Cyber Survey Pre-Analysis Plan

Kathryn Hedgecock, Lauren Sukin, Leah Matchett

February 2021

1 Introduction

The advent of the widespread use of cyber operations challenges existing theories of international conflict. The non-physical nature of cyber operations may alter the dynamics of attacks, as well as retribution for them. In this research, we ask: Will governments and citizens respond to cyberattacks in similar ways as they respond to kinetic attacks? Or, are cyberattacks a unique means, prompting unique patterns of responses?

Because of their non-physical and clandestine nature, cyberattacks are inherently less visible than kinetic ones, and their attribution is often more difficult.[1] The scale of damage that can be caused by cyberattacks may also vary from kinetic attacks.[2] These features suggest that the public may be less likely to support or to push for a military response to cyberattacks. On the other hand, cyber operations are also a new form of weapon, which may increase the public's perception of danger when they are used.[3] Cyber operations often rely on public infrastructure (e.g. the internet),[4] therefore blurring the lines that distinguish combatants from non-combatants and military objects from civilian ones. This may increase civilian perceptions of their own vulnerability to cyber operations.[5] This would suggest that the public should be more eager to see their government respond to cyberattacks than they would to similar, physical attacks.

Previous research suggests that public support for retaliation to an armed attack may create pressure, or at least opportunities, for politicians to take military actions.[6] In this way, public attitudes in the wake of attacks and other events are critical to assess. While there is a robust literature assessing public attitudes in reaction to physical attacks—whether conventional or nuclear—few studies have addressed these questions in the cyber realm.

Building on the literatures of cyber security and public attitudes about military retaliation, we propose a conjoint experiment. This experiment will examine how the public's support for government retaliation to an attack varies with the nature of the attack (i.e. whether it is a physical attack or a cyberattack), the level of damage caused by an attack, the timing of the attack, and the certainty of attribution for the attack. This research also addresses questions about the threshold at which respondents feel that the damage caused by an attack should elicit a response, a central issue in debates on cyber defense and deterrence.[7]

[1] Erik M Mudrinich. "Cyber 3.0: The department of defense strategy for operating in cyberspace and the attribution problem". In: *AFL Rev.* 68 (2012), p. 167; Jon R. Lindsay. "Tipping the scales: the attribution problem and the feasibility of deterrence against cyberattack". In: *Journal of Cybersecurity* 1.1 (Sept. 1, 2015). Publisher: Oxford Academic, pp. 53–67. ISSN: 2057-2085. DOI: 10.1093/cybsec/tyv003. URL: http://academic.oup.com/cybersecurity/article/1/1/53/2354517 (visited on 08/27/2020).

[2] Thomas Rid. "Cyber war will not take place". In: *Journal of strategic studies* 35.1 (2012), pp. 5–32.

[3] Paul Slovic. *The feeling of risk: New perspectives on risk perception.* Routledge, 2010.

[4] James Andrew Lewis. *Assessing the risks of cyber terrorism, cyber war and other cyber threats.* Center for Strategic & International Studies Washington, DC, 2002.

[5] Neil C Rowe. "Challenges of civilian distinction in cyberwarfare". In: *Ethics and Policies for Cyber Operations.* Springer, 2017, pp. 33–48; Nadiya Kostyuk and Carly Wayne. "Communicating Cybersecurity: Citizen Risk Perception of Cyber Threats". In: (2019).

[6] Ifat Maoz and Clark McCauley. "Threat, dehumanization, and support for retaliatory aggressive policies in asymmetric conflict". In: *Journal of Conflict Resolution* 52.1 (2008), pp. 93–116; Michael Tomz, Jessica LP Weeks, and Keren Yarhi-Milo. "Public opinion and decisions about military force in democracies". In: *International Organization* 74.1 (2020), pp. 119–143.

[7] Sarah Kreps and Debak Das. "Warring from the virtual to the real: Assessing the public's threshold for war over cyber security". In: *Research & Politics* 4.2 (Apr. 1, 2017). Publisher: SAGE Publications Ltd. ISSN: 2053-1680. DOI: 10.1177/

Table 1: Experimental Conjoint Treatments

Variable	Treatment	Text	Effect
Type of Attack	Cyber	Cyberterrorist operatives have conducted a cyberattack	-
	Physical	Terrorist operatives have conducted a physical, in-person attack	+
Target of Attack	Archive	municipal archives / theft of large amounts of personally identifiable information	Baseline
	Bank	financial institutions / theft of large amounts of money	+
	Grid	electrical grid / large power outages	++
	Death	electrical grid / a large number of deaths	+++
Confidence of Attribution	High	Government officials have high confidence	+
	Low	Government officials have low confidence	-
Timing of Attack	Late	recently	+
	Early	more than a year ago	-

2 Experimental Design

We propose a conjoint survey experiment that tests how the characteristics of attacks on the United States affect public attitudes of retaliation in response to those attacks. Additionally, we seek to better understand the mechanism underpinning these attitudes.

2.1 Treatment and Independent Variables

The treatment is delivered in the form of a paragraph description of an attack on New York City. Respondents are asked to imagine a terrorist attack on New York City. The terrorist operatives are part of a group called "Red Square," sponsored by the Russian government.

Respondents are given information about the date of the attack (i.e. whether the attack was recent or whether it occurred more than a year ago), the mode of the attack (i.e. whether the attack was conducted physically or with cyber operations), the target and result of the attack, as well as the confidence level with which the attack has been attributed to the organization, Red Square (i.e. low or high confidence).

For example, the following text details the treatment group where respondents are assigned to a ***recent cyber*** attack involving the ***theft of information*** with ***low confidence*** in the attribution of the attack. Language in bold varies depending on the treatment.

> Imagine that **cyberterrorist** operatives have **recently** conducted a **cyber** attack against the United States. The operatives targeted the **municipal archives** in New York City, resulting in the **theft of large amounts of personally identifiable information.** Government officials have **low confidence** that the attack was perpetrated by an organization of **cyberterrorists**, called Red Square, that is sponsored by the Russian government.

Table 2.1 details the different treatment conditions for each of the four treatments (attack type, target, attribution, and timing.) The text associated with the treatments is also provided. In addition, the table details the expected effect of each treatment on respondents' support for retaliation in reaction to the attack.

The full text of the survey will be made available in the paper's appendix. Each of the four treatments constitutes an independent variable, which our survey will analyze. We can also conceptualize each respondent as being assigned to one of 32 possible combinations of the treatment conditions; as a robustness check, we will also analyze these combinations as separate independent variables.

2053168017715930. URL: https://doi.org/10.1177/2053168017715930 (visited on 08/27/2020); Sarah Kreps and Jacquelyn Schneider. "Escalation firebreaks in the cyber, conventional, and nuclear domains: moving beyond effects-based logics". In: *Journal of Cybersecurity* 5.1 (Jan. 1, 2019). Publisher: Oxford Academic. ISSN: 2057-2085. DOI: 10.1093/cybsec/tyz007. URL: http://academic.oup.com/cybersecurity/article/5/1/tyz007/5575971 (visited on 08/27/2020).

2.2 Dependent Variable

The main dependent variable of interest is public support for retaliation. Respondents are asked: "Should the U.S. military retaliate against the attack?" Respondents are asked to indicate yes or no as well as to rate the strength of their response. We will use a binary variable indicating respondents' support for retaliation, a scaled variable reflecting strong opposition to strong support for retaliation, and a binary version of the scaled variable comparing only strong opposition and strong support.

We will also analyze several additional, alternative dependent variables. For example, respondents are asked how they would most prefer the U.S. to respond to the attack and are given the following options:

- Conduct a physical attack against Russia
- Conduct a cyberattack against Russia
- Conduct a physical attack against Red Square
- Conduct a cyberattack against Red Square
- Introduce economic sanctions against Russia
- Publicly denounce the attack, but do nothing else

Each option will be treated as a binary variable. In addition, we will create a scale indicating the aggressiveness of respondents' preferred retaliatory strategy, from physical attack on Russia to public denouncement.

There are several other questions that respondents are asked that can serve as alternative dependent variables. First, respondents are asked how effective they believe a response would be. Respondents are also asked to rate their level of approval if the government responded to the attack by taking each of the above actions. In addition, they are asked if the U.S. should criminally prosecute either the organization that conducted the attack or Russian government officials involved in the sponsorship of that organization. They are also asked what kinds of attacks they would support, e.g. "boots on the ground," "unmanned aircraft," or the use of "low-yield nuclear weapons," among other choices. Respondents are asked whether the actors carrying out the incident intended to cause deaths. All of these alternative dependent variables can be analyzed in a binary or a scaled format.

In many cases, while these variables are important to understand as distinct dependent variables, they also may be mechanisms affecting the relationship between the type of the attack (e.g. Cyber of Physical) and support for retaliation to the attack.

For example, while we may want to know what features of an attack cause respondents to think the attack was intended to kill Americans (Intention), we may also view this variable not as a separate dependent variable, but as a mechanism through which respondents may react differently to cyber and physical attacks. That is, it is possible that the deaths that result from cyberattacks are seen as 'accidental', because of perceptions that cyber operations cause minimal damage. We might also expect respondents to be more likely to retaliate to attacks they perceive as intending to kill. Thus, one reason to expect higher rates of retaliation in response to physical attacks than cyberattacks may be because of perceptions about the attacks' intentions.

2.3 Sample

We will use the Lucid Survey company to collect a nationally representative sample of approximately 3000 respondents in the United States.

2.4 Exclusion Criteria

Subjects can be excluded from the sample if they meet one or more of the following criteria:

1. They do not consent to take the survey.
2. They are identified by the survey company or researchers as 'speeders,' i.e. if they take the survey at an unreasonable pace.
3. They find the situation identified in the experiment to be "very unbelievable."
4. They fail either of two attention check questions, one of which asks a respondent to answer 'Neither Agree nor Disagree,' and the other asks whether 1910 occurred before 1920.
5. They fail the main manipulation check question, asking them whether the main treatment took the form of a cyber or a physical attack.
6. They fail secondary manipulation checks asking about the target, timing, and confidence level of attribution for the attack.

The authors will assess how dropping or including the different groups of respondents listed above affects the survey's results.

3 Hypotheses

Our experiment is primarily designed to identify how responses to physical and cyberattacks differ. To this end, we develop several hypotheses, which are detailed below.

We test these hypotheses in different ways. Some hypotheses can be tested by the inclusion of the relevant variables in regression analyses, described in more detail in Section 4. We may use various data visualization approaches to illustrate the effects of certain mechanisms. In other cases, we will use two-stage regressions to test the effects of post-treatment mechanisms (e.g. by regressing the treatment on the mechanism and the mechanism on the dependent variable.)

In these two-stage regression tests, if both coefficients are positive, then we can conclude the treatment is leading to the dependent variable through the mechanism. If the first coefficient is not significant and the second coefficient is significant, then the variable is not a mechanism and can be treated as a control. If the first coefficient is significant and the second is not, then the mechanism can be treated as an alternative dependent variable; the treatment would affect the mechanism but the mechanism would not effect the dependent variable. If both coefficients are not significant, then the variable is irrelevant to the model and can be dropped.

The following sections detail several hypotheses that we will test, information about how we will operationalize these hypotheses, as well as information about how we will test certain hypotheses.

3.1 Type of Attack

Our survey is primarily designed to distinguish how the *type* of an attack, e.g. whether the attack is conducted through cyber or kinetic means, affects attitudes in response to the attack. Informed by existing survey research, we argue that:

H1 **The public will be more likely to support retaliation against physical attacks than cyber attacks.**

3.2 Magnitude of Effects

The magnitude of the effects of the attack is an important feature to consider; we expect that larger attacks will elicit greater support for retaliation to those attacks.

H2 **As the magnitude of the attack increases, public support for retaliation will increase.**

Some work suggests that cyberattacks do not cause as much damage as as physical attacks. Empirically, cyber operations lead to a low threshold of damage and there is a preponderance of espionage operations in the cyber domain. As a result, respondents may be sensitized to this empirical reality and therefore less likely to respond to the magnitude of the attack when it occurs through cyber, rather than kinetic, means. We will therefore test the following hypothesis:

H3 **The magnitude of the attack will have a smaller effect for attacks conducted through cyber, rather than kinetic, means.**

We will test this hypothesis both directly (by examining the presence or absence of an interaction between an attack being cyber in nature and the magnitude of effect), and through an investigation of a survey question about the experimental scenario's plausibility. That is, we will assess whether large-scale physical attacks are seen as more plausible than large-scale cyberattacks.

3.3 Attribution Certainty and Timeline

Another widely discussed aspect in which cyberattacks may differ from physical attack is the certainty and speed of attribution. We expect that attribution certainty should have a positive effect on support for retaliation:

H4 **As the attribution certainty of an attack increases, public support for retaliation will increase.**

We also anticipate that this effect should vary between cyber and physical attacks. Because cyberattacks are seen as more uncertain, attribution certainty should have a greater effect on respondents' willingness to retaliate than with non-cyberattacks.

H5 **The attribution certainty of an attack will have a greater effect for attacks conducted through cyber, rather than kinetic, means.**

We also expect that the more recent an attack is, the more support there will be for retaliation. That is:

H6 **As the timing of an attack becomes more recent, public support for retaliation will increase.**

Like with the certainty of attribution, we expect that the timing of attribution should have a greater effect on respondents' willingness to support retaliation in response to cyber, rather than physical attacks.

H7 **The timing of an attack will have a greater effect for attacks conducted through cyber, rather than kinetic, means.**

While we directly vary the timing of an attack and the certainty of attribution, we also assess how important respondents think it is to respond promptly to an attack. We expect respondents to value prompt responses more when attacks are kinetic, rather than cyber, when attacks have larger effects, and when attacks are more recent. We expect that, when timely retaliation is important to respondents, they will be more likely to support retaliation:

H8 **As timely retaliation to an attack becomes more important, public support for retaliation will increase.**

3.4 Novelty of Cyber

Another factor that's been associated with cyberattacks is their novelty. Cyber is a relatively new domain. Novelty could make retaliation more or less likely for a number of reasons, described in this section.

We measure perceptions of novelty through the knowledge each respondent has on cyberattacks. Specifically, we ask respondents 'How much would you say that you know about cyberterrorism?' We also ask respondents if they believe the attack intended to kill Americans; respondents with higher knowledge will be more likely to know that cyberattacks rarely cause deaths.

Respondents with higher knowledge of cyber operations should view cyber operations as relatively low-risk. They are more likely to know that cyber operations rarely cause significant damage and more likely to know that cyber operations rarely cause deaths. This would suggest the following hypothesis:

H9 **As individuals have more knowledge about cyber operations, they will be less likely to support retaliation to cyberattacks.**

However, in our survey, we control the actual risk associated with a cyberattack by varying the magnitude of the attack's effect, so we expect that, controlling for the attack's impact, novelty should either have no effect or should a positive effect on support for retaliation, as the decreased risk perception associated with knowledge about cyberattacks should cause individuals to discount the possibility of escalation.

A positive association between knowledge about cyber operations and support for retaliation could also be caused by the fact that high-knowledge individuals think more about cyber risks, i.e. these risks have a higher salience for such individuals, while low-knowledge individuals do not consider these risks regularly.

Thus we will also test the following hypothesis:

H10 **As individuals have more knowledge about cyber operations, they will be more likely to support retaliation to cyberattacks.**

Having more information about cyberattacks could contribute to respondents' perceptions of their own vulnerability to cyberattacks, which may, in turn, make respondents more likely to support retaliation. We expect vulnerability to relate to respondents' support for retaliation in both the *cyber* and *physical* attack conditions, as follows:

H11 **As individuals have a greater perception of vulnerability to cyber (terrorist) operations, they will be more likely to support retaliation to cyber (physical) attacks.**

3.5 Escalation Risk

Another factor affecting the willingness to respond to an attack is the perceived likelihood of escalation. While each of the above factors is likely to influence the respondent's perceived escalation risk, there are also a host of other factors which affect the likelihood of escalation in both the physical and cyber domains. Because we cannot measure each of these, we instead directly measure respondents' perception of the escalation risk associated with retaliation. We expect that this should be negatively associated with willingness to retaliate. Stated formally:

H12 **As perception of the risks of escalation increase, support for retaliation will decrease.**

We measure escalation risk three ways: 1) by asking respondents how effective they think retaliation will be at preventing future attacks, 2) by asking if respondents anticipate more attacks against the U.S. if no retaliation occurred, and 3) by asking: "Imagine that the US decided to retaliate against the attack by targeting Red Square. In response to the US military retaliation, Red Square carries out an attack against the US. This new attack is similar to the first attack. Knowing this, do you think that the US military should have retaliated against the original attack?"

Because we expect that retaliation to physical attacks will be perceived as more likely to escalate than retaliation to cyberattacks, we also predict that expectations about escalation will encourage more support for retaliation to physical attacks than to cyberattacks.

While we focus on military escalation, retaliation or lack thereof can escalate in other ways as well. We assess respondents' reasoning for supporting retaliation by looking at how they evaluate the consequences of their choices. We ask respondents to 'Imagine the U.S. did not respond to the attack. How likely or unlikely are the following outcomes?' We will examine answers to each of the following statements:

- Recruitment: 'Red Square, and organizations like it, would gain more members'
- Economy: 'It would be better for the U.S. economy.'
- Reputation: 'It would help the United States' reputation in the world.'
- GlobalNorms: 'More countries would act aggressively in international politics.'

We will assess if respondents perceive the consequences of inaction to an attack as different when the attack uses cyber, as compared to kinetic, means. We expect that inaction to physical attacks will be perceived as more costly than inaction to cyberattacks. We expect that the perceived costs of inaction will be tied to support for retaliation, as follows:

H13 **As the perceived consequences of inaction in response to an attack increase, support for retaliation will increase.**

3.6 Norms

Cyber weapons are a relatively recent development in international relations. It is not yet clear what norms govern their behavior or how these norms might work. However, we might expect that different norms govern cyberattacks than govern physical attacks.

H14 **Differing norms around cyber and physical attacks will affect support for retaliation in different ways.**

We survey respondents on their attitudes about a variety of normative statements, including: "An eye for an eye is never enough"; "It is always unacceptable to steal private information or money"; "Anyone that kills Americans deserves to be punished"; "It is always morally wrong to destroy private property or public infrastructure"; and "Anyone that steals private information or money will use it to cause more harm later on." We generally expect that agreement with the ideas that "An eye for an eye is never enough" and that "Anyone that kills Americans deserves to be punished" will be associated with increased support for retaliation to attacks. Other norms may apply only to certain types of attacks (e.g. a norm against stealing may make retaliation more likely in the Bank and Archive treatment conditions.) Some norms (such as norms about globalism or the value of international law) may be associated with decreased support for retaliation to attacks.

3.7 Offensive Advantage

Whether or not a respondent believes retaliation will deter future attacks is also important. The perception that retaliation deters should increase respondents' support for retaliation, as follows:

H15 **As the likelihood of retaliation deterring future attacks increases, public support for retaliation will increase.**

Importantly, respondents may believe that the use of cyber weapons is difficult or impossible to deter. That is, the 'offensive advantage' associated with these weapons may prevent effective countering of their use. In this way, the ability to deter future attacks is a distinguishing feature of cyber and kinetic attacks.

H16 **The perceived likelihood of retaliation deterring future attacks will be lower for attacks conducted through cyber, rather than kinetic, means.**

We measure this with the following question: "Imagine the U.S. did not respond to the attack. How likely or unlikely are the following outcomes....There would be more attacks against the U.S." This perception may be higher in individuals who have greater self-professed knowledge about cyber operations.

We will measure directly whether the perceived likelihood of retaliation deterring future attacks is lower when the attack is conducted through cyber, rather than kinetic means. If this is the case, as we predict, (and if the likelihood of successful deterrence increases support for retaliation), then we can conclude that one reason why support for retaliation to physical attacks is higher than support for retaliation to cyberattacks is because of the perceived difference in deterrence in the two domains.

4 Analysis Plan

4.1 Initial Analysis

Our initial analysis will first make sure every variable is scaled in the intuitive direction. For example, variables measuring threat perception on a 5 point scale should run from 1 (low threat perception) to 5 (high threat perception.)

We will then create indices and binary versions of relevant variables. The standard procedure for creating a binary variable from scaled responses is to have "very" and "somewhat" in comparison to "neither," "not somewhat," and "not very." In some cases, alternate approaches may be more accurate. The researchers will provide justification for any alternate approach used.

We will examine summary statistics for each variable in the dataset, and we will perform a balance test on each axis of treatment. If necessary, we will create a separate version of the dataset that re-weights the data to achieve appropriate balance across treatments. We will assess relevant summary statistics, create visualizations of, and determine correlations between each of the main independent variables (see Table 4.1). We will then assess and visualize the size and significance of the effects of each variable on the main and alternate dependent variables.

Table 2: Expected Effects of Main Models

Treatment	**Expected Effect**
Cyber	-
Physical	+
Archive	Baseline
Bank (1)	+
Grid (2)	++
Death (3)	+++
High	+
Low	-
Early	-
Late	+

4.2 Main Models

We will run bivariate and multivariable logit, probit, and linear probability models (LPMs) to test effects of the treatments on the dependent variables.With each regression model, we will evaluate fit. We will use alternate model specifications, such as bootstrapping or the inclusion of polynomial terms, if issues with fit arise. We will visualize the results of each model in various ways, including through coefficient plots.

4.3 Treatment Interactions

As a robustness check, we will re-run the main models with interactions between attack type (e.g. Cyber or Physical) and the other treatments. We expect few significant effects. The treatment interactions that we will assess are described here:

- Type x Attribution
- Type x Scale
- Type x Timing
- Type x Scale x Attribution
- Type x Scale x Timing
- Type x Timing x Attribution
- Type x Scale x Attribution x Timing

We anticipate a possible significant interaction between Type and Attribution as well as between Type and Scale, as described in our hypotheses.

4.4 Analysis of Open-Ended Questions

To further understand our survey responses, we will compile the text-based responses to two open-ended questions: "WhyNot" and "Why." These questions ask respondents to describe the reasoning behind their support for or opposition to retaliation. We will use a random sample and qualitative open-coding to develop categorizations and associated dictionaries; we will apply these dictionaries to all responses. We will analyze the content of the open-ended responses qualitatively, as well as provide summary statistics for the categories identified through open-coding and through the application of dictionaries. We will assess the distribution of responses for different groups of respondents.(For example, we may identify the types of responses associated with pro-retaliation respondents who received the Cyber treatment condition.) We will validate categorization created through dictionary-based methods by hand-coding a random sample of responses. Additionally, we will apply a topic model to validate our open-coding approach, and we will run a fictitious prediction to find most distinct words associated with certain groups of responses. (For example, we may identify distinct words associated with the pro-retaliation responses of respondents who received the Cyber treatment condition.)

4.5 Controls

Table 4.5 depicts control variables that may be included in the analysis. The first column identifies the control. If the control corresponds to a singular question, the label for that question is given in parentheses. Details on the text of each question are provided in the appendix. For control variables that are indexes, no name in parentheses is given. For many of these variables, we predict a directional effect on support for retaliation. For some variables, we also expect an interaction with certain treatments.

4.5.1 Indexes

This section outlines how certain indexes will be constructed.

- Vulnerability to Cyber is constructed as an index of four questions:
  - How likely do you think it is that there will be a cyberterrorist attack targeting the United States next year? (CybLikelyInf)

Table 3: Possible Controls and Predicted Effects

Control	**Predicted Effect:**	**Prediction Interaction**
Knowledge about Cyber (OftenCyber)		
Knowledge about Terrorism (OftenTerror)		
Threat from Cyber (Threats_C)	+	Cyber (+)
Threat from Russia (Threats_R)	+	
Threat from Terrorism (Threats_T)	+	
Vulnerability to Cyber	+	Cyber (+)
Irrationality of Terrorists	+	Physical (+)
Irrationality of Cyberterrorists	+	Cyber (+)
Retaliation	+	
Nationalism (Nationalism)	+	
Importance of International Law (ILaw)	-	
Globalism	-	
Age (Age)		
Education (Education)	-	
Gender (Female)	-	
Region (Region)		
Hispanic (Hispanic)		
Race (Race)		
PartyID- Republican Party (PID)	+	
Kids (Kids)	-	
Military Service (MilService)		
Household Income (HHI)	-	
Conservative	+	
US Citizen (Citizen)		
Support for Disaster Relief (DisasterReliefFunding)		Death (+) ; Bank (+)
Support for Disaster Preparedness (DisasterPrepFunding)		
Support for Infrastructure Funding (InfFunding)		Grid (+); Grid+Death (+)
Support for Financial Institutions Funding (FinanInstFunding)		Bank (+)
Trust in Government (Trust)		
Concern over Hospital Cost (HosMoney)		Bank (+)
Concern over Hospital Closure (HosClosed)		Grid (+)
Concern over Hospital Deaths (HosDeaths)		Death (+)
Concern over Hospital Future (HosMoreEarthquakes)		
Concern over Hospital Records (HosRecords)		Archive(+)

  - How likely do you think it is that you or someone you know well will be affected by a cyberterrorist attack next year? (CybLikelyYou)
  - Have you or someone you know ever been the victim of a cyberterrorist attack? (CybExposure)
  - How concerned are you about protecting personally identifiable information, such as your address and social security number? (PersonalVulnerable)

- Perceptions of Actors The variables for the Irrationality of Terrorists and Irrationality of Cyberterrorists are constructed from indexes of the personality traits respondents associated certain actors. Respondents indicating that an actor is 'radical,' 'evil,' 'ideological,' and 'dangerous' result in a higher irrationality score, while indicating that an actor is 'predictable' and 'rational' result in a lower score. The index derives from respondents' responses to the following questions: CyberPerceptions and TerrorPerceptions. Respondents were also asked about their opinions on Russian military officials (RussianPerceptions) as a baseline from which to compare their attitudes about terrorists and cyberterrorists.

- Retaliation is an average of a respondent's support for the death penalty and for waterboarding as asked in the questions labelled 'DeathPenalty' and 'Waterboarding.'

- Globalism is an index constructed from respondents attitudes about three statements:
  - The United States needs to play an active role in solving conflicts around the world. (WorldPoliceman)
  - Generally speaking, the United States can trust other countries. (IntlTrust)
  - The use of military force only makes problems worse. (Dove)

- Conservative indicates a respondents' ideology on a scale from liberal to conservative. Respondents were randomly assigned a question that asked them either to select their ideology from a scale running from liberal to conservative (Conservative) or a scale running from conservative to liberal (Liberal). We will combine these responses into a single variable.

5 Other Analyses

This section details two additional research areas that we will explore with this survey data.

5.1 Cyber Policy

We will also examine how respondents' answers map onto other key policy areas. In particular, we will examine how the treatments affect respondents' support for policy solutions to challenges in the cyber domain.

We ask each respondent: "Imagine that the United Nations wants to limit the use of cyber weapons. Which of the following policies do you most support?" (Ban) We hope to understand how the controls and mechanisms we outline co-vary with support for international bans and other policies as a means of solving problems associated with the use of cyber weapons. This is mainly an exploratory exercise. However, we expect that respondents primed to consider cyberattacks will be more likely to support strict measures restricting cyber weapons. We also predict that the severity of the cyberattack described in the treatment (e.g. more recent, more damaging, and more clearly attributed attacks) will be associated with higher levels of support for strict policies against the development and proliferation of cyber weapons.

5.2 SolarWinds Natural Experiment

While our survey was fielding, news of the SolarWinds Hack broke in U.S. media. Upon hearing of this, we added three questions to the survey to capture how this event (one of the largest cyber operations in history) affected responses to our fictional cyber (and physical attacks).

To determine whether respondents had heard of the event, we asked respondents how much they had heard about the event, and provided them with a series of true/false questions about the event to assess their knowledge. We will create an index of the respondent's knowledge based on how many of these are answered correctly.

We also asked respondents how they would 'most prefer' the U.S. respond to the attack, using the same categories we use for our experiment: "Conduct a physical attack against Russia," "Conduct a cyberattack against Russia," "Introduce economic sanctions against Russia," "Publicly denounce the attack, but do nothing else", and "Do not acknowledge the attack." We will analyze respondents' preferred responses to the SolarWinds attack, using the same operationalizations discussed for the parallel question in the main experiment.

We plan to analyze our results on public attitudes about SolarWinds in two ways:

1. SolarWinds as a Salience Instrument: We will compare support for retaliation to our fictional situation among similar respondents who have and have not heard of the SolarWinds hack. We expect that respondents who have heard of the attack will view cyberattacks as more salient and more threatening, and they will therefore be more supportive of a retaliation against cyberattacks in our fictional vignette.

2. SolarWinds as a Baseline: It may be that respondents are more willing to support the use of force in our fictional survey experiment than in 'real life.' We will compare the same respondent's willingness to respond to the conjoint vignette and the real-world example.

We will also examine how support for retaliation to SolarWinds varies with respect to demographics as well as other controls and mechanisms identified in our survey.

References

Kostyuk, Nadiya and Carly Wayne. "Communicating Cybersecurity: Citizen Risk Perception of Cyber Threats". In: (2019).

Kreps, Sarah and Debak Das. "Warring from the virtual to the real: Assessing the public's threshold for war over cyber security". In: *Research & Politics* 4.2 (Apr. 1, 2017). Publisher: SAGE Publications Ltd. ISSN: 2053-1680. DOI: 10.1177/2053168017715930. URL: https://doi.org/10.1177/2053168017715930 (visited on 08/27/2020).

Kreps, Sarah and Jacquelyn Schneider. "Escalation firebreaks in the cyber, conventional, and nuclear domains: moving beyond effects-based logics". In: *Journal of Cybersecurity* 5.1 (Jan. 1, 2019). Publisher: Oxford Academic. ISSN: 2057-2085. DOI: 10.1093/cybsec/tyz007. URL: http://academic.oup.com/cybersecurity/article/5/1/tyz007/5575971 (visited on 08/27/2020).

Lewis, James Andrew. *Assessing the risks of cyber terrorism, cyber war and other cyber threats*. Center for Strategic & International Studies Washington, DC, 2002.

Lindsay, Jon R. "Tipping the scales: the attribution problem and the feasibility of deterrence against cyberattack". In: *Journal of Cybersecurity* 1.1 (Sept. 1, 2015). Publisher: Oxford Academic, pp. 53–67. ISSN: 2057-2085. DOI: 10.1093/cybsec/tyv003. URL: http://academic.oup.com/cybersecurity/article/1/1/53/2354517 (visited on 08/27/2020).

Maoz, Ifat and Clark McCauley. "Threat, dehumanization, and support for retaliatory aggressive policies in asymmetric conflict". In: *Journal of Conflict Resolution* 52.1 (2008), pp. 93–116.

Mudrinich, Erik M. "Cyber 3.0: The department of defense strategy for operating in cyberspace and the attribution problem". In: *AFL Rev.* 68 (2012), p. 167.

Rid, Thomas. "Cyber war will not take place". In: *Journal of strategic studies* 35.1 (2012), pp. 5–32.

Rowe, Neil C. "Challenges of civilian distinction in cyberwarfare". In: *Ethics and Policies for Cyber Operations*. Springer, 2017, pp. 33–48.

Slovic, Paul. *The feeling of risk: New perspectives on risk perception*. Routledge, 2010.

Tomz, Michael, Jessica LP Weeks, and Keren Yarhi-Milo. "Public opinion and decisions about military force in democracies". In: *International Organization* 74.1 (2020), pp. 119–143.