an adversary. Whereas the Obama Administration's response was cross domain, broad, and inexact, the Trump administration used an exquisitely targeted, narrow in design and implementation preemptive strike to impose costs to change the way in which Russians frame or conceive of actions against the USA. US actions against the IRA were made possible through a revision in US policy codified in in National Security Presidential Memorandum 13 (NSPM-13) [88]. NSPM-13 altered constraints on the use of military operations in cyberspace initially established in Presidential Policy Directive 20 (PPD-20) and in the words of then National Security Advisor John Bolton effectively untied the hands of the US Department of Defense [88].

The attack against the IRA was well-coordinated and part of a joint plan formulated by the NSA and USCYBERCOMMAND that resulted in the temporary disconnection of the IRA from the Internet [89]. Russian news channel, Federal News Agency (FAN), reported on the attacks and claimed that US hackers had infected one of the IRA's internal IT servers and destroyed a RAID controller and half of the drives attached to it [90]. FAN further stated that the vector of the attack originated via social engineering in the form of a Trojan laced emailed [90]. US Defense officials claimed that the objective of the attack was to "throw a little curveball, inject a little friction, sow confusion" [91]. The narrow-targeted nature of the attack allows for the analysis the decision to attack within a constraints of risk modeling and the uncertainties present.

By nearly all measures, the US attack against the IRA was highly constrained. The attack was focused on disrupting a nonstate, but state affiliated actor with limited capacity for cyber response actions. The actor was of limited strategic or tactical importance to the overall defensive or offensive capabilities of the principal adversary, the Russian Federation. The attack itself was temporary and the IRA was expected to recover to full operational capacity within a short recovery period following the attack. The attack was, according to US Defense sources combinatory in nature featuring both cyber effects and information effects in the form of targeted messages sent to employees of the IRA [91].

The limited nature of the attack's effects, its public acknowledgement, its selective targeting of nonstate, noncritical, limited value infrastructure all combine to highlight a strategy of extreme risk aversion. The effects of the operation were quantifiable, the uncertainty of the effects and likelihood for a potential adversary response were minimal. Rob Knake sums it up best: "Our response to a very hostile act is we're going to cause connectivity problems? That's not a terribly strong signal. If you shut off the internet for all of Russia, that's a signal. Isolating one building I don't think is much of one" [89]. The resultant risks of the US actions against the IRA were comparatively low.

The divergence in risk models between states engaged in cyber activities in these two linked incidents could not be starker. The Russian Federation, hacked, US election infrastructure, political parties, and engaged in information operations with hacked data. These operations combine to form risk acceptant behavior, expecting that each individual act, viewed discretely and after completion through the lens of forensic analysis would in all probability not result in a strong response, yet the uncertainty, the unquantifiable effects and tangible results of the actions was by all accounts unknown [80].

When looking at the two decisions, one to engage in *Ad hoc*, semi-coordinated attacks against the critical infrastructure of a foreign adversary with little to no known ultimate benefit versus the structured, coordinated, highly organized attacks against a nonstate, noncritical infrastructure party within the territory of a foreign adversary. We see two fundamentally different conceptualization of risk and levels of tolerance for uncertainty. While the often-used term for Russian activities is adventurism, thus leading to claims of some potential cult of the offensive within cyberspace, the reality is more banal. Russian cyber actors have learned through consistent and repeated cyber actions taken in Ukraine [92, 93], Estonia [94], Georgia [95], and elsewhere, that with limited effects and limited consequences that cyber activities are a cheap substitute or addition to conventional military power, and further maintain the relevance of the state within the international system [87]. Adventurous, perhaps, but a highly constrained type of adventurism with effects substantially less than those associated with more conventional forms of conflict, and certainly less relative risk. Even operating under

conditions of uncertainty, the calculable risks provide substantial room for action.

The USA, in contrast, having engaged in similar behaviors to that of the Russian Federation arrived at a different level of tolerance for both risk and uncertainty. This is best illustrated through a brief examination of its early adventurism against the Islamic Republic of Iran.

Losing cyber precision guided munitions

Many books and articles have been written on Stuxnet [38, 63, 96–101]. I will not rehash their findings here but rather focus on the effects of Stuxnet and the way in which it informed US learning about risk and uncertainty regarding cyber conflict. It is important to note that at the time of its discovery, Stuxnet was by far the most advanced malware ever discovered, leveraging multiple 0 days. Its construction was itself immensely risk averse and constructed with substantial legal oversight [102]. Its effects have been judged to have been limited [63]. When Stuxnet was designed every effort was undertaken to mitigate risk and to manage uncertainty [38]. Its code was novel, its implementation deliberate, its specification to systems precise. One might imagine commanders engaged in an operational planning meeting discussing how they had minimized the risks for nearly every contingency. It is further possible to imagine that they felt certain after substantial testing that the malware they had designed would work with the intended effects. With uncertainty low and risks minimized, the action could proceed. Yet, unlike more risk acceptant actors, the progress of implementing Stuxnet was laborious. While risks had been accounted for, there were still unknowns that could not be quantified and therefore while uncertainty was low, it was still present. With each iteration and progression into Natanz the confidence in the probabilities for success increased. Yet, despite the enormous care given to the implementation of this advanced malware by 2010, the unthinkable had happened. A small Belarusian cybersecurity firm VirusBlokAda had identified the malware and within weeks, it was making the rounds through various security companies and eventually to Symantec where the malware was deconstructed. Yet, beyond issues posed by having a highly targeted malware detected and deconstructed, the certainty thought to have been baked into the code through risk mitigation strategies designed to contain the virus to Iran failed and the malware spread around the world. Engineers at Symantec and elsewhere discovered numerous 0-day exploits within the malware [103]. Although the code itself was highly targeted, its susceptibility to reverse engineering was high [38]. Within a year of its release, the wild variants of the malware were being utilized and seen on networks around the world [104]. While the effects of the malware on nontargeted countries were minimal, the impact of the malware's spread on the decision-making processes of the USA was likely substantial.

The USA, in an attempt to achieve geopolitical gains (the slowing or halting of Iranian Highly Enriched Uranium production), in a covert, quasi legal manner (not impacting nonmilitary infrastructures), saw its efforts derailed and its covert nature and geopolitical ramifications expand. Widely considered as the first serious "digital weapon" [38], its spread beyond its intended use posed a substantial technical danger to USA and allied infrastructures through potential reverse engineering as well created geopolitical problems for the suspected perpetrators of the attack. Finally, its failure opened the door for the Iranian regime to respond, which it did with attacks against US interests in the Middle East through a devastating attack against Saudi Aramco [105], and attacks against US interests at home [106]. Within 2 years of the disclosure of the Stuxnet attacks, PPD-20 had imposed robust constraints on US cyber activities, requiring presidential authorization for their use.

Perceptions of risk and uncertainty as informed by heuristic biases result in divergent state behaviors in cyberspace. While these biases operate similarly to those in conventional domains, the underlying characteristics are often obfuscated within a virtual domain. The final section below pulls the concepts of risk and uncertainty in cyberspace together and examines the heuristic for how states arrive at decisions.

Framing and confounding uncertainty and risk in cyber conflict

Cyberspace as a complex system of systems that is constantly changing, whose artifacts are dynamic in logical and at times geospatial existence, do not lend themselves to static assessments of risk and uncertainty. Generally, as noted by Gomez and Villar, fixed normative frameworks are disadvantaged when seeking to learn about decisions pertaining to cyberspace [107]. The lived experience of cyberspace by states conditions their perceptions and informs their biases. The concrete nature of physical space, of bombs, missiles, soldiers and tanks, and their potential first-order effects is subject to misinterpretation, a bias in representing their intention, based on time and space. Their presence, and likely capability clear, still results in poor assessments of risk bounded by differing levels of uncertainty. Iterative interactions with adversaries can provide a measure of learning that in many ways mitigates biases and facilitates learning thus fostering better decisions with more quantifiable risks under conditions with manageable levels of uncertainty. Black swan events and those events that occur outside of quantifiable notions of what is and are not possible are present in all interactions; however, as Erik Gartzke notes they form the "error term" of conventional conflicts.

Several millennia of experience have honed state notions of risk and uncertainty on decisions pertaining to conventional conflict, and yet conflicts occur. Several decades of activity in a rapidly evolving domain of interactions, cyberspace, has resulted in divergent perceptions and tolerances for risk and uncertainty by states. Some states, such as the Russian Federation, China, Iran, and North Korea, find the uncertainty of cyberspace tolerable within an international context and the risks of actions undertaken manageable and perhaps liberating. Other states, particularly the USA and like liberal democracies find the uncertainties of cyberspace constraining, and the risks difficult to fully accept, outside of extremely limited and controlled interactions.

Moreover, the frequency of interactions in cyberspace increases perceptions of risk in some actors in cyberspace, while decreasing the perceptions of risk in others. Similarly, some actors see the increased number of activities in cyberspace as increasing the uncertainty of state intentions and behaviors, while others see it as decreasing the uncertainty of state intentions. This is best thought of as a divergence of uncertainty of intent. As noted by Russel Leng, "in a world where realism remains the dominant diplomatic culture, successful statesman cannot escape the requirement to demonstrate resolve…" [108]. As a consequence, when the Russian Federation and China continue to engage in espionage and attacks against the USA or others globally with little consequence, they grow more certain that those states they are targeting, despite their rhetoric, are unlikely to respond in any meaningful way. Their perceptions of risk

and reward indicate that with limited risk they can continuously reap rewards. In contrast, the USA and her allies grow increasingly uncertain as their experience has shown them that even with highly quantifiable risks and supposedly low uncertainty, actions within cyberspace can result in unintended consequences. Unintended consequences are best thought of as uncertainty of effect. What is unique and fascinating is that while there are significant discrepancies in certainties of intent among states in cyberspace, where the two groups more substantially diverge is on issues pertaining to uncertainty of effect. Because some actors are risk tolerant, they are able to undertake actions with uncertain effects, whereas other states, particularly the USA, are risk averse and only willing to undertake actions with certain effects.

The imbalance in risk perception coupled the inherent uncertainties of intent and effect undertaken by adversaries within the domain, the potential responses of other states, the potential for unintended consequences results in divergent decision structures. The most prominent example of this is the 2018 US Department of Defense Cyber Strategy [109]. The strategy's proponents argue that all of the challenges listed above require that states engage in what amounts to constant contact, a means by which to create friction and signal adversaries [110, 111]. Moreover, because states are unable to affect the risk calculus of other states through conventional deterrence mechanisms as has been demonstrated in a variety of instances [111–116], the next best thing is to foster an environment in which states consistently interact to provide information to one another. The notion of fostering friction has received some push back and is thought to increase risk and uncertainty, while being concurrently escalatory [117]. Yet, as demonstrated in multiple studies, actions in cyberspace are not inherently escalatory [118].

What remains is an environment or "domain" of interaction of states that has many of the same characteristics residents within conventional physical environments and domains of interaction, but yet responds differently for a variety of reasons. The result of these differences is a domain of constant and evolving contention on everything from definitional issues over cyber war and information war to more nuanced issues that impact the way in which states evaluate risk and uncertainty. The present state heuristic is to learn by doing. By pushing the bounds of acceptable behavior, states are learning what the attendant risks of cyberspace are. Through repeated actions against one another states are learning and, in some instances, reducing uncertainty. While other states, in pushing the bounds of state-on-state interaction, further increase their uncertainty. The challenge remains, as illustrated by Kreps and Schneider, that activities in cyberspace are perceived differently than comparable activities in the physical world [119]. This combined with the proliferation of technology as well as its continued advancement into nearly all aspects of national security is that the breadth and potentially the depth of uncertainties associated with actions in cyberspace continue to grow. Because there remains a high degree of uncertainty about both the perceptions and the effects of activities in cyberspace, the ability to formulate informed, if bounded, that risk calculations will remain highly imperfect in the near future. The advantage of an environment in which risks of escalation are low, is that learning, while potentially costly, is not costly in a way equivalent to more conventional state-on-state interactions. The result is that states can, to some degree, embrace what feels right to them within the present learning by doing heuristic. However, if states seek to shift or alter the behavior of others within the international system, it will be increasingly important to understand how a target of their manipulations arrived at their present calculations of risk and perceptions of uncertainty. With this knowledge, states can then begin to formulate policies to alter a potential adversary's behavior.

Conflict of interest statement. None declared.

References

1. Berejikian J. Model building with prospect theory: a cognitive approach to international relations. *Polit Psychol* 2002;**23**:759–86.
2. Jervis R, Lebow RN, Stein JG. *Psychology and Deterrence*. Baltimore, MD: Johns Hopkins University Press, 1985.
3. Huth P, Bennett DS, Gelpi C. System uncertainty, risk propensity, and international conflict among the great powers. *J Confl Resol* 1992;**36**: 478–517.
4. Brantly AF. *The Decision to Attack: Military and Intelligence Cyber Decision-Making*. Athens: University of Georgia Press, 2016.
5. Lin HS. Operational considerations in cyber attack and cyber exploitation. In: Reveron DSs (ed.), *Cyberspace and National Security*. Washington, D.C.: Georgetown University Press, 2012.
6. Von Neumann J, Morgenstern O. *Theory of Games and Economic Behavior*. Princeton, N.J, Woodstock: Princeton University Press, 2004.
7. Levy JS. Learning and foreign policy: sweeping a conceptual minefield. *Int Org* 1994;**48**:279–312.
8. Knight FH. *Risk, Uncertainty and Profit*. Boston: Houghton Mifflin Company, 1921.
9. Moravcsik A. Taking preferences seriously: A liberal theory of international politics: Erratum. *Int Org* 1998;**52**:229–229.
10. Damasio AR. *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam, 1994.
11. Taleb N. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House Trade Paperbacks, 2010.
12. Jervis R. Political implications of loss aversion. *Polit Psychol* 1992;**13**: 187.
13. Levy JS. Applications of prospect theory to political science. *Synthese* 2003;**135**:215–241.
14. Berejikian JA. Cognitive theory of deterrence. *J Peace Res* 2002;**39**: 165–183.
15. O'Neill B. Risk aversion in international relations theory. *Int Stud Q* 2001;**45**:617–640.
16. Kim W, de Mesquita BB. How perceptions influence the risk of war. *Int Stud Q* 1995;**39**:51–65.
17. Bueno de Mesquita B. *The War Trap*. New Haven: Yale University Press, 1983.
18. Quattrone G, Tversky A. Contrasting rational and psychological analyses of political choice. *Am Polit Sci Rev* 1988;**82**:719–36.
19. Kahneman D, Tversky A. Prospect theory: An analysis of decision under risk. *Econometrica* 1979;**47**: 263–92.
20. Fearon JD. Rationalist explanations for war. *Int Organ* 1995;**49**: 379–414.
21. Mintz A, Wayne C. *The Polythink Syndrome: U.S. Foreign Policy Decisions on 9/11, Afghanistan, Iraq, Iran, Syria, and ISIS*, Stanford University Press, 2016.
22. Kissinger H. *Diplomacy*. USA: Simon & Schuster, 1994.
23. Pevehouse J, Nordstrom T, Warnke K. The correlates of war 2 international governmental organizations data version 2.0. *Confl Manage Peace Sci* 2004;**21**:101–119.
24. Allison GT. *Essence of Decision; Explaining the Cuban Missile Crisis*. Boston: Little Brown, 1971.
25. Jervis R. *Perception and Misperception in International Politics*. Princeton, N.J.: Princeton University Press, 1976.
26. Gartzke E. War is in the error term. *Int Organ* 1999;**53**:567–587.
27. Warner M. *Intelligence in Cyber-and Cyber in Intelligence*. Washington, DC: Georgetown University Press, 2017.
28. Jervis R. Reports, politics, and intelligence failures: The case of Iraq. *J Strateg Stud* 2006;**29**:3–52.

29. Marrin S. Preventing intelligence failures by learning from the past. *Int J Intell CounterIntell* 2004;**17**:655–672.
30. Hart J. Three approaches to the measurement of power in international relations. *Int Organ* 1976;**30**:289–217.
31. Branch J. What's in a name? Metaphors and cybersecurity. *Int Organ* 2020;1–32. doi:10.1017/S002081832000051X.
32. Reveron DS. *Cyberspace and National Security: Threats, Opportunities, and Power in a Virtual World*. Washington, DC: Georgetown University Press, 2012.
33. Nye JS. *Cyber Power*. Cambridge, MA: Harvard Kennedy School, Belfer Center for Science and International Affairs, 2010.
34. Valeriano B, Maness RC. *Cyber War versus Cyber Realities: Cyber Conflict in the International System*. New York, NY: Oxford University Press, 2015.
35. Rid T. Cyber war will not take place. *J Strateg Stud* 2012;**35**:5–32.
36. Healey J. *A Fierce Domain: Conflict in Cyberspace, 1986 to 2012*. Arlington, VA: Cyber Conflict Studies Association, 2013.
37. Lindsay JR, Cheung TM, Reveron DS. *China and Cybersecurity: Espionage, Strategy, and Politics in the Digital Domain*. New York: Oxford University Press, 2015.
38. Zetter K. *Countdown to Zero Day: Stuxnet and the Launch of the World's First Digital Weapon*. New York: Crown Publishers, 2014.
39. Brantly AF. The violence of hacking state violence and cyberspace. *Cyber Defense Rev* 2017;**2**:73–92.
40. Valeriano B, Jensen BM, Maness RC. *Cyber Strategy: The Evolving Character of Power and Coercion*. Oxford: Oxford University Press, 2018.
41. Cyber Security Strategy for Germany. *Interior FMots*. Berlin: Cyber Security Strategy for Germany, 2011, 1–20.
42. Clarke RA, Knake R. *Cyber War: The Next Threat to National Security and What to Do about It*. New York: HarperCollins Publishers, 2010.
43. Axelrod R, Iliev R. Timing of cyber conflict. *Proc Natl Acad Sci U S A* 2014;**111**:1298–1303.
44. Brantly AF. Cyber actions by state actors: Motivation and utility. *Int J Intell CounterIntell* 2014;**27**:465–484.
45. Schelling T. *Arms and Influence*. New Haven, Conn: Yale University Press, 1966.
46. Brantly AF. Aesop's wolves: The deceptive appearance of espionage and attacks in cyberspace. *Intell Natl Security* 2016;**31**:674–685.
47. Gartzke E, Lindsay JR. Weaving tangled webs: Offense, defense, and deception in cyberspace. *Secur Stud* 2015;**24**:316–348.
48. Buchanan B. *The Cybersecurity Dilemma Hacking, Trust and Fear between Nations*. Oxford: Oxford University Press, 2017.
49. Gartzke E, Lindsay JR. Thermonuclear cyberwar. *J Cybersecur* 2017;**3**: 37–48.
50. Tang S. Fear in international politics: Two positions. *Int Stud Rev* 2008; **10**:451–471.
51. Borghard ED, Montgomery M. Defend forward as a whole-of-nation effort. *Lawfare* 2020.
52. Carson A, Yarhi-Milo K. Covert communication: The intelligibility and credibility of signaling in secret. *Secur Stud* 2017;**26**:124–156.
53. Fearon JD. Signaling foreign policy interests: Tying hands versus sinking costs. *J Confl Resolut* 1997;**41**:68–90.
54. van der Meer S. *Signalling as a Foreign Policy Instrument to Deter Cyber Aggression by State Actors*. The Hague: Netherlands Institute of International Relations, 2015.
55. Sanger DE. *The Perfect Weapon: War, Sabotage, and Fear in the Cyber Age*. New York, NY: Broadway Books, 2018.
56. McDermott R. *Risk-Taking in International Politics: Prospect Theory in American Foreign Policy*. Michigan: University of Michigan Press, 2001.
57. Clark RM. *Intelligence Analysis: A Target-Centric Approach*. Los Angeles, CA: CQ Press/SAGE, 2020.
58. Greenberg A. *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers*. New York: Doubleday, 2019.
59. Stevens T. *Cyber Security and the Politics of Time*. Cambridge, UK: Cambridge University Press, 2015.
60. Gartzke E. The myth of cyberwar: Bringing war in cyberspace back down to earth. *Int Secur* 2013;**38**:41–73.
61. Smeets M. A matter of time: on the transitory nature of cyberweapons. *J Strateg Stud* 2018;**41**:6–32.
62. Rid T, McBurney P. Cyber-weapons. *RUSI J* 2012;**157**:6–13.
63. Slayton R. What is the cyber offense-defense balance? Conceptions, causes, and assessment. *Int Secur* 2017;**41**:72–109.
64. Kahneman D. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux, 2013.
65. Kaplan FM. *Dark Territory: The Secret History of Cyber War*. New York: Simon & Schuster, 2016.
66. Brenner J. *Glass Houses: Privacy, Secrecy, and Cyber Insecurity in a Transparent World*. London: Penguin Books, 2013.
67. Valeriano B, Jensen B. *The Myth of the Cyber Offense*. CATO Institute: CATO Institute, 2019, 1–16.
68. Baldor LC. *Army officer: China, Russia Don't Fear US Cyber Retaliation*. Washington, DC: The Washington Post, 2018.
69. Jardine E. Mind the denominator: towards a more effective measurement system for cybersecurity. *J Cyber Policy* 2018;**3**: 116–24.
70. Jardine E. The trouble with (supply-side) counts: the potential and limitations of counting sites, vendors or products as a metric for threat trends on the Dark Web. *Intell Natl Secur* 2019;**34**:95–111.
71. Jardine E. Global cyberspace is safer than you think: Real trends in cybercrime. *SSRN Electron J* 2015;1-32.
72. Gross ML, Canetti D, Vashdi DR. Cyberterrorism: Its effects on psychological well-being, public confidence and political attitudes. *J Cybersecur* 2017;**77**:138–110.
73. Kostyuk N, Zhukov YM. Invisible digital front. *J Confl Resolut* 2019; **63**:317–347.
74. 2017 Cost of Data Breach Study. MI: Ponemon Institute, 2017, 1–35.
75. Loughran TA, Paternoster R, Weiss D. Hyperbolic time discounting, offender time preferences and deterrence. *J Quant Criminol* 2012;**28**: 607–628.
76. Greenberg A. A Critical Intel Flaw Breaks Basic Security for Most Computers. *Wired, Conde Nast*, www.wired.com/story/critical-intel-flaw-breaks-basic-security-for-most-computers/ (6 January 2018, date last accessed).
77. Gibson W. *Neuromancer*. London: Harper Voyager Publishers, 2013.
78. Cline E. *Ready Player One*. New York: Crown Publishers, 2011.
79. Kahneman D. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux, 2011.
80. Rid T. *Active Measures: The Secret History of Disinformation and Political Warfare*. New York: Farrar, Straus and Giroux, 2020.
81. Lipton E, Sanger DE, Shane S. The perfect weapon: How Russian Cyberpower Invaded the U.S. *The New York Times,* 2016.
82. Fridman O. *Russian 'Hybrid Warfare': Resurgence and Politicisation*. New York, NY: Oxford University Press, 2018.
83. Brantly AF, Cal NM, Winkelstein DP. *Defending the Borderland: Ukrainian Military Experiences with IO, Cyber, and EW*. West Point, NY: US Army Cyber Institute, 2017, 1–60.
84. Fitzgerald CW, Brantly AF. Subverting reality: The role of propaganda in 21st century intelligence. *Int J Intell CounterIntell* 2017;**30**:215–240.
85. Carlin JP, Graff GM. *The Dawn of the Code War: America's Battle against Russia, China, and the Rising Global Cyber Threat*. New York, NY: PublicAffairs, 2019.
86. United States Senate. *Senate Intelligence Committee Report on Russian Interference in the 2016 United States Presidential Election*. Washington, DC: United States Senate, 2019.
87. Jasper S. *Russian Cyber Operations: Coding the Boundaries of Conflict*. Washington, DC: Georgetown University Press, 2020.
88. Nakashima E. White House authorizes 'offensive cyber operations' to deter foreign adversaries. *The Washington Post*. Washington, DC, 2018.
89. Greenberg A. US Hackers' Strike on Russian Trolls Sends a Message—but What Kind? *Wired.com* 2019.
90. Cimpanu C. US wiped hard drives at Russia's 'troll factory' in last year's hack. *ZDnet* 2019.
91. Nakashima E. U.S. Cyber Command operation disrupted Internet access of Russian troll factory on day of 2018 midterms. *The Washington Post*. Washington, DC, 2019.

92. Greenberg A. How An Entire Nation Became Russia's Test Lab for Cyberwar. *Wired*, 2017.
93. Brantly AF, Collins L. A bear of a problem: Russian special forces perfecting their cyber capabilities. *Army Magazine* 2018;**68**, https://www.ausa.org/articles/bear-problem-russian-special-forces-perfecting-their-cyber-capabilities.
94. Kaiser R. The birth of cyberwar. *Polit Geogr* 2015;**46**:11–20.
95. Beehner L, Collins L, Ferenzi S, *et al*. *Analyzing the Russian Way of War: Evidence from the 2008 Conflict with Georgia*. https://mwi.usma.edu/wp-content/uploads/2018/03/Analyzing-the-RussianWay-of-War.pdf 2018, (March 15, 2021, date last accessed).
96. Maathuis C, Pieters W, Berg Jvd. Cyber Weapons: A Profiling Framework. In *2016 International Conference on Cyber Conflict (CyCon U.S.)* 2016, 1–8.
97. Simonenko M. Stuxnet and nuclear enrichment of the cyber security regime. *Secur Index Russian J Int Secur* 2013;**19**:85–97.
98. Kushner D. The real story of stuxnet. *IEEE Spectrum* 2013;**50**:48–53.
99. Barzashka I. Are cyber-weapons effective? *RUSI J* 2013;**158**:48–56.
100. Farwell JP, Rohozinski R. Stuxnet and the future of cyber war. *Survival* 2011;**53**:23–40.
101. Langner R. Stuxnet: Dissecting a cyberwarfare weapon. *IEEE Secur Privacy* 2011;**9**:49–51.
102. Lindsay JR. Stuxnet and the limits of cyber warfare. *Secur Stud* 2013;**22**:365–404.
103. Falliere N, Murchu LO, Chien E. W32.Stuxnet Dossier. Symantec, 2010.
104. Zetter K, Son of Stuxnet Found in the Wild on Systems in Europe. *Wired*. Wired.Com: Wired, 2011.
105. Bronk C, Tikk-Ringas E. The Cyber Attack on Saudi Aramco. *Survival* 2013;**55**:81–96.
106. Kagan FW, Stiansen T. *The Growing Cyberthreat from Iran: The Initial Report of Project Pistachio Harvest*. Washington, DC: American Enterprise Institute, 2015.
107. Gomez MA, Villar EB. Fear, uncertainty, and dread: Cognitive heuristics and cyber threats. *Polit Governance* 2018;**6**:61–72.
108. Leng RJ. Escalation: Competing perspectives and empirical evidence. *Int Stud Rev* 2004;**6**:51–64.
109. Department of Defense Cyber Strategy. Defense USDos. Washington, DC: US Department of Defense, 2018.
110. Fischerkeller MP, Harknett RJ, Vićić J. The Limits of Deterrence and the Need for Persistence. In: AFs Brantley (ed). *The Cyber Deterrence Problem*. London: Rowman and Littlefield, 2020.
111. Fischerkeller MP, Harknett RJ. Deterrence is not a credible strategy for cyberspace. *Orbis* 2017;**61**:381–393.
112. Brantly AF. The cyber deterrence problem. In: *10th International Conference on Cyber Conflict (CyCon)*, Tallinn, Estonia: NATO CCDCOE, 2018, 31–54.
113. McKenzie TM. *Is Cyber Deterrence Possible?* USA: Air University Press, 2017, 1–33.
114. Nye JJ. Deterrence and Dissuasion in Cyberspace. *Int Secur* 2017;**41**:44–71.
115. Buchanan B. Cyber Deterrence Isn't MAD; It's Mosaic. *Geo J Int Aff* 2014;130–140.
116. Glaser C. *Deterrence of Cyber Attacks and U.S. National Security*. Washington, DC: Cyber Security Policy and Research Institute, 2011, 1–8.
117. Healey J. The implications of persistent (and permanent) engagement in cyberspace. *J Cybersecur* 2019;**5**.
118. Jensen B, Valeriano B. *Observations from Simulations and Surveys*. Washington, DC: The Atlantic Council, 2019.
119. Kreps S, Schneider J. Escalation firebreaks in the cyber, conventional, and nuclear domains: moving beyond effects-based logics. *J Cybersecur* 2019;**5**:1–11.