Seeing is Believing: Quantifying and Visualizing Offensive Cyber Operations Risk

Major Michael Klipstein, Ph.D.

ABSTRACT

This paper presents an integration of decision-maker preferences, quantitative risk analysis, and simulation modeling to aid commanders in choosing a course of action (COA) for conducting offensive cyber operations (OCO). It incorporates information from subject matter experts (SMEs) to parameterize a simulation model which provides decision support to mission planners when evaluating different COAs. The methodology is exercised and evaluated by cyberwarfare practitioners. The research findings demonstrate its value for increasing the ability of inexperienced personnel to make COA selections on par with experienced personnel, providing greater perceived understanding of risk defined as meeting the constraints of both cost and effectiveness, mitigating confusion or ambiguity resulting from subjective terms, and providing greater consensus of COA selection among practitioners in the aggregate. The advantages of this approach are significant as it produces a portrait of each COA that reveals the effect of the uncertainties that the SMEs admit pertaining to each of their outcome estimates. Given the value functions and trade-off weights of the commander, these translate into a meaningful portrayal of the risk to the decision maker in each COA.

INTRODUCTION

Military commanders and their staff below the national command level are ill-prepared to assess risks for conducting offensive cyber operations (OCO) (Department of Defense, 2017a, Department of Defense, 2017b). The man-made cyber domain exhibits four unique traits that differentiate itself from the traditional military domains. First, a lack of permanence exists for objects within the domain as they appear, disappear, or change at rapid speed. Next, the domain lacks measures of effectiveness for operations. The view of the virtualized battlefield is limited, and an accurate feedback loop for actions and

MAJ Michael Klipstein, Ph.D. has been involved with cyberspace operations since his 2010 assignment to USCYBERCOM. Following this, Michael worked for two years in Tailored Access Operations (TAO) as part of the National Security Agency (NSA). Michael then led the Army's National Mission Team; the offensive cyber team tasked to conduct operations to protect the critical infrastructure and key resources of the nation and to answer intelligence requirements as assigned by the President. Following success with the NSA, he created and trained two national level cyber protection teams charged with protecting the DoD's networks from nation-state actors. In 2014, Michael attended the Naval Postgraduate School in California to attain his Ph.D. in Information Sciences. Currently, Dr. Klipstein is assigned as a senior research scientist and the Chief of Outreach at the Army Cyber Institute at West Point. He assists the Joint Staff and ARCYBER with research problems facing the international cyber community.

effects does not exist. Third, actions within this domain occur at computational speed, or near the speed of light. The last unique characteristic is the ability of an attacker to remain anonymous or even to masquerade as another entity. Proxy servers and The Onion Router (TOR) make attribution of an attack difficult, if not impossible (Kallberg and Cook 2017). Further compounding the attribution problem, cyber operations are characterized by a lack of detection by the targets for intelligence gathering and destructive effects until it is too late to defend themselves.

Commanders are guided by doctrine drawn from personal education, experience, or historical context to create an analogy for the current environment (Department of Defense 2011a; Department of the Army 2012). Traditionally, operational commanders come from a combat arms backgrounds at higher levels. Examples of these backgrounds include Infantry, Armor, fighter pilots, naval surface and subsurface fleet commanders. These commanders traditionally lack a computer science or telecommunications education or experience in Signals Intelligence (SIGINT) gathering and analysis. Additionally, commanders are reliant on subjective risk measures that are foundationally based on the experience and education of the commander to assess the operational risks.

Therefore, it is reasonable that commanders and their respective staffs are unable to adequately assess the risks involved, particularly with second and third order effects, for OCO. For example, if an OCO capability is used against a target, several considerations must be considered. First, the capability cannot be used elsewhere globally as an anti-virus company will likely see it and create a signature for it. Next, the target will investigate and remediate the vulnerability used in the OCO. Compounding this consideration is the potential for the vulnerability to vanish globally through remediation. Third, the OCO capability could

potentially be used by the adversary against US targets. Unlike bombs and missiles, OCO capabilities can be reassembled from a forensic investigation and reused. This paper asserts that a new risk assessment technique is needed, one based on quantitative measures that account for the commander's desired operational end-state.

For this research, offensive operations consist of both OCO and intelligence gathering operations. The latter has been at various times identified as computer network exploitation (CNE); intelligence, surveillance, and reconnaissance (ISR), and; surveillance and reconnaissance (S&R). The reason for this deviation from current doctrine is that, from the adversary perspective, attack and intelligence operations look similar, if not the same, until the point of an attack payload is released for effect. This deviation from doctrine also forces consideration on the potential ramifications of detection, attribution, and compromise from adversary actions regardless of the operation.

Current cyber operations

In 2010, U.S. Cyber Command (USCYBERCOM) was established at Ft. Meade, MD, and collocated with the National Security Agency (NSA). Personnel within USCYBERCOM are mostly military with government civilians, and some contractors. Military personnel make up the preponderance of the planning teams for the organization and typically are assigned to USCYBERCOM for three years before returning to their military service career field. It is not unusual for military personnel to be unable to articulate the mechanics of how the Internet works before arriving at USCYBERCOM. However, these same military personnel are on planning teams that support national level interests and support the geographic military combatant commands (CCMD). Currently, CCMDs are responsible for all military operations and therefore, the security of portions of the planet. In February 2014, then Chief of Staff of the Army, GEN Ray Odierno stated that: "We have to be able to do that and potentially be able to conduct tactical offensive cyber operations, because I think in the future, that'll be another way for us to maneuver in the battlespace that we might be in. So I think we have to develop those techniques" (Council on Foreign Relations 2014). However, if the personnel at USCYBERCOM do not understand the risks involved with OCO, how can the CCMDs be expected to make a meaningful assessment of the risks?

Risk assessment methods

Current risk assessment and decision-making for OCO consists of a combination of subjective measurements and other cognitive mechanisms are used in daily routine or simple tasks. However, as complexity rises, or experience diminishes, these cognitive mechanisms begin to fail and initiate other problems. Examples of these mechanisms are group dynamics, heuristics, bias, affect, and overestimation or underestimation of risk.

The systems for risk analysis such as ones used in the Department of Defense (DoD) require extensive experience and knowledge of the risks and consequences involved. The DoD

explains this requirement in the Joint Operations manual that: "Commanders compare similarities of the existing situation with their own experiences or history to distinguish unique features and then tailor innovative and adaptive solutions to each situation." (2017b, II-4, c). Because commanders and their staffs lack experience, education, and expertise in cyberspace operations, these decision-makers are incapable of assessing the risks involved in OCO. Cyber operations have the potential to be considered mixed gambles (Holt and Laury 2002; de Langhe and Puntoni 2015; Kahneman 2013; Kahneman and Lovallo 1993), where both gains and losses may occur simultaneously. This is in contrast with single-domain gambles where only gains or losses may occur (de Langhe and Puntoni 2015).

No existing doctrine for commanding and controlling military operations, much less cyber operations, include the application of multi-criteria decision making for weighing and assessing risks and rewards. Thus, commanders and their staffs are incapable of trading off between reward in operations and the associated costs. This is more vital in cyberspace operations as a superbly executed operation may still not yield the desired end-state of the commander as they will lack perfect knowledge of a target configuration or hardware. The DoD uses fourteen different systems to analyze and assess operational risk (Army War College, personal communication, February 2016). Of these, only four potential systems for assessing risk in cyberwarfare exist: one each from the Army, Navy, Air Force, and Joint doctrine.

The remaining four risk assessment methodologies use subjective terms to convey risk. These systems use terms such as "high," "moderate," or "low" to convey an understanding of the risks and to describe the severity of the risk (Department of the Army 2013; Broder and Tucker 2012). These terms have no clearly defined meaning or context. Often, the definitions of these terms include qualitative descriptions such as "unlikely to occur," "severe impact," and "highly likely" that offer no discrete boundaries to divide and define the areas. Different people may observe the same data and arrive at different conclusions. Consistent metrics do not exist for these measures, which makes this situation even more inexplicable. These risk analysis methodologies are qualitative and ambiguous at best.

Qualitative scales lack standardization and meaning. Two people with different experience levels and backgrounds would likey have different interpretations of what is "severe" or "high impact" (Bennett 2000). This is because non-numeric descriptions lead to different interpretations of data. Budescu, Broomell, and Por (2009) found participants even applied their subjective meaning to the nominal scales, even though a quantified definition existed. However, these subjective meanings were based on the heuristics of each person. Another example of these heuristics at play is the decision maker mentally assigning values, numbers, or probabilities when none exist (Ellsberg 1961). These heuristics consider the bias, past experiences, and cognitive understanding of each person. Therefore, it is impossible for a group of disparate people from different backgrounds and experiences to arrive at the same definition of what constitutes for each level of risk.

Two other flaws of these qualitative systems are range compression and the presumption of regular intervals. In range compression, if numbers are assigned to risk assessments using as an example, a 1–5 or a 1–10 scale, a small incremental movement can have a large impact on the alternatives or consequences. As the scale range decreases, the magnitude of impact conversely increases, that is, if the numbers and the corresponding meanings have regular intervals. With the presumption of regular intervals between levels, a 1–2-3-4-5 scale implies that a 4 is twice as good/bad as a 2; this is not necessarily true (Hubbard 2009; Savage 2012). Alternate methods of overcoming these challenges present their own dilemmas. For example, the Analytical Hierarchy Process (AHP) is often used in multi-criteria decision making. However, AHP suffers from multiple criticisms for use in this manner such as producing arbitrary results (Dyer, 1990) along with a lack of standardized scales for decision maker preferences and an assumption of criteria independence (e.g., no correlation) (Ishizaka, 2009). These flaws make this method substandard for multiple reasons but most importantly, since three of the objectives in the cyber operations hierarchy are dependent on a fourth criterion. The objective hierarchy used in this paper will be discussed in a later section. Since different backgrounds and experiences create different heuristics used to assess the severity of a situation, the current risk assessment systems are inadequate. These inadequate risk assessment systems coupled with the cognitive pitfalls create potential failure when used in new operations where the decision maker and support staff lack the experience and education in understanding the risks and consequences involved.

Cognitive mechanisms

Group dynamics are the interactions of a group setting where one person oversees a decision, but others inform the decision. Two potential problems occur in this situation. First, a strong personality will overrun people that disagree with an opinion. This is a form of confirmation bias. Another potential group dynamic problem is that subordinates will sometimes withhold critical information and defer to the leader even in an emergency. This phenomenon has been identified in multiple workplace environments, to include investigations using flight data recorders of crashed airplanes (Asch 1956, 1955; Gilovich 1991; Garvin and Roberto 2001; de Dreu, Nijstad, and van Knippenberg 2008; Foushee 1982).

Heuristics are the mental rules of thumb and analogies used in everyday life to make sense of new information or to fill in the gaps when information is missing. However, heuristics requires comparable base knowledge for comparison (Kahneman 2003; Dowd, Petrocelli, and Wood 2014; Kane and Webster 2013; Davis, Kulick, and Egner 2005; Grifffin et al. 2002). If a commander perceives the risk of offensive cyber operations as the same as the risk involved in kinetic operations by tanks, aircraft, or ships, this is a flawed comparison. Cyber operations have the potential of the adversary being within your sanctuary to witness and counter your operations on commencement. This aspect does not exist typically in kinetic operations.

Bias is the subjective perception lens that the individual interprets information. Each person uses multiple biases daily. Biases are formed from experiences, education, assumptions, prejudices, and correctly or incorrectly, our observations. Biases are important to consider when data is interpreted to become information. However, multiple people viewing the same data can arrive at different interpretations and contrasting versions of the same information (Kahneman and Tversky 1984; Kahneman 2013, 2003, Tversky and Kahneman 1981, 1974; Davis, Kulick, and Egner 2005; Milkman, Chugh, and Bazerman 2009; Heilbronner, Hayden, and Platt 2010; Dowd, Petrocelli, and Wood 2014; Kane and Webster 2013).

Affect refers to emotions or feelings that sway the judgment of the decision-maker. Examples of such emotions or feelings are fear, anger, surprise, or dread and have a personal value of "goodness" or "badness" (Clore, Gerald L & Huntsinger, Jeffery R, 2007). Cognitive psychology research illustrates how angry people make more aggressive and risk-seeking decisions while fearful or unsure decision makers are more risk-averse. This implies that as decision-makers may make choices that otherwise would be different in other circumstances (Arceneau 2012; Girodo 2007; Kahneman and Tversky 1979; Buelow and Suhr 2013; Bruyneel et al. 2009; Figner et al. 2009; Weber and Chapman 2005; Kahneman and Lovallo 1993; Nygren et al. 1996).

Decision-makers may overestimate risk or be overconfident in the circumstances. A popular example of this phenomenon in research are the people who habitually purchase lottery tickets, but not flood insurance while living in a flood-prone area (Davis, Kulick, and Egner 2005; Heilbronner, Hayden, and Platt 2010; Kahneman and Tversky 1984; Ludvig, Madan, and Spetch 2013; Tversky and Kahneman 1974). Kahneman and Lovallo (1993) describe how individuals manifest overconfidence in themselves when assessing the risk associated with multiple choices. In their study, participants assessed that they were correct approximately 99% of the time when the success rate hovered around 80%. Part of this discrepancy stemmed from optimism.

Operational risks in offensive cyber operations

In military operations, as in the public sector, risk minimization is required. To meet this requirement, the problem and solution set must be optimized to maximize the reduction of risk. Risk management is the process of incorporating the assessment and reduction of risk into decision making. Effective risk management requires the identification of the attributes of concern for the commander and gauging success or failure of each alternative. In OCO, two overarching objectives exist: Maximizing Effectiveness and Minimizing Costs. Effectiveness is a function of the following concerns: Maximizing Intelligence Gained, Maximizing Damage Inflicted, Minimizing Detection of Operations, Minimizing Attribution Given That Detection Occurred, and Minimizing Compromise Given That Detection Occurred (Klipstein 2017). Please refer to Figure 1. Each of these objectives can be further broken down into sub-objectives. However, only the first level of the objective was used in this research.

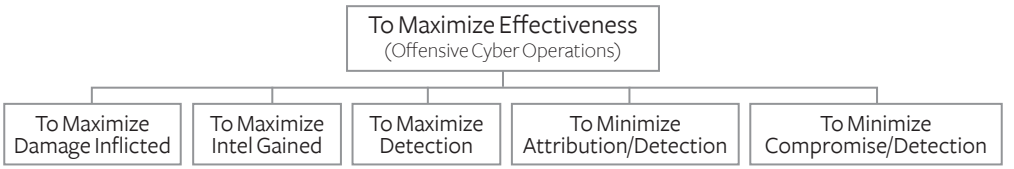

Figure 1. Objective Hierarchy for Maximizing Effectiveness in Offensive Cyber Operations

Maximizing Intelligence Gained and Maximizing Damage Inflicted are self-explanatory for effects the commander wishes to invoke on an adversary. However, operations may be exclusive of each other or in a sequence, depending on the intent of the operation. Minimizing Detection of Operations for this research is defined as the adversary not becoming aware that an intruder has entered their networks. These three elements are value independent of each other.

The next two elements, however, are value dependent on Minimizing Detection of Operations. Minimizing Attribution Given Detection is defined as the adversary being able to reasonably blame a nation or organization for intruding into the adversary network. Minimizing Compromise Given Detection is defined as other friendly operations, by one or more organizations, being discovered and mitigated by the adversary because of initial detection and subsequent investigation. Of note, to maximize the effectiveness of the operation, a minimization may occur as seen in the last three elements.

Similarly, Minimizing Costs can be broken down into Minimizing Personnel Costs, Minimizing Equipment Costs, Minimizing Infrastructure Costs, and Minimizing Time Costs. Personnel Costs are defined as the wages and other costs needed for a workforce. Equipment costs are defined as the distributed resources available for more than one individual. Equipment Costs entail the associated costs of the hardware and software required for creating the software capabilities and modeling the adversary network. Infrastructure Costs include technical actions taken to conduct and protect the cyber operations infrastructure from attribution, including the eventual replacement of infrastructure for redundancy or because of attribution. Time Costs are the last element of the hierarchy. Although time may be monetized to arrive at an incurred cost, such as labor rates, this approach uses a non-monetized definition. In this research, Time Costs are viewed as the length, in days, for a capability to be prepared before an operation commences. Because the first three elements of this hierarchy are classified for cyberwarfare operations by the DoD, only non-monetized time was used as a cost consideration for minimization as shown in Figure 2.

Figure 2. Objective Hierarchy for Minimizing Costs in Offensive Cyber Operations

Current risk assessment techniques are inadequate for commanders to understand the risks involved in cyber warfare. What is needed is a system in which subjective qualitative measures are discarded for quantification. Achieving quantification is best served by multi-criteria decision-making (MCDM) when multiple considerations are used and must be balanced against the decision-maker's values and priorities. For OCO, the risk may be best defined as the failure to meet minimally acceptable measures of effectiveness or to exceed a maximally acceptable level of cost.

Framework

This framework harnesses the experiences of subject-matter experts (SMEs). To qualify as a SME, participants had to possess a minimum of five years with national-level cyber operations. Participant experience in this effort ranged from five to eighteen years with an average of 8.8 years of national-level operations. SME opinions were modeled for each of three courses of action (COA) offered in each scenario with a truncated triangular distribution. This distribution captured what the SME expected to see 90% of the time in the real world. SMEs provided the most likely rate of success to occur, the highest success rate realistically to be expected, and the lowest success rate to be realistically expected. Each SME provided these assessments for any hierarchical element involved. Examples of this are, "What is the likelihood of COA 1 achieving all the required damage?" or "What is the likelihood of COA 1 being detected?"

SME elicitations ranged from 0, never happening, to 1, will always occur, graduated into one-tenth increments. SME uncertainty manifested in the range of the estimation scores provided. For example, if the SME provided the scores: .4, .55, and .6 for lowest value, most likely, and high value respectively, this person has less uncertainty than a SME that provided the scores of .2, .6, and .9 for the same scenario. Therefore, the wider the range or window of scores, the more uncertainty is involved in the elicitation.

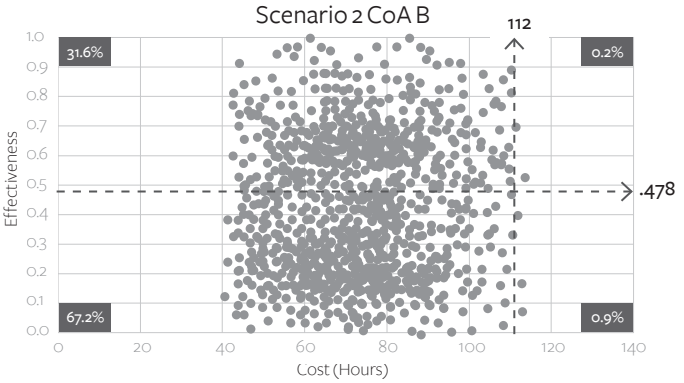

Figure 3. Sample Graphical COA

Since SMEs potentially exhibit the negative cognitive characteristics previously discussed, all SME opinions were equally weighted and then used in a Monte Carlo simulation. The simulation randomly draws SME inputs for each hierarchical concern of the commander. Additionally, the Commander relatively weights concerns to one another. This simulation used the constraints for minimum Effectiveness and the maximum Cost for this operation. Simulations were limited to 3,000 iterations for this research so that later participants could see individual iteration points and how these individual iterations measured against Effectiveness and Cost requirements. Simulations constructed with over 100,000 iterations provided similar distributions; however, the individual iterations of these outputs were indistinguishable. Please see a 3,000 iteration COA simulation output used in Figure 3.

Graphical simulation outputs shown in the Sample COA Evaluation are divided into four regions starting with Region 1 in the upper left corner and then progressing in a clockwise manner. Region 1 is the desired region. In this area, the evaluated course of action meets or exceeds the minimum effectiveness and does not exceed the maximum cost. At the top right is Region 2, where the COA meets the minimum effectiveness but has broken the cost constraint. Below Region 2 is Region 3. In this area, the minimum effectiveness has not been met and the maximum cost has been breached. This is the worst area for a course of action. In the bottom left is Region 4, where the minimum effectiveness has not been met, but the maximum cost has not been exceeded.

Experiment

This research effort elicited the participation of offensive cyber planners at each CCMD, resulting in 60 of the 61 available planners participating. Participants were given a scenario set in five years in the future. In these scenarios, authority to conduct OCO had been delegated to the CCMD with USCYBERCOM conducting deconfliction. Adversaries ranged from peer-state, less advanced nations, and non-nation state actors. Participants were presented with three attacking and three intelligence gathering scenarios. In each scenario, participants had to rank the order of the three COA's based on the commander's guidance for operational goals, desired end-state, and concerns.

Planners read each scenario and the written descriptions of each COA before rank ordering the COA. Participants were then presented with a second group of COA in a graphical format. Participants were told that the graphical COA had no bearing on the written COA. In truth, the graphical COA was the mathematical representation of the written COA based on SME elicitations. Graphical COA were placed in a randomized order to further obfuscate the relation between the two groups. Participants then rank ordered the graphical COA based on the same commander's guidance from the written COA.

Commander preferences for operational goals and tolerances of risk were mathematically modeled using mid-value splitting techniques (Kirkwood 1997). This allows for tradeoff values between operational goals, as defined by the hierarchical objectives previously dis-

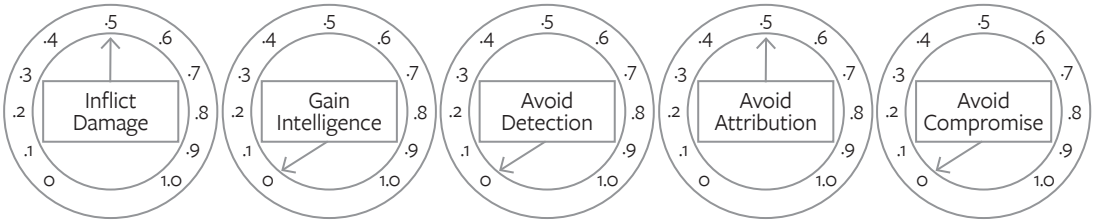

Figure 4. Example Dials for Adjusting for Decision Maker Weight for a Given Operation (Klipstein, 2017)

cussed. The result is a language that allows the commander to fine-tune objectives concerning each other. In the example illustrated in Figure 2, the commander places equal weight on inflicting damage and avoiding attribution. All other hierarchical goals are accounted for with a weight of zero. Commanders may weigh any objective as they wish so long as the total of the five objectives does not exceed 1, the total amount of "care" of the commander.

RESULTS

This research effort investigated how useful a framework with graphical outputs of risk is for aiding personnel who lack the necessary experience. In this effort, the personnel examined were two groups: personnel with national level cyber experience and personnel without national cyber level experience. The experiment focused on the amount of change between the rankings of written and graphical COAs.

This effort determined that 22 of the 36 analyses undertaken met or exceeded statistical significance, suggesting that a framework built on SME knowledge and expertise that incorporated the uncertainty that SMEs acknowledge allowed decision makers to make more informed assessments of risk, and consequently, better decisions regarding unfamiliar and new operations within their organizations. This research succeeded in creating a tailor-made expression of risk based on the Commander's preferences and desires.

Each scenario was analyzed in six different ways. The first three ways are as follows: all participants with no one group, either national level experienced or inexperienced, held constant; all participants with personnel with national level experience held constant; and all participants with personnel lacking national level experience held constant. These three analyses are used for two reasons. The first is to determine if the framework benefits the population. The second is to determine if the increased participant size affects the outcome of the experiment.

The fourth scenario consisted solely of participants with previous national-level experience. The fifth is the converse: personnel lacking national level experience. The sixth analysis focused on USCYBERCOM participants. This analysis examines how effective this framework is for planners currently working at the national level, in addition to being used as the control for comparison against inexperienced personnel. For this paper, only a comparison of

experienced personnel as a group, inexperienced personnel as a group, and USCYBERCOM planners occurs.

Scenario 1A

Scenario 1A introduced the participants to a future scenario in which CCMDs have been partially delegated authority to conduct intelligence and conduct offensive cyber operations. In this first scenario, the combatant commander needs intelligence to ascertain the intentions of an adversary that is threatening a US ally and escalating tensions. Intelligence from other sources indicates the adversary may invade the ally, and the combatant commander wishes to confirm the reports. In this scenario, the commander places values of 60% for avoiding detection, and 40% for gathering the required intelligence. Success in this operation is defined as the exfiltration of a Microsoft Word document outlining the adversary's attack plans, at a minimum.

In retrospect, the COAs for this scenario may have been too similar in their predicted probability of success. Multiple participants noted the potential for detection in the written COAs. COA B, the most popular first choice, was not detected in virtualized testing. The next most popular written COA choice was COA C, which had been used in operations in the past undetected, but virtualized testing indicated that it would be detected. The least popular choice, COA A, was a modified open-source tool with a known signature that offered a 50/50 chance of detection. This information also aligns with the Graphic COA choices.

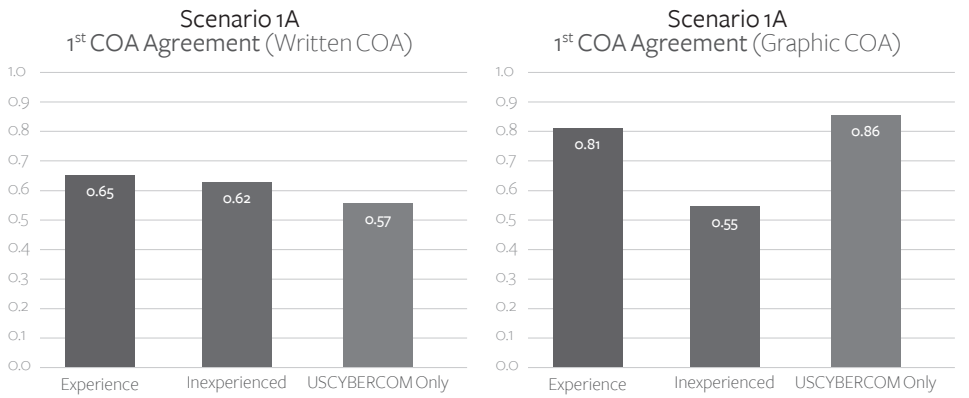

Figure 5. Percentage of 1st Choice COA for Written and Graphic COAs in Scenario 1A

Analysis of the graphical choices made by participants demonstrates that COA B, presented to the participants as COA 3, was the overwhelming first choice in every analysis. COA B had a combined 27.7% predicted effectiveness when Regions 1 and 2 were combined. The second choice in all but two analyses was COA A, also presented as COA A. Not enough data exists in this scenario to accurately account for choices made between the other two COAs when examining second and third choices. As statistical significance was not attained in any

of the analyses in this scenario, no further analysis will be conducted to illustrate support for the advanced hypothesis. Please refer to Figure 5 for the rate of first choice COA agreement for both written and graphic COAs. Although not statistically significant, both the experienced personnel and the USCYBERCOM only groups increased in the aggregated consensus of what the first COA for a recommendation for implementation should be.

Scenario 1B

Scenario 1B is the escalation of Scenario 1A. In this scenario, the commander has attained the required information. Analysis has determined that the adversary intends to erode the trust between the US and its ally by conducting small-scale guerilla attacks. The commander wishes to conduct OCO for two purposes: to disrupt the planning for guerrilla attacks and to demonstrate the network vulnerabilities to the adversary, suggesting that the US is aware of its intentions. The commander places 60% of his value on destruction and 40% on avoiding attribution. Success in this operation is achieved when all information residing on the target containing a one terabyte hard drive is rendered inaccessible and unrecoverable.

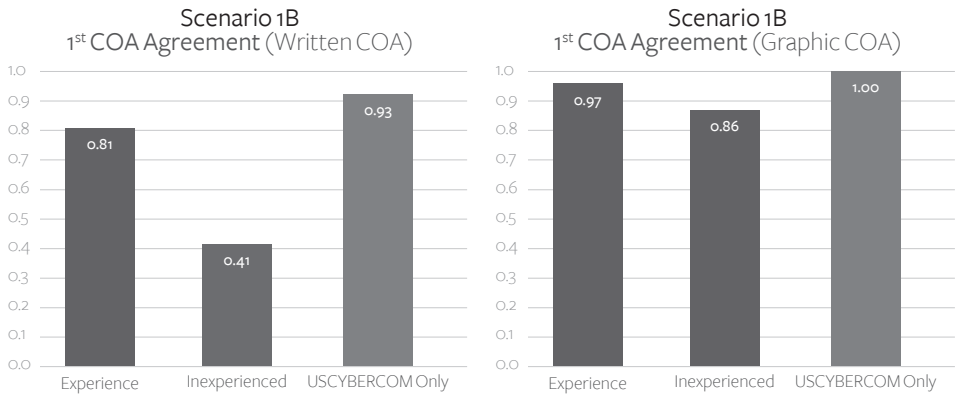

Figure 6 - Percentage of 1st Choice COA for Written and Graphic COAs in Scenario 1B

Figure 6 demonstrates the ability of this framework to aid all personnel with national level experience in understanding risk, not just the inexperienced. In this scenario, the result was not only an increase in aggregated consensus across all three groups but also a change in the primary recommended COA to the commander. Additionally, this scenario demonstrated how a COA might be interpreted as feasible in a written format while having little to no potential for success when mathematically modeled. Remember that thirty SMEs evaluated the different COAs and provided their 90% confidence intervals. This insight further demonstrates the need for quantified risk analysis. Participants, regardless of the method of analysis, continued to make decisions of preference ranking based on Region 1 of the charts, as was observed in Scenario 1A.

Scenario 2

Scenario 2 changes the focus to combating a non-state actor. In this scenario, the actor in

question uses the Internet to recruit, to spread propaganda, and to orchestrate command and control of operations. The non-state actor escalates the situation by posting a video of a captured US military member being killed as a propaganda tool. The combatant commander, working in coordination with the Theater Special Operations Command (TSOC), designated five personnel as high payoff targets. The targeted personnel are instrumental to operations and are believed to be directly connected to the service member's death. For this operation, the combatant commander orders that online intelligence operations are to commence to gather information for ascertaining the patterns of life of the five targets. Once enough information has been attained, the TSOC will coordinate for capture/kill operations to commence. The combatant commander has placed equal value on gaining intelligence while avoiding detection.

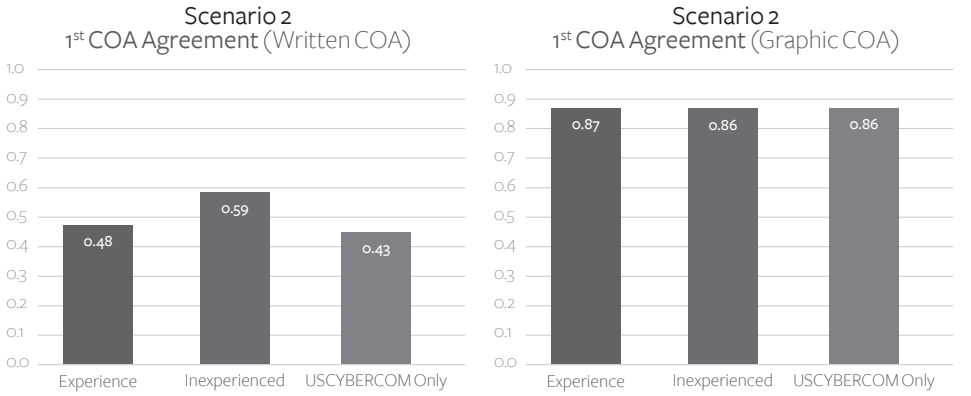

Figure 7. Percentage of 1st Choice COA for Written and Graphic COAs in Scenario 2

As in previous scenarios, the results suggest that participants use Region 1 of the graphics as a tool for assessing preference. This scenario further reinforces the hypothesis that a framework built on SME insights, which quantifies risk, and that presents results in a graphical output can mitigate the inexperience of cyber planners when compared to those with national level experience. In the graphic COAs, 86% of the inexperienced personnel chose the same first preferred COA. This percentage is comparable to the 87% of the overall national level experienced planners and 85% for the USCYBERCOM only planners.

Scenario 3

Scenario 3 presents the participants with another intelligence-gathering operation. In this scenario, an adversarial government uses state-sponsored contracted companies to work on the government's behalf in an attempt to avoid attribution. Intelligence indicates that the contracted company has infiltrated the combatant command networks and exfiltrated documents that update the Theater Security Cooperation agreements, to include personnel and equipment movement schedules and locations.

The commander orders an intelligence operation to confirm or deny the presence of sensitive U.S. military documents within the adversary's network. Confirmation in this operation

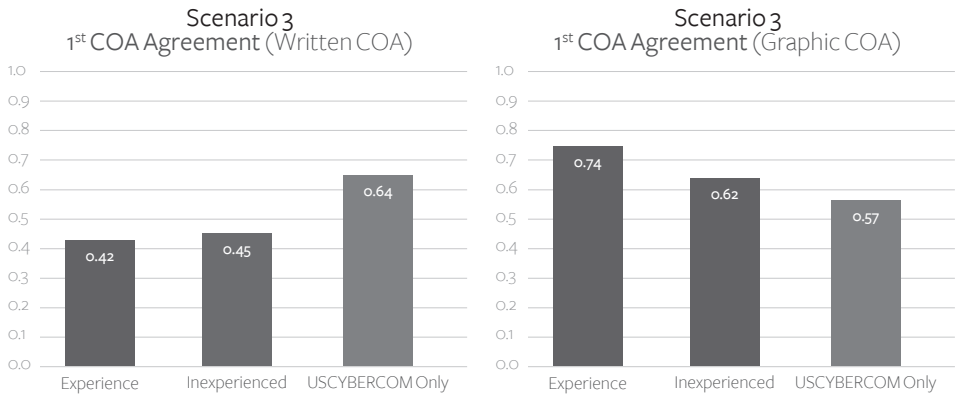

Figure 8. Percentage of 1st Choice COA for Written and Graphic COAs in Scenario 3

is defined as the identification of the 400MB of the non-public portion of the Theater Security Agreement, which ranges in classification from SECRET to TOP SECRET. This operation will be considered a success if all 400MB of the sensitive portion of the document is identified, copied, and downloaded. Notably, OCO action is not authorized at this time.

Analysis of the command's networks indicates that at least two adversary entry points exist and that more are probable. As such, the commander places a value of 60% on avoiding attribution due to the sophistication of the adversary. As the adversary uses state-sponsored contracted companies for operations, the commander also wishes to prevent attribution to the company that works on the adversary's behalf. The remaining 40% of the commander's value comes from the intelligence potentially gained.

As in previous scenarios, indications suggest that participants continue to use Region 1 of the graphics as a tool for assessing preference. All three groups again shifted in the primary COA selection from A to B. In this scenario; the recommended graphic COA had only a 12% predicted success from the simulation compared to 7.1% for the second choice and 2.3% for the third. In the graphic COAs, 62% of the personnel lacking national level experience chose the same first preferred COA. This result is comparable to the 74% of the overall national level experienced planners and 57% for the USCYBERCOM only planners. Due to the groups' 33% increased agreement on COA B being the recommended COA, Scenario 3 again supports the hypothesis advanced by this research.

Scenario 4

In this scenario, the CCMD, in coordination with the CIA, plans to conduct OCO against a non-state actor's online magazine before being published in two weeks. This operation serves two purposes: to prevent disseminating bomb-making information in the magazine and to facilitate the CIA identification of the magazine's readers. Due to other unrelated CIA activities within web forums, the commander has been directed not to bring attribution to US or CIA efforts. Because of this directive, the commander values the outcomes of this operation

at 40% for the destruction or denial of the online material, 30% for avoiding attribution, and 30% for avoiding compromise.

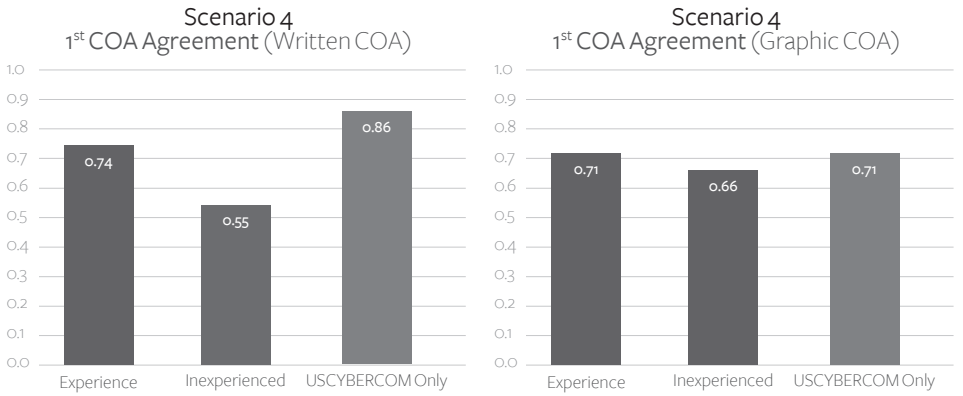

Figure 9. Percentage of 1st Choice COA for Written and Graphic COAs in Scenario 4

Analysis of the outcomes of this scenario suggests two pieces of information were used to rank order COAs. First, the graphic Region 1 prediction matches the written COA ranking. Second, the participant packets showed that participants indicated—using underlining, circling, and highlighting—key information in the written COAs used for decision making. This information pertained to the likelihood of a capability being detected during the operation. The rankings are given in order from least likely to be detected to most likely, matched the written rankings and the graphical Region 1 prediction of success, from most likely to least. Thus, the participants were able to assume the proper ranking of COAs most likely to be based on the written format, suggesting that this scenario suffers from a design flaw. See Figure 9.

Scenario 5

The last scenario for the participants portrays another OCO operation. An adversary of the US uses a state-sponsored business to conduct operations on its behalf. The business in question has targeted US and allied systems with malware for intelligence gathering and denial of service. Additionally, these attacks have been highly publicized in the media but not publicly attributed due to US intelligence equities.

The planned OCO operation will demonstrate to both the adversary and the state-sponsored business that the US is knowledgeable of the adversary's activities. However, US cyber operations must prevent the adversary from discovering and attributing the network infrastructure used for these operations. For these reasons, the commander places 50% of the value of the operation on attaining destruction, 30% on avoiding detection, and 20% on avoiding attribution.

This scenario suggests that participants use Region 1 and Region 2 of the graphics as a tool for assessing preference as observed in Scenario 1A. Again, participants in the aggregate

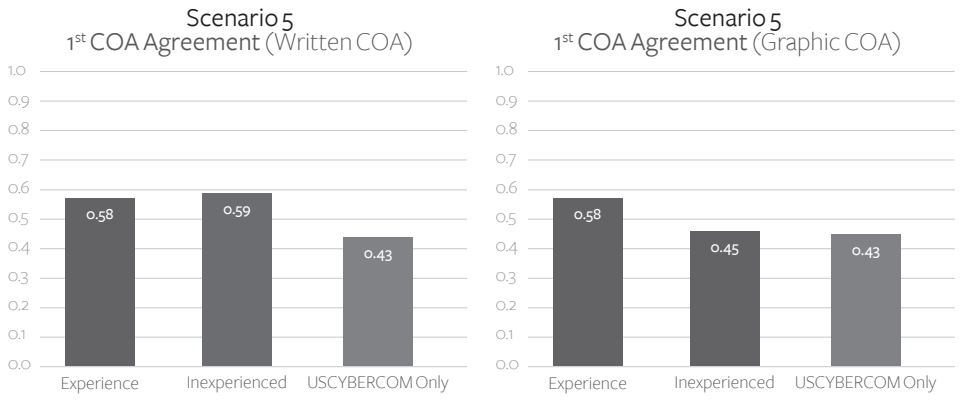

Figure 10. Percentage of 1st Choice COA for Written and Graphic COAs in Scenario 4

changed what COA would be recommended for implementation, from COA B to COA A. This method in which decision makers prioritize data for decision merits further research. In the graphic COAs, 44% of the inexperienced personnel chose the same first preferred COA. This is lower than the 58% of the overall national level experienced planners. Since the USCYBERCOM-only planner analysis was not statistically significant and will not be compared, this result further suggests that the hypothesis is supported. Please refer to Figure 10.

ANALYSIS

The analysis of the collected data uncovered three trends. First, inexperienced personnel overcame their lack of national-level experience and made decisions on par with experienced personnel. Next, using the framework, experienced personnel more often expressed preferences for the same decision and subsequent action. Third, the data suggest that Region 1 in the graphics was the primary determining factor for the decision.

Inexperienced personnel overcome their lack of experience

The first trend identified was the goal of the research, namely, to overcome the lack of national-level experience at organizations below the national-level for OCO. Inexperienced personnel lack a thorough understanding of the environment, along with second and third order effects of operations. For this analysis, inexperienced offensive cyber planners made decisions on par with experienced offensive cyber planners, in addition to offensive planners currently working at USCYBERCOM.

Inexperienced personnel were more likely to agree on a recommended COA in graphical form versus written form in four of six scenarios. The increase in agreement ranged from 18% in Scenario 4 to 47% in Scenario 2. Inexperienced personnel recorded a negative change in agreement, -11%, for the preferred COA while the experienced group and the USCYBERCOM planners both recorded a 50% increase and 25% increase, respectively. Interestingly, the inexperienced personnel bested the experienced and USCYBERCOM personnel in Scenario 1B.

In this scenario, the inexperienced personnel recorded a 33% increase in agreement on the preferred COA while the experienced personnel decreased by 28% and the USCYBERCOM planners decreased by a remarkable 46%.

Scenario 5 was the other scenario in which the inexperienced personnel did not increase in agreement on the preferred COA. In Scenario 5, the inexperienced personnel decreased in agreement by 23%, and as a group changed their preferred COA selection from the written to the graphic. Conversely, the experienced personnel registered no change in the level of agreement, but a change in COA selection. The USCYBERCOM planners as a subset also had no change in their level of agreement but a change in COA selection.

Additionally, the data suggests that the graphics produced by this quantitative framework mitigate the lack of national-level experience possessed by the inexperienced personnel. In the four scenarios that exceeded a 95% confidence interval, inexperienced personnel selected the same COA in comparable numbers to the experienced personnel and the USCYBERCOM planners. For Scenario 1B, the inexperienced personnel chose COA B at a rate of 55%, on par with 58% of the experienced personnel and 50% of the USCYBERCOM planners. Scenario 2 resulted in 86% of the inexperienced personnel choosing COA A along with 87% of the experienced personnel and 85% of the USCYBERCOM planners. Scenario 3 resulted in the USCYBERCOM planners not meeting or exceeding a 95% confidence interval; however, 62% of the inexperienced personnel selected COA B while 74% of the experienced group also did. In Scenario 5, the USCYBERCOM planners again did not exceed a 95% confidence interval. However, 44% of the inexperienced personnel opted for COA C as the primary choice while 58% of the experienced group chose COA A. This analysis suggests that although the hypothesis is supported regarding mitigating the lack of national-level expertise, the framework may also aid experienced personnel.

Value for experienced personnel

Experienced personnel exhibited greater agreement in selecting the first recommended COA when comparing the amount of consensus from the written to the graphic COA. They increased in agreement in four scenarios. Most remarkably, in Scenario 2, they increased in agreement by 80% and in Scenario 3 by 53%. Additionally, the USCYBERCOM planners increased their agreement in three scenarios. Most notably, the agreement for COA recommendation in Scenario 2 doubled. In Scenario 1A, the agreement increased by 50%. Additionally, a quantified framework may be of use in USCYBERCOM if offensive planners continue to rotate out of the organization at the current rate of two to three years.

USCYBERCOM is a military organization working at the national-level of cyber operations, employing both military and civilian personnel. As such, the average military planner leaves this assignment in three years, sometimes two. It is also not unusual for planners to come from diverse backgrounds into USCYBERCOM with no prior experience in cyber operations. Given this, the mean USCYBERCOM experience at the national-level is 3.78 years, less than

the five years needed for an expert status by this research effort and by many other mainstream researchers (Ericsson, Prietula, and Cokely 2007; Prietula and Simon 1989; Macnamara, Hambrick, and Oswald 2014; Ericsson, Krampe, and Tesch-Romer 1993). Only five of the 14 USCYBERCOM planners have a minimum of five years' experience to meet this standard. Three of the five personnel who meet this five-year, expert-level standard are civilians. This unexpected result suggests that the framework is useful for the less experienced national level personnel as well.

Use of Region 1 for decision making

As mentioned in the previous chapter, the data suggests that participants, both with and without national-level experience, typically relied on Region 1 of the graphic representation for a rank preference decision. Recall that Region 1 is the quadrant of the graph that satisfies both the effectiveness and cost requirements of the commander. This suggestion was further reinforced by an examination of the COA ranking choices participants made. The selections of inexperienced personnel, experienced personnel, and USCYBERCOM personnel aligned with the highest Region 1 value in the CoAs for Scenarios 1B, 2, and 4. Additionally, the second and third COA ranking aligned to the second and third highest percentages of predicted success in Region 1 of the CoAs. Furthermore, the USCYBERCOM planners' choices aligned to the highest Region 1 value in Scenario 5.

In two of the scenarios, participants combined the predicted success scores of Regions 1 and 2 to rank their preferences. Region 2 meets the minimum effectiveness of the commander but goes past the maximum time allowed. In Scenarios 1A and 5, except for the USCYBERCOM planners in Scenario 5, participant rankings aligned with the combined scores of Regions 1 and 2. The first preferred COA aligned with the highest combined score, the second with the next highest, and so on. This suggested technique would focus on the effectiveness of a capability without regard to the cost in time. Therefore, the participant only thinks about the end state, not the cost. These findings must be subject to formal testing, however, if they are to be taken as indicative of general decision behavior.

CONCLUSION

This research effort set out to test the hypothesis that a quantifiable framework could mitigate the lack of national-level expertise for OCO at the CCMDs. The outcome is a highly effective framework that considers the operational desires, risk tolerances, and personal values for individual decision makers. This framework uses insights of SME expertise to give a more complete and unbiased view of the probability of success regarding mission effectiveness and the predicted costs. Not only did this research support the hypothesis, but it also has its own unexpected utility for experienced personnel in organizations below the national level and the current USCYBERCOM planners. This framework demonstrated that inexperienced organizations have the potential for making decisions on par with experienced organizations, given that a quantifiable framework and SME insights are available.

NOTES

Arceneau, Kevin, 2012, Cognitive Biases and the Strength of Political Arguments, *American Journal of Political Science,* 56 (2), 271–85.

Asch, Solomon E., 1955, Opinions and Social Pressure, *Scientific American* 193 (5), 2–6.

——. 1956 Studies of Independence and Conformity: I. A Minority of One against a Unanimous Majority, *Psychological Monographs: General and Applied* 70 (9), 1–70. https://doi.org/http://dx.doi.org.libproxy.nps.edu/10.1037/h0093718.

Bennett, Ruth, 2000, Risky Business The Science of Decision Making Grapples with Sex, Race, and Power, *Science News* 158 (12), 190–91, http://www.jstor.org/stable/3981298.

Broder, James F, and Eugene Tucker, 2012, *Risk Analysis and the Security Survey,* 4th ed. Oxford: Butterworth-Heinemann.

Bruyneel, Sabrina D, Siegfried Dewitte, Philip H Franses, and Marnik G Dekimpe, 2009, I Felt Low and My Purse Feels Light : Depleting Mood Regulation Attempts Affect Risk Decision Making, *Journal of Behavioral Decision Making* 170 (October 2008), 153–70, https://doi.org/10.1002/bdm.

Budescu, David V. (University of Illinois at Urbana-Champaign), Stephen Broomell (University of Illinois at Urbana-Champaign), and Han-Hui Por (University of Illinois at Urbana-Champaign), 2009, Improving Communication of Uncertainty in the Reports of the Intergovernmental Panel on Climate Change, *Psychological Science* 20 (3), 299–308.

Buelow, Melissa T., and Julie A. Suhr, 2013, Personality Characteristics and State Mood Influence Individual Deck Selections on the Iowa Gambling Task, *Personality and Individual Differences* 54 (5), 593–97, https://doi.org/10.1016/j.paid.2012.11.019.

Clore, Gerald L (University of Virginia), and Huntsinger, Jeffery R. (University of Virginia), 2007, How Emotions Inform Judgment and Regulate Thought, *Trends in Cognitive Sciences* 11 (9), 393–99, https://doi.org/10.1016(J.tics.2007.08.005.

Council on Foreign Relations, 2014, Amid Tighter Budgets, U.S. Army Rebalancing and Refocusing - Council on Foreign Relations." *Transcripts.* http://www.cfr.org/united-states/amid-tighter-budgets-us-army-rebalancing-refocusing/p32373.

Davis, Paul K, Jonathan Kulick, and Michael Egner, 2005, Implications of Modern Decision Science for Military Decision-Support Systems, Arlington, VA: RAND.

de Dreu, Carsten K W, Bernard A Nijstad, and Daan van Knippenberg, 2008, Motivated Information Processing in Group Judgment and Decision Making, *Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc 12* (1), 22–49, https://doi.org/10.1177/1088868307304092.

de Langhe, Bart, and Stefano Puntoni, 2015, Bang for the Buck: Gain-Loss Ratio as a Driver of Judgement and Choice, *Management Science* 61 (5), 1137–63.

Department of Defense, 2017a, *Joint Operations,* Washington D.C.

——, 2017b, Joint Operations, Washington D.C.: Department of Defense.

Department of the Army, 2012, ADP 6-0 (*Mission Command*), Washington D.C.: Army Publishing Directorate, https://armypubs.us.army.mil/doctrine/index.html.

——, 2013, AR 385-10 *The Army Safety Program*, Washington D.C.: Army Publishing Directorate.

Dowd, Keith W, John V Petrocelli, and Myles T Wood, 2014, Integrating Information from Multiple Sources: A Metacognitive Account of Self-Generated and Externally Provided Anchors, *Thinking & Reasoning* 20 (3). Department of Psychology, Wake Forest University, Winston-Salem, NC, US petrocjv@wfu.edu; Petrocelli, John V.,P.O. Box 7778,Winston-Salem,US,27109,Department of Psychology, Wake Forest University,petrocjv@wfu.edu: Taylor & Francis, 315–32, https://doi.org/http://dx.doi.org/10.1080/13546783.2013.811442.

Ellsberg, D., 1961, Risk, Ambiguity, and the Savage Axioms, *Quarterly Journal of Economics* 75 (November), Unlisted:643–69, http://libproxy.nps.edu/login?url=http://search.proquest.com/docview/864366624?accountid=12702.

Ericsson, K. Anders, Ralf Krampe, and Clemens Tesch-Romer, 1993, The Role of Deliberate Practice in the Acquisition of Expert Performance, *Psychological Review* 100 (3), 363–404, https://doi.org/0033-295X/93.

Ericsson, K Anders, Michael Prietula, and Edward Cokely, 2007, The Making of an Expert, *Harvard Business Review* July-Augus, 1–7.

Figner, Bernd, Rachael J. Mackinlay, Friedrich Wilkening, and Elke U. Weber, 2009, Affective and Deliberative Processes in Risky Choice: Age Differences In Risk Taking in the Columbia Card Task, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35 (3), 709–30, https://doi.org/http://dx.doi.org/10.1037/a0014983.

Foushee, H C. 1982, The Role of Communications, Socio-Psychological, and Personality Factors in the Maintenance of Crew Coordination, *Aviation, Space, and Environmental Medicine* 53 (11), 1062–66, http://www.researchgate.net/publication/16049243_The_role_of_communications_socio-psychological_and_personality_factors_in_the_maintenance_of_crew_coordination.

NOTES

Garvin, David A, and Michael A Roberto, 2001, What You Don't Know About Making Decisions, *Harvard Business Review* September, 22–32.

Gilovich, Thomas, 1991, *How We Know What Isn't So.* New York, NY: The Free Press.

Girodo, Michel, 2007, Personality and Cognitive Processes in Life and Death Decision Making: An Exploration into the Source of Judgment Errors by Police Special Squads, *International Journal of Psychology* 42 (6), University of Ottawa, Ottawa, ON, Canada girodo@uottawa.ca; Girodo, Michel,145 Jean-Jacques Lussier Street,Ottawa,Canada,K1N 6N5,School of Psychology, University of Ottawa,girodo@uottawa.ca: Taylor & Francis Wiley-Blackwell Publishing Ltd., 418–26, https://doi.org/http://dx.doi.org/10.1080/00207590701436728.

Grifffin, Robert J, Kurt Neuwirth, James Giese, and Sharon Dunwoody. 2002, Linking the Heuristic-Systematic Model and Depth of Processing, Communication Research. 2002, https://doi.org/10.1177/009365002237833.

Heilbronner, Sarah R., Benjamin Y. Hayden, and Michael L. Platt, 2010, Neuroeconomics of Risk-Sensitive Decision Making, In *Impulsivity: The Behavioral and Neurological Science of Discounting,* edited by Gregory Madden, PhD and Warren Bickel, PhD, 159–87, American Psychological Association.

Holt, Charles A, and Susan K Laury, 2002, Risk Aversion and Incentive Effects, *The American Economic Review* 92 (5). Nashville: American Economic Association, 1644–55. http://libproxy.nps.edu/login?url=http://search.proquest.com/docview/233033790?accountid=12702.

Hubbard, Douglas W. 2009, Worse Than Useless: The Most Popular Risk Assessment Method and Why It Doesn't Work, In *The Failure of Risk Management: Why It's Broken and How to Fix It,* 117–43, Hoboken, NJ: John Wiley & Sons.

Kahneman, Daniel, 2003, A Perspective on Judgement and Choice, *American Psychologist* 58 (9), 697–720, https://doi.org/10.1037/0003-066X.58.9.697.

——. 2013, Thinking Fast and Slow, 2nded, Farrar, Straus and Giroux.

Kahneman, Daniel, and Dan Lovallo, 1993, Timid Choices and Bold Forecasts - a Cognitive Perspective on Risk Taking, *Management Science* 39 (1), 17–31, http://libproxy.nps.edu/login?url=http://search.proquest.com/docview/38481183?accountid=12702.

Kahneman, Daniel, and Amos Tversky, 1979, Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2), 263–91, http://libproxy.nps.edu/login?url=http://search.proquest.com/docview/56137572?accountid=12702.

——, 1984, Choices, Values, and Frames, *American Psychologist* 39 (4):341–50.

Kallberg, Jan, and Thomas Cook, 2017, The Unfitness of Tradional Military Thinking in Cyber, IEEE Access 5, 8126–30.

Kane, Joanne, and Gregory D. Webster, 2013, Heuristics and Biases That Help and Hinder Scientists: Toward a Psychology of Scientific Judgment and Decision Making, In *Handbook of the Psychology of Science,* edited by Gregory PhD Feist and Michael PhD Gorman, 1sted., 437–59, New York, NY: Springer.

Kirkwood, Craig, 1997, *Strategic Decision Making,* Edited by Carl Hinrichs, Belmont, CA: Wadsworth Publishing Company.

Klipstein, Michael. 2017, Quantifying Risk for Offensive Cyber Operations, Naval Postgraduate School.

Ludvig, Elliot A. (Princeton University), Christopher R. Madan (University of Alberta), and Marcia L. Spetch (Universtiy Medical Center Hamburg-Eppendorf), 2013, Extreme Outcomes Sway Risky Decisions from Experience, *Journal of Behavioral Decision Making* 27, 146–56, https://doi.org/10.1002/bdm.1792.

Macnamara, Brooke, David Hambrick, and Frederick Oswald, 2014, Deliberate Practice and Performance inMusic, Sports, Education, and Professions: A Meta-Analysisle, *Association for Psychological Science*2 25 (8), 1608–18.

Milkman, Katherine L, Dolly Chugh, and Max H Bazerman, 2009, How Can Decision Making Be Improved?, *Perspectives on Psychological Science* 4 (4), University of Pennsylvania, Philadelphia, PA, US kmilkman@wharton.upenn.edu; New York University, New York, NY, US ; Harvard University, Cambridge, MA, US; Milkman, Katherine L.,500 Jon M. Huntsman Hall,Philadelphia,US,19104,Wharton School, University of: Wiley-Blackwell Publishing Ltd. Blackwell Publishing Sage Publications, 379–83, https://doi.org/http://dx.doi.org/10.1111/j.1745-6924.2009.01142.x.

Nygren, Thomas E, Alice M Isen, Pamela J Taylor, and Jessica Dulin, 1996, The Influence of Positive Affect on the Decision Rule in Risk Situations: Focus on Outcome (and Especially Avoidance of Loss) rather than Probability, *Organizational Behavior and Human Decision Processes* 66 (1), New York: Elsevier Science Publishing Company, Inc., 59, http://libproxy.nps.edu/login?url=http://search.proquest.com/docview/223179153?accountid=12702.

Prietula, Michael, and Herbert Simon, 1989, The Experts in Your Midst, *Harvard Business Review,* no. January-February 1989, 120–24.

NOTES

Savage, Sam, 2012, *The Flaw of Averages*, Hoboken, NJ: Wiley.

Tversky, Amos, and Daniel Kahneman, 1974, Judgment under Uncertainty: Heuristics and Biases, *Science* 185 (4157), Hebrew U, Jerusalem, Israel: American Assn for the Advancement of Science, 1124–31, https://doi.org/http://dx.doi.org/10.1126/science.185.4157.1124.

Tversky, Amos, and Daniel Kahneman, 1981, The Framing of Decisions and the Psychology of Choice, *Science* 211 (4481), Stanford U: American Assn for the Advancement of Science, 453–58, https://doi.org/http://dx.doi.org/10.1126/science.7455683.

Weber, Bethany J, and Gretchen B Chapman, 2005, Playing for Peanuts: Why Is Risk Seeking More Common for Low-Stakes Gambles?, *Organizational Behavior and Human Decision Processes* 97 (1), Psychology Department, Rutgers University, Piscataway, NJ, US bweber@eden.rutgers.edu; Weber, Bethany J.,152 Frelinghuysen Road,Piscataway,US,08854-8020,Psychology Department, Rutgers University,bweber@eden.rutgers.edu: Elsevier Science, 31–46, https://doi.org/http://dx.doi.org/10.1016/j.obhdp.2005.03.001.