On Cyber-Enabled Information/Influence Warfare and Manipulation

March 21, 2017
Draft 3

Herbert Lin (Stanford University) and Jackie Kerr (LLNL)

Abstract

The United States has no peer competitors in conventional military power. But its adversaries are increasingly turning to asymmetric methods for engaging in conflict. Much has been written about cyber warfare as a domain that offers many adversaries ways to counter the U.S. conventional military advantages, but for the most part, U.S. capabilities for prosecuting cyber warfare are as potent as those of any other nation. This paper advances the idea of cyber-enabled information/influence warfare and manipulation (IIWAM) as a form of conflict or confrontation to which the United States (and liberal democracies more generally) are particularly vulnerable and are not particularly potent compared to the adversaries who specialize in this form of conflict. IIWAM is the deliberate use of information against an adversary to confuse, mislead, and perhaps to influence the choices and decisions that the adversary makes. IIWAM is a hostile activity, or at least an activity that is conducted between

two parties whose interests are not well-aligned, but it does not constitute warfare in the sense that international law or domestic institutions construe it. Cyber-enabled IIWAM exploits modern communications technologies to obtain benefits afforded by high connectivity, low latency, high degrees of anonymity, insensitivity to distance and national borders, democratized access to publishing capabilities, and inexpensive production and consumption of information content. Some approaches to counter IIWAM show some promise of having some modest but valuable defensive effect. But on the whole, there are no good solutions for large-scale countering of IIWAM in free and democratic societies. Development of new tactics and responses is therefore needed.

1. Introduction

From the standpoint of traditional military conflict, the United States is unmatched by any other nation. Other nations have taken note of U.S. conventional military prowess and sought other "asymmetric" methods for confronting the United States and other Western nations—that is, they seek to confront the United States and other Western nations targeting their weaknesses and vulnerabilities. Cyber warfare is one asymmetric counter to Western (and especially U.S.) military advantages that depend on the use of cyberspace.[1]

"Cyber warfare" spans a broad spectrum. At the high end, cyber conflict threatens critical national infrastructure, e.g., information technology systems that are vital to society or national interests, such as the computers controlling the electric grid or air traffic control systems; undetected alteration of financial data held by major financial institutions; and computerized weapons systems unable to hit their targets because they have lost their ability to access GPS.

Much of high-end cyber conflict amounts to war by any standard. In turn, war has connotations of hard power: armed conflict, violence, death and destruction, shooting, kinetic weapons, and clear transitions between war and peace. The patron saint of war is Clausewitz, who wrote that "War . . . is an act of violence to compel our opponent to fulfill our will"[2] and in war, "the fighting forces must be destroyed."[3]

But not all cyber conflict resembles war in the Clausewitzian sense. Lower-level cyber conflict involves credit-card fraud; intellectual property theft involving blueprints, business data, and contract negotiating positions; compromises of personal information such as credit reports and medical data; denial of service attacks that prevent rightful users from accessing online resources. Such activities can have significant effects on nations over time, but they do not rise to the level of war.

This paper extends the spectrum of cyber conflict to a domain that is not even necessarily home to activity that is illegal under either domestic or international law but that nevertheless has profound threat implications for modern democracies—that domain is cyber-enabled information/influence warfare and manipulation.

2. Information/Influence Warfare and Manipulation

Information/influence warfare and manipulation (IIWAM) is the deliberate use of information by one party on an adversary to confuse, mislead, and ultimately to influence the choices and decisions that the adversary makes. IIWAM is a hostile non-kinetic activity, or at least an activity that is conducted between two parties whose interests are not well-aligned. At the same time, IIWAM is not warfare in the Clausewitzian sense (nor in any sense presently recognized under the laws of war or armed conflict), which accounts for the "manipulation" part of the term. IIWAM has connotations of soft power: propaganda, persuasion, culture, social forces, confusion, deception. The patron saint of IIWAM is Sun Tzu, who wrote that "The supreme art of war is to subdue the enemy without fighting."[4]

Note that IIWAM is a methodology or an approach to how one party (Party A) might deal with another party (Party B) seen as an adversary. Party A and Party B can be nations, nonstate actors, or domestic populations, and in principle IIWAM could entail an adversarial relationship in any combination (that is, nations against other nations, against nonstate actors, or against its domestic population; nonstate actors against nations, against other nonstate actors, or against its domestic population; or populations against their home nations, against nonstate actors, or against other domestic populations).

2.1 The Information Environment

The battlespace of IIWAM is the information environment. The information environment is the aggregate of individuals, organizations, and systems that collect, process, disseminate, or act on information.[5] The information environment has three interrelated dimensions—physical, informational, and cognitive/emotional—in and through which individuals, organizations, and systems continually interact.

- The physical dimension is composed of command and control systems, software, key decision makers, and supporting infrastructure that enable individuals and organizations to create effects.
- The informational dimension specifies where and how information is collected, processed, stored, disseminated, and protected.[6]
- The cognitive/emotional dimension encompasses the minds and emotions of those who transmit, receive, and respond to or act on information.

2.2 Strategy and a Theory of Victory in Information/Influence Warfare

In IIW, victory is achieved by A when A succeeds changing B's political goals so that they are aligned with those of A. But such alignment is not the result of B's "capitulation" or B's loss of the ability to resist—on the contrary, B (the losing side) is openly willing. That is, IIWAM victory shares the Clausewitzian focus on the opponent's will, but not its focus on destroying military forces.

IIWAM mostly uses words and images to persuade, inform, mislead, and deceive so that the adversary does not use the (fully operational) military assets it does have, and the military outcome is the same as if those military assets had been destroyed. IIWAM operations also provide additional options for action when it is undesirable for some reason to refrain from using kinetic military operations. Most importantly, IIWAM takes place below legal thresholds of "use of force" or "armed attack," and at least in an international legal sense does not trigger the use of military force in response.

The targets of IIWAM are the adversary's perceptions, which reside in the cognitive dimension of the information environment. IIWAM focuses on damaging knowledge, truth, and confidence, rather than physical or digital artifacts; the former reside in "brain-space" rather than 3-D space or cyberspace. IIWAM seeks to inject fear, anger, anxiety, uncertainty, and doubt into the adversary's decision making processes. Successful IIWAM practitioners alter adversary perceptions and are able to predict how altered perceptions increase the likelihood that the adversary will make choices that are favourable to the IIWAM practitioner.

IIWAM seeks to influence individuals, organizations, news media, government agencies, political leadership and segments of society. Furthermore, these entities are not only military entities—there are no "noncombatants" that enjoy immunity from IIWAM attack. IIWAM attacks the legitimacy of entities larger than ad hoc groups of individuals— government and other institutions that promote a larger societal cohesion (e.g., schools, news media) are particularly important targets from this perspective.

IIWAM perpetrators may also find that the sowing of chaos and confusion in an adversary for its own sake serves their interests. For example, an adversary whose government is in chaos and whose population is confused is unlikely to be able to take decisive action about anything, at least not without extended delay, thus affording the IIWAM user more freedom of action. Sowing chaos and confusion is thus essentially operational preparation of the information battlefield—shaping actions that make the information environment more favourable for actual operations should they become necessary. In addition, introducing sufficient chaos into the information environment may reveal targets of opportunity that can be exploited.

2.3 Operations in Information/Influence Warfare

2.3.1 How IIWAM Operations Achieve Their Objectives

IIWAM operations are activities that seek to affect (change) the information environment in any one, or all, of its three dimensions (physical, informational, and cognitive/emotional) in ways that provide advantages over the adversary. IIWAM operations can be (and mostly have been) conducted outside the explicit context of military operations (e.g., when traditional military operations are not going on) by entities without affiliation to military forces or military command and control.

IIWAM operations are primarily psychological in nature. IIWAM operations convey selected information and indicators to adversary audiences to influence their emotions, motives, objective reasoning, and ultimately the behaviour of adversary governments, organizations, groups, and individuals. Their purpose is to induce or reinforce adversary attitudes and behaviour in ways favourable to the originator's objectives.[7]

The key term in the definition of IIWAM operations is the conveyance of **selected** information to adversary audiences. The selected information may be mostly false, mostly true, or some mix of the two, and "selected information" stands in contrast to "all relevant information," a phrase that might be used in normal discourse regarding, for example, honest educational efforts. In IIWAM operations, information is selected for conveyance on the basis of whether it will influence the audience's attitudes and behaviour in favourable manner, rather than on whether it contributes to a fair or balanced or objective presentation in which the audience can decide for itself. (Of course, it may be in the interest of the originator to appear that the operation is all of the latter.)

IIWAM operations may be white, grey, or black.[8] White IIWAM operations clearly and correctly identify the originator: a white IIWAM operation publicly associated with Nation A is in fact conducted by Nation A. Grey IIWAM operations are not publicly associated with any actor at all. Nation A may originate an IIWAM operation but if the operation is grey, no national actor is identified. Black IIWAM operations are publicly associated with a nation or actor other than the true originator: thus, black IIWAM operations are by definition "false-flag" operations. If Nation A originates a black IIWAM operation, Nation A may be construct it so that it is publicly associated with Nation C.

Depending on the purpose of the IIWAM operation and the risks entailed, a white, grey, or black operation may be more suitable. For the United States, grey or black IIWAM operations targeting certain audiences (e.g., U.S. citizens) are constrained by law and/or policy.

IIWAM operations may also involve deception. Deceptive IIWAM operations can be executed to induce adversaries to take (or fail to take) specific actions that will advantage the IIWAM originator and/or disadvantage the adversary. Deceptive IIWAM operations seek to reinforce the adversary's preconceived beliefs, focus the adversary's attention on unimportant activities so that important activities go unnoticed; create the illusion of strength where weakness exists; overload the adversary's information collection and analytical capabilities; and reduce the adversary's situational awareness.

The impact of IIWAM operations can be significantly increased in two types of use:

- When IIWAM operations are used to channel or influence other preexisting forces in society. Here, the actual large-scale impact is the direct result of economic forces, cultural forces, social forces, psychological forces, organizational or bureaucratic forces rather than anything specific impact resulting directly from a particular IIWAM operation.
- When IIWAM operations are used in a pre-existing atmosphere of uncertainty and doubt. The side using IIWAM operations knows what its intentions are, what it hopes to

> accomplish, and what its future plans and moves are. By contrast, a doubtful or uncertain adversary is likely to dither in determining the scope and nature of the actual threat and about what should be done about it. Dithering consumes valuable time, during which the IIWAM attacker can create new facts on the ground and may even change the adversary's strategic calculus.[9]

IIWAM is not likely to be a supremely powerful instrument of conflict in the same sense as nuclear weapons. Because IIWAM is primarily psychological in nature, there will always be people in a target population that are immune to its effects—this is most true in populations that have strong institutions and traditions dedicated to the rule of law and relatively sane and trustworthy (i.e., not corrupt) political leaders. But in instances when only a small number of people need to behave differently because of IIWAM conducted against them (e.g., in close electoral contests), IIWAM can prove decisive.

2.3.2 The Psychological Basis for IIWAM Operations

2.3.2.1 Cognitive biases

IIWAM operations usually take advantage of cognitive biases in human beings. These biases result from human use of intuitive reasoning strategies rather than analytical strategies. One of the most important intuitive reasoning strategies are heuristics that substitute simple judgments for complex inferential tasks, resulting in cognitive biases that sometimes lead to erroneous conclusions.[10]

For IIWAM purposes, some of the most important heuristics are the availability heuristic (people judge events or objects as frequent, probable, or causally powerful by the ease with which examples of those events or objects can be brought to mind);[11] the representativeness heuristic (people categorize events or objects on the basis of their resemblance to the underlying category characteristics); the anchoring heuristic (people give excessive weight to initial estimates in subsequent adjustments of those estimates); and the affect heuristic (people judge the risks and benefits of an event or a course of action depending on the positive or negative feelings that they associate with it).[12]

A variety of cognitive biases arise from the use of these heuristics. Here are a few illustrative examples:

- Fluency bias arises when the ease with which an individual processes information about an idea, object or event fuels the expectation of being able to give a positive response to it. Simplistic and one-sided messaging takes advantage of the fluency bias.
- Confirmation bias is an individual's preference for seeking and interpreting new information in ways that are consistent with their beliefs, attitudes, and decisions, and to steer away from inconsistent information.[13] Media channels such as Fox News play to this bias for individuals with a right-of-centre orientation, and similarly for MSNBC for those with a left-of-centre orientation.

- Illusory truth bias is an individual's perception of greater truth for statements that are easier to process, for example, as the result of repetition. IIWAM operations thus often convey the same message repeatedly.
- Loss aversion bias is an individual's greater sensitivity to loss than to gain.[14] In many instances, people will take reckless gambles to recoup a loss but proceed cautiously when trying to improve their situation. IIWAM operations thus often emphasize how bad a situation is to prime people for acting more recklessly.
- Recency bias is a tendency to rely upon memories that are easily accessed, which can encourage the use of recently presented information even when it is inaccurate.[15]

Biases such as these (a more complete list of biases can be found in Jonathan Baron's work, *Thinking and Deciding*.[16]) are vulnerabilities in the cognitive armor of otherwise rational and analytical individuals, and designing IIWAM operations against these vulnerabilities is likely to enhance their effectiveness.

2.3.2.2 Emotional biases

The cognitive biases described above suggest how the judgments and conclusions of actual human beings may differ from those of the hypothetical maximally rational person due to a reliance on fallible mental heuristics. But emotional factors also affect the judgments and conclusions that people make. Emotional biases can be seen when an individual has a motivation for believing (i.e., an emotional investment) in a particular answer or outcome or view that prevents him or her from achieving the benefits of rational consideration.

For example, a variety of studies have found that individuals are uncomfortable (an emotional reaction) to inconsistencies between their behaviour and their beliefs or attitudes, and are motivated to eliminate those inconsistencies.[17] A most common way to do so is for them to change their perception of inconsistency regarding their behaviour. They may rationalize their behaviour so that they can see the behaviour as consistent with their beliefs and attitudes or avoid exposure to information that challenges their beliefs and seek information that bolsters their beliefs.[18]

People are also more likely to arrive at conclusions that they want to arrive at (i.e., conclusions that they feel motivated).[19] Their reasoning is also motivated by a desire to protect their status within an affinity group whose members share defining cultural commitments.[20]

People subject arguments that are favourable to their own position to a less rigorous and critical analysis compared to arguments that are unfavourable.[21] In the political context, an individual's emotional stance towards a political candidate is more important than his or her view about that candidate's policies[22] or the facts known about the candidate.[23]

Findings such as the preceding suggest that IIWAM operations that stimulate the emergence of strong emotion such as fear, ethnocentrism, and pride are likely to make those targeted more resistant to factual information and less willing to engage in reflective rational consideration.

2.3.3 A Typology of IIWAM operations

This paper explores three distinct kinds of IIWAM operation: propaganda operations, leak operations, and chaos-producing operations.

2.3.3.1 Propaganda operations

A debate exists within the social science literature about the definition of "propaganda." Some scholars assert that all types of mass persuasion constitute propaganda.[24] Other scholars defines propaganda as "The organized attempt through communication to affect belief or action or inculcate attitudes in a large audience in ways that circumvent or suppress an individual's adequately informed, rational, reflective judgment."[25]

These contrasting definitions have in common an emphasis on conveying information to large audiences to influence opinion, attitudes, and emotion in ways that help the originator. In this context, Hitler's ideas on propaganda remain relevant today—propaganda should attract broad public attention, provide the most simple formulations of essential ideas, focus on appealing to the emotions of the public rather than their reasoning powers, and repeat the conveyed messages continually.[26]

There is also no requirement that the information conveyed be true. Hitler was an advocate of "the big lie," [27] believing that the broad masses would "more readily fall victims to the big lie than the small lie, since . . . It would never come into their heads to fabricate colossal untruths, and they would not believe that others could have the impudence to distort the truth so infamously."

2.3.3.2 Chaos-producing operations

Chaos-producing operations are operations that confuse and disrupt by means of misinformation for no purpose other than the creation of chaos. Such operations disorient without seeking a specific behavioural outcome but serve useful purposes by lowering an adversary's situational awareness and increasing the uncertainty in the environment.

For example, on September 11, 2014, St. Mary Parish in Louisiana was the subject of a well-coordinated and professionally produced IIWAM chaos-producing operation claiming that a powerful explosion had occurred at the local Columbian Chemicals plant.[28] This operation included hundreds of Twitter accounts documenting the disaster, still images and videos of the explosion and flames, text messages to many local residents, a screen shot of CNN's home page discussing the event, and a YouTube video in which ISIS claimed credit for the attack.

It was all fake. The perpetrator had gone to enormous efforts to stage this operation, simply to create some hours and perhaps days of chaos and concern in the St. Mary Parish. Had this been a one-time event, it could have been a mere blip on the national scene, the equivalent

of "a tasteless prank," in the words of the director of the St. Mary Parish Office of Homeland Security and Emergency Preparedness. But it was not—rather, it was one of several such events orchestrated in the second half of 2014.

Although chaos-producing operations and propaganda operations share a lack of concern for truth, the latter are conducted to convey a particular political point of view to the target audience. The former have no such goal—taken in isolation and by themselves, they are not political at all, at least not explicitly.

Chaos-producing operations also have the important virtue that their messaging need not be consistent—for myriad messages to be inconsistent with each other helps rather than hurts the spread of chaos. Moreover, inconsistent messages need not be coordinated with each other—which means they can be produced in large volume very rapidly by a variety of different sources.

2.3.3.3 Leak operations

If the information conveyed is mostly true, an IIWAM operation is most similar to a leak of information. Leaks convey information to the target audience information that the adversary might wish to keep out of public view, and when disclosure occurs in the context of disclosing secret information, it gains notoriety and attracts attention disproportionately to its actual importance. Paraphrasing an editorial in the New York Times,[29] there is a difference between treating a piece of information as newsworthy even though it was leaked and treating a piece of information as newsworthy because it was leaked. It is also worth noting that Wikileaks in particular has skillfully exploited this phenomenon and can entice even mainstream media into reporting on any claim that Wikileaks wishes to make, because of the expectation that some leaked documents will underlie that claim.[30]

A mix of true and false information may be more efficacious than pure truth or pure lies. Pure truth may be inconvenient in the sense that true statements may not be available to support the message that the IIWAM operation wishes to convey.[31] A listener who recognizes lies as lies is likely to become more sceptical of subsequent statements, whereas a listener who recognizes statements as true is more likely to believe that subsequent statements are true—one aspect of a cognitive bias known as truth bias in cognitive and social psychology.[32] This phenomenon is also manifested even when people have good reason to refrain from assuming truth.

3. Cyber-Enabled Information/Influence Warfare

Modern information technology—i.e., computers and communications technology, that is, the "cyber" portion of "cyber-enabled IIW"—afford IIWAM practitioners a variety of new opportunities. Unlike information technologies of the past (e.g., books, film), modern information technologies effectively separate information (represented as ones and zeros, i.e.,

as bits) from the physical substrate (e.g., paper) needed in the past to convey information. The following characteristics of today's information environment are noteworthy.

- High connectivity. In 2016, the number of Internet users globally approached 3.5 billion people,[33] and nearly every user on the Internet is connected to every other one through a relatively small number of links.
- Low latency. Users that are directly linked can be notified in milliseconds of new communications and information rather than the hours or days that characterized radio, telephone, or newspaper communication.
- Anonymity. Information represented in digital form always be physically separated at some point from identifying information, at which point any party can be associated with it.
- Low cost. The marginal cost of conveying more bits of information is essentially zero in most instances today using modern information technology, which more or less eliminates volume as a constraint on the information people can send and receive.
- Multiple distribution points. There are numerous content providers on the Internet, ranging in size from single individual teenagers and automated bots to government agencies, that supply information.
- Many-to-many bi-directional communications. Consumers and content providers easily engage in reciprocal dialogue and the lines between consumer and provider are often indistinct.
- Disintermediation. Today's information environment is far less reliant on established intermediaries than the environment of a few decades ago. In the past, intermediaries such as newspapers played editorial roles helped their readers to manage, interpret, and evaluate large volumes of information. Today, more users depend on the newsfeeds of social media and technological tools to filter and sift information, but these tools lack serious editorial judgment.
- Insensitivity to distance and national borders. It is just as easy to send information across the ocean as across the street, and national borders are much more porous to information than they are to physical objects.
- High availability of personal information. Large quantities of personal information of individuals are available to interested parties, either for free or for a nominal price.
- Information insecurity. All information is subject to risks related to compromises of confidentiality, integrity, availability, and authenticity, but digitally recorded information arguably suffers these risks to a greater degree. A full discussion of such risks is beyond the scope of this paper, but it suffices to say that recording information digitally often engenders a false sense of security (likely because protecting bits of information is different from protecting a physical artefact storing bits), and people continue to be surprised when the security of their information is compromised.

These characteristics of the information environment writ large have a number of important implications for the prosecution of IIWAM.

Perhaps the most significant observation about cyber-enabled IIWAM is that unlike the cyber warfare described in Section 1, cyber-enabled IIWAM operations need not be particularly sophisticated to be effective, as happened in the Russian email hacks described above.

Furthermore, and as described in Section 2.3.1, the impact of cyber-enabled IIWAM operations can be enhanced by channelling larger forces to amplify their effects. At the same time, enhanced impact does not come for free—planning for and predicting psychological, legal, organizational, societal, and economic effects, especially on a large scale, is an exercise in predicting second order effects, that is, effects that go beyond the technical effects of a cyber operation. This constitutes a significant expansion of the space that planners of an IIWAM attack must account for—and IIWAM defenders as well.

For example, IIWAM originators can engage in a very high tempo of operations—it is fast, easy and cheap to send out tweets and Facebook notifications, and tsunamis of information can be generated rapidly with little warning. Responses to noteworthy events in the real world can also be issued rapidly. Rapid response and a high tempo of operations means that the IIWAM originator can obtain first-mover advantages that allow him or her to set the initial terms of the messaging narrative.

A high tempo of operations is particularly useful for IIWAM chaos-producing operations. A great deal of experience with the Internet over the past several decades suggests that information suppression by removing it is a difficult if not impossible task. Attempts to remove information often (and arguably usually) leads to drawing more attention to that information, because it is impossible to destroy all copies of digitally stored information once a copy has become public. But another method to suppress a message that is almost as effective is to drown it out with competing messages (i.e., by creating messaging chaos with a flood of mutually inconsistent messages) instead of trying to remove it.

High connectivity also means that even actors whose voice would have been small before the rise of the Internet now have megaphonic reach to large audiences. Communities of like-minded "fringe" individuals are much easier to form under such circumstances, where such individuals can and do receive social reinforcement for their views.

High connectivity has particular relevance to today's political campaigns, which are a mix of "official" campaigns controlled by candidates and unofficial (and formally unrelated) "informational" campaigns conducted by supporters (and opponents) of those candidates. The Internet has encouraged the proliferation of politically oriented Web sites in the United States and elsewhere established by private citizens that are not subject to government regulation regarding campaign financing or fairness, and some of these sites are as influential as any traditional political or media outlet.

IIWAM originators can operate in relative anonymity, which eliminates the possibility of negative social consequences from engaging in such activities and reduces social inhibitions about engaging in such behaviour. Free of inhibitions, the number of individuals willing to engage in IIWAM operations expands.

IIWAM originators can leverage their large numbers to intimidate parties expressing views contrary to theirs. Most ordinary citizens are easily identifiable through publicly available information and thus anyone can reach them. Critical public postings often generate a flood of personally abusive and threatening but anonymous communications to the poster. Such

communications can be psychologically intimidating to the poster and inhibiting to others who might otherwise express their views. In some cases, posters have had their physical safety threatened.

Disintermediation helps the IIWAM originator. Those who use the online equivalents of traditional information intermediaries and rely on their editorial services to cope with the information deluge have at least some tools to cope with some IIWAM operations because they continue to be exposed to useful and factual information from multiple points of view. But those who rely on social media and search engines to filter the information ocean are less likely to be exposed to information that contradicts their prior beliefs. These users are exposed preferentially (or almost exclusively) to information that conforms to their own individual predilections, and hence they reinforce their existing confirmation biases.

Today's information environment enables crowdsourcing—the use of large numbers of individuals acting in loose cooperation and often without central guidance to achieve certain purposes. IIWAM originators can draw on the cooperation, witting or unwitting, of individuals whom they have been successful in influencing. In many instances, it only takes a retweet or a "like" to achieve a many-fold amplification of the message embedded in an IIWAM operation that has influenced an individual.

Because IIWAM operations can easily cross borders, IIWAM operators can take advantage of different laws in different geographic regions, engaging in IIWAM operations targeted against one national jurisdiction from the comparative safety of another jurisdiction that allows such behaviour. In addition, IIWAM originators can operate from the territories of their target nation with minimal infrastructure and gain protective benefits that the target nation confers upon its residents.

The easy availability of multiple distribution points gives rise to automated social chatbots that can be used in IIWAM operations. A social chatbot is a computer programme that generates content for and interacts with human users on social media but conceals its identity as a non-human entity. Chatbots have had a measureable impact on political dialogue.[34]

Lastly, IIWAM operations can exploit weak information security. Such operations can obtain information meant to be confidential or forge or alter print, audio, and video documents. The products of these operations can then be disseminated strategically to support the IIWAM originator's objectives. An example of this approach was the Russian hacking operation conducted in 2016 to access confidential emails of the Democratic National Committee and key staffers of Hillary Clinton's campaign.

4. An Exemplar Practitioner of Information/Influence Warfare—Russia

In the lead-up to the U.S. presidential election of November 2016, the American media audience was barraged by a display of confidential information and correspondence stemming from hacked private and organizational emails and other records, most notably from the

Democratic National Committee (DNC) and John Podesta, a key member of the Clinton campaign. After months of speculation concerning Russian involvement in the hacking which led to the release of private documents and data on the sites WikiLeaks, DCLeaks, and Guccifer 2.0, in early October the Obama administration formally announced its belief that the Russian Federation was behind the disclosures and that these were intended to interfere with the U.S. election cycle.[35]

For those familiar with Russian politics, the strategic release of "compromising material" concerning political rivals does not appear so unusual, with so-called "kompromat" having been used to tarnish reputations and undermine opponent messages for years. Recent Russian examples have included leaked recordings of private phone conversations by the opposition leaders and video footage of prominent critics in bed with prostitutes. The international deployment of such a tactic to influence the domestic politics of another country, while a little more novel, likewise draws upon a rich history of Russian military strategy and is particularly exemplary of recent developments in Russian military strategic thinking.[36]

[Note that Russia is not at all the only practitioner of IIWAM. A planned revision of this paper will address its use by the Islamic State and the alt-Right in the United States.]

4.1 The Russian Art of Strategy

Russia has long excelled at some aspects of the use and manipulation of information discussed in this paper. Soviet era theories of "reflexive control," cybernetics, and "maskirovka" – focusing on the use of information, deception, and psychological manipulation have influenced the development of current approaches to military strategy.

In recent years, Russia has further refined an explicit strategic approach to the use of IIWAM campaigns to achieve political and military goals at home and abroad. Asymmetry, ambiguity, indirect or deniable actions, and sophisticated information campaigns have become integral components of the country's military strategy – exemplified by what has been described as "next generation warfare" or the "Gerasimov Doctrine."

Elements of this strategy have been evident since Russian conflicts with Estonia (2007) and Georgia (2008), and have grown increasingly apparent in the Russian handling of the Crimea Annexation and ongoing conflict in Ukraine, the Russian involvements in the Syrian civil war, and Russian meddling in the U.S. election in 2016.[37] Aspects of the same approaches have likewise been used against protest movements, opposition leaders, and independent media within the country's own domestic sphere.

Explicit formulations of the current turn in Russian military doctrine have emerged over the last few years, indicating a period of significant strategic thought concerning the role of information. In a December 2013 article in a professional military journal, chief of the general staff, General Valery Gerasimov, laid out a vision of the current geostrategic and military-technological challenges facing Russia, perceived threats, and potential strategic adaptations to respond to these global challenges.[38] The article, which focused particularly on the novel type

of threat posed by events such as the Arab Spring and the Colour Revolutions in states of the former Soviet Union, suggested that the rules of war and the relationship between overtly military and non-military "means" in "achieving political and strategic goals" had changed and that Russia's own approach must also adapt to these new forms of "modern warfare." "The focus of applied methods of conflict," Gerasimov explained, "has altered in the direction of the broad use of political, economic, informational, humanitarian, and other non-military measures – applied in coordination with the protest potential of the population." These non-military measures were, in turn, to be "supplemented by military means of a concealed character, including informational conflict and the actions of special operations forces."

The Gerasimov Doctrine speaks explicitly of the need to find and exploit vulnerabilities even of the most militarily powerful opponents. "We must not copy foreign experience and chase after leading countries," he argued, "but we must outstrip them and occupy leading positions ourselves." He describes the use of "information spaces" as playing a critical role in this process, "[opening] wide asymmetrical possibilities for reducing the fighting potential of the enemy."

The doctrine stresses the importance of "cognitive-psychological forms of influence" in addition to "digital-technological" mechanisms,[39] that is, information/influence war in addition to what we understand in the West as cyber war. These tools are likewise to be applied regardless of binary distinctions between wartime and peacetime, being used to shape perception, deter, delay, or compel opponent actions, and influence perception, combined with special operations, and diplomatic and economic forms of influence, as well as nuclear and conventional military deterrence, but preferably reducing the need for outright use of military force to achieve desired strategic goals.

Adamsky argues that "it is difficult to overemphasize the role that Russian official doctrine attributes to the defensive and offensive aspects of informational struggle in modern conflicts," a point reinforced by Gerasimov's view that the appropriate ratio of non-military to military operations is 4 to 1 (i.e., the former is of greater importance than the latter).

As a strategy of influence, rather than of brute force, Russia's current next generation warfare approach both deemphasizes kinetic force and relies heavily on the "information struggle" as a core component of successful military campaigns. It can likewise be used against both individual actors and organizations and even entire populations within opponent countries, internationally, and at home. In a turn modelled upon Western use of soft power and public diplomacy for the promotion of democratic values, the strategy seeks to shape and leverage popular opinion and protest potential in targeted populations as one lever in achieving strategic influence on rival countries.

4.2 IIWAM In-Action: Russian Annexation of Crimea

Russia's 2014 annexation of Crimea from neighbouring Ukraine demonstrates the country's developing approach to the use of IIWAM in conflict and pre-conflict situations.

Integrated campaigns of media and social media coverage sought to influence public opinion on the topic, both in Russia and Ukraine and the international community. Special operations and false flag or unattributed actions (black and grey operations) involving "polite people" and "little green men"[40] were paired with official denial of Russian military involvement, causing other countries to pause before attributing the source of personnel and weapons observed in Crimea and other regions of Ukraine experiencing protest and violence. As the question of attributing actual Russian military involvement loomed, in the face of official Russian denials, there was also uncertainty as to whether any Russian actions in Crimea or Ukraine more broadly rose to the level of acts of war to which some international response might have been appropriate.

As the events in Crimea were orchestrated as a rapidly unfolding peaceful protest for independence and referendum concerning the region's return to Russia, Russian media coverage and diplomatic rhetoric emphasized the democratic nature of the transition (denying comparisons with prior infamous land grabs in European history). Meanwhile, lacking absolute certainty as to the nature of the threat or absolute binding security arrangements with Ukraine, Western states that had stood in solidarity with the Maidan protesters and rebuked Russian aggression stalled, concerned over escalating the crisis. By the time the nature of Russian activities in Ukraine became clearer, the annexation of Crimea was a fait accompli.

Domestic and regional Russian media coverage and viral social media during the crisis played on the emotions and biases of particular populations, emphasizing the "Russianness" of the local Crimea population, the supposed threat of violence towards Russian speakers in the region, and the role of soldiers as peacekeepers protecting the Russian-ethnic population from the menace of Ukrainian nationalist extremist violence. Coverage varied from the plausible to the implausible (such as a story describing the crucifixion of a three-year old Russian toddler), but was artfully mixed with real stories and footage. Nightly news footage showed long caravans of trucks bringing "humanitarian aid" to the beleaguered regions, and Western resistance to such efforts were portrayed as an effort to obstruct assistance to fellow Russians facing ethno-national oppression and atrocities.

While the irredentist logic of the land grab was less acceptable to Western audiences, other arguments were emphasized in international statements and media output, relying upon the rhetorical tactic of "what-about-ism" where Crimea's "protection" was compared with US or NATO-led efforts in Kosovo or Libya, and emphasizing the illegitimacy of the "coup" that had recently displaced democratically-elected President of Ukraine, Viktor Yanukovich, placing Crimea (and the rest of Ukraine) under supposedly illegitimate and anti-Russian rule.

5. Vulnerabilities of Liberal democracies to IIW

Liberal democracies are particularly vulnerable to IIWAM for a number of reasons. First and foremost, liberal democracies are inherently open societies, at least by comparison to many of the other nations of the world. They make available to their publics more information about their societies, and that information tends to be more truthful and accurate. They have media outlets for carrying information to the public that are more independent than in authoritarian nations. Most importantly, they are subject to periodic, peaceful regime change according to

the outcome of popular elections. Elections and political campaigns are thus particularly lucrative targets for IIWAM operations.

Democracies are willing to do certain things in war that they are unwilling to do in peacetime and vice versa. Law, regulation, and societal institutions (both government and nongovernment) are often organized around this distinction, and thus democracies must make explicit decisions about transitioning between the two. They do not do well (and often do not take decisive action) in responding to hostile actions taken against them that fall below the threshold of war—and IIWs are just such actions. By contrast, authoritarian states that believe in a continuous struggle with other nation states do not organize themselves this way and are able to develop institutions that operate in an integrated manner and with equal facility and authority across these conditions.

Democracies also tend to believe in the rule of law. For example, the United States operates under the auspices of the First Amendment to the U.S. Constitution, which guarantees freedom of speech and expression against government intervention except under very limited and specific conditions. Domestic political speech and expression receive the highest levels of protection, even when such speech is factually inaccurate and inflammatory. And governments generally do not assert extraterritorial control over content hosted outside their borders.

Another exacerbating factor within the U.S. government and especially within its military institutions is that information operations—deception, psychological operations, and so on—are somehow considered less important because of its unchallenged traditional military strength. For example, Steven Metz observes that "the American military is not as strong at psychological precision [i.e., psychological operations] as it should be, in part because technological advantages appear to make psychological effectiveness unnecessary."[41]

Such sentiments are at least suggestive of a public reticence towards IIWAM operations, at least by the United States. But irrespective of policy judgments about whether such operations are appropriate or helpful against adversaries of the United States, so-called mirror-imaging of an adversary—attributing to an adversary our own values and sentiments—may well contribute to an insensitivity and lack of awareness of adversary efforts in this regard.

6. Responding to IIW

Citizens in modern societies live an IT-enabled information deluge. A fast-moving information deluge is the ideal battleground for using IIWAM. Rapid information flow gives recipients (i.e., the targeted populace) little time to process and evaluate new information. Large volumes of information are cognitively disorienting and can be confusing. Opportunities for emotional manipulation abound.

Any coherent response strategy to IIWAM involves two critical elements: identifying IIWAM when it is in use and taking action to counter it or its effects.

6.1 Identifying IIWAM as It Occurs

One of the most insidious effects of IIWAM is that words and images do not have the same kind of obviously destructive effect on a society as do kinetic weapons or even cyber weapons. Indeed, successful IIWAM operations by actor X against society Y should be able to persuade large segments of society Y that X is not their adversary.

One point of departure for recognizing IIWAM operations is knowing the parties that have something to gain from them. As described above, Russia has adopted an approach to conflict that emphasizes IIWAM as a domain of strength. But nonstate actors such as the Islamic State also demonstrate high degrees of media sophistication in promulgating their messages and advancing their causes. Even political movements have caught on to the power of IIW, as one can see in the rise of the alt-right in the United States and Europe. Since the Internet and cyberspace point the way towards a much more powerful IIW, cyber-enabled IIWAM is a useful tool for many different types of adversary and a useful instrument for political combat and competition.

A second characteristic of IIWAM operations are efforts to undermine the legitimacy of the institutions that provide societal stability and continuity. In normal times, citizens argue over politics and the meaning of various events. Under IIWAM attack, citizens do not even agree on the events that have happened—each side has its own version of the facts to drive their own narratives. IIWAM also attacks institutions, such as established media outlets that adhere to journalistic standards and ethics, that seek to inform the public.

A third signal could be the automated detection and identification of IIWAM weapons in use. For example, the rapid emergence of large numbers of automated social chatbots all promulgating with similar political messages could signal the start of a concerted IIWAM campaign. Research is underway to identify such chatbots automatically.[42]

6.2 Countering IIWAM

As noted above, users that have abandoned traditional intermediaries (and their online equivalents) tend to be exposed preferentially (or almost exclusively) to for information that conforms to their own individual preferences. These individuals are not what the Founding Fathers of the United States had in mind when they placed their trust in a well-informed citizenry.

Since these parties are the most likely targets of IIW, what can be done to protect them when they do not know they are being targeted and have no particular wish to be protected from IIWAM operations that reinforce their prior beliefs and attitudes?

It is instructive first to consider some ideas that are nevertheless unlikely to help very much. For example, "naming and shaming" is probably ineffective against many nation states conducting IIW, especially those that have chosen to engage in international relations in ways that are not consistent with the behavioural norms of liberal democracies. Nor is naming and shaming effective against parties that engage in white IIWAM operations.

The U.S. response to the Soviet use of IIWAM operations in the Cold War—the United States launched Radio Free Europe/Radio Liberty and Voice of America to provide alternative information sources to those behind the Iron Curtain—is another model. These broadcast services operated as independent journalism outlets providing truthful information generally unfiltered by the U.S. government, though of course they were not seen that way by the Soviets.

But it is hard to imagine such an approach helping very much today. One reason is that the target audiences of IIWAM are today often the liberal democracies, where individuals have—and are supposed to have—considerable freedom as well as the legal right to choose their own information sources. Any approach to countering IIWAM will have to refrain from exercising government control over private-sector content provision.

Also, the velocity of information flow gets in the way of thoughtful reflection. Russia, foreign terrorist groups, and extreme political movements use of cyber-enabled IIWAM that encourage and celebrate the public expression of raw emotion—anger, fear, anxiety—and thereby channel powerful destructive and delegitimizing forces against existing institutions such as government and responsible media. Moreover, users of IIWAM are under no obligation to be consistent in their messaging, which means that they can promulgate messages much more rapidly than if they had to ensure consistency. Against this rapid-fire information deluge, the pace of communication vehicles operating during the Cold war would be completely inadequate in countering the hostile narratives offered today.

A second important reason is that any effort to coordinate and synchronize government-wide communications will take time. The desire for government-wide coordination is understandable—IIWAM operations benefit from consistency, and uncoordinated responses may well be mutually inconsistent. But rapid response—made especially important because responding to adversary IIWAM operations is by definition reactive—is arguably incompatible with coordination through an entity as large as a national government. If so, rapid government responses to adversary IIWAM operations will almost certainly have to be grey in nature rather than white.

On the citizen side, efforts to improve civic participation and engagement are always important to pursue. But the scale of the effort needed to move the needle towards thoughtful and informed civic engagement is enormous, especially in light of the fact that people are known to resist the absorption of knowledge and information that disturbs their prior beliefs about the world.

Consider, for example, the phenomenon that people are generally predisposed to believe in ideas that they hear, and reject them only after exerting mental effort to evaluate the ideas.[43] Rapp has found that encouraging the retrieval of accurate knowledge during reading can reduce the influence of misinformation;[44] however, such retrieval is effortful and individuals are less likely to undertake such effort if left to their own devices. Thus, if refuting a lie requires that the lie be repeated, refutation may well backfire since the repetition of the lie may well reinforce it.

Research on the psychology of communications suggests that people can be "inoculated" against fake news. Such inoculation consists of simultaneous delivery of an initial message and also a pre-emptive flagging of false claims that are likely to follow and an explicit refutation of potential responses.[45]

This is easier said than done, however, and many other common-sense techniques to reduce reliance on misinformation apparently offer even less promise.[46] Individuals warned about the potential falsity of a statement are not less reluctant to subsequently rely on that statement subsequently. Waiting so that people can no longer easily recall misinformation also does not help, as the reliance of many readers on misinformation increases over time. Presenting materials more slowly and decreasing the complexity of text content, both of which should reduce processing burdens that can impede careful evaluation, do not help substantially either.

As for the private sector, some major private sector actors have indeed acknowledged a degree of responsibility to counter certain kinds of IIWAM operations. For example, Facebook is deploying a new protocol for its users to flag questionable news sites. Google bans fake news web sites from using its online advertising service. Twitter, YouTube, and Facebook shut down accounts that they determine are promoting terrorist content.

Many argue that such measures are helpful but inadequate to stem the rising tide of misinformation conveyed through cyber-enabled IIWAM. For example, a recent Facebook letter from Mark Zuckerberg states that "Our approach will focus less on banning misinformation and more on surfacing additional perspectives and information, including that fact checkers dispute an item's accuracy."[47] But one must wonder about the value of the latter approach given the cognitive biases and requirements for effortful mental processing described above.

Others would advocate more intrusive or aggressive steps, such as cutting off prominent users that are "obviously" disseminating misinformation. The interaction between private companies and users is generally governed by the Terms of Service (TOS) agreement rather than by law—for the most part, private companies have no legal responsibility to protect the expression of all points of view. So far, goes the argument, these companies have interpreted TOS agreements narrowly, so narrowly that a lot of misinformation and inflammatory rhetoric does flow because their enforcement efforts are inadequate. But these private companies also respond to shareholder and advertiser concerns, and in the end, quite properly intend to make a profit from their efforts—and that profit generally increases as more people generate more message traffic. What is "obviously" misinformation to one user may not be obvious to others, and broad interpretations of TOS agreements run the risk of antagonizing a large part of their customer base, with all the financial consequences that such action might entail.

To sum up, some of the approaches described above have some promise of having some valuable defensive effect against IIWAM. But taken as a whole, the discussion of this section suggests that there are no good comprehensive solutions for countering IIWAM in free and democratic societies. Development of new tactics and responses is therefore needed.

7. Conclusion

IIWAM is one of the oldest forms of conflict known to humanity, and democracy itself has an ancient pedigree as well. In its older forms, democracy has rested on an underlying foundation of an enlightened, informed populace engaging in rational debate and argument to sort out truth from fiction and half-truth in an attempt to produce the best possible policy and political outcomes.

But even before Twitter and Facebook and the World Wide Web, the match between this idealized view of democracy and reality has been has been questioned by a number of scholars.[48] And if the match between ideal and reality was not entirely perfect in those days, today's information environment and cyber-enabled IIWAM have certainly rendered it much more questionable. The institutions of democracy are also poorly adapted to dealing with IIWAM operations, and especially cyber-enabled IIWAM operations because of their speed and reach.

Cyber-enabled IIWAM is a new kind of threat to democratic nations—a threat that evades established laws and conventions and turns the strengths of democracies, namely their openness and guaranteed freedoms, against them. In this regard, the threat from IIWAM is much like the threat from traditional cyber weapons that affect the confidentiality, integrity, and availability of information and information systems—cyber weapons pose a greater threat to nations that are more advanced users of information technology than to less-developed nations.

Lastly, it is worth noting that the cyber aspect of cyber-enabled IIWAM is critical, but it need not to be particularly sophisticated for cyber-enabled IIWAM to be effective. Cyber-enabled IIWAM takes advantage of fundamental characteristics of modern information technology—namely, vulnerabilities that will always be present in any kind of information technology regardless of sophistication—and that in turn allows the IIWAM attacker to control larger forces that have little to do with cyber per se. The significance of this point is that wherever good responses to IIWAM are to be found, a better, stronger, and more robust cybersecurity posture per se is not likely to be much help.

REFERENCES

Dmitry Adamsky, *Cross-Domain Coercion: The Current Russian Art of Strategy,* Institut Français des Relations Internationales, Paris, France, November 2015, http://www.ifri.org/sites/default/files/atoms/files/pp54adamsky.pdf.

John Banas and Stephen A. Rains, "A Meta-Analysis of Research on Inoculation Theory," *Communication Monographs* 77(3): 281-311, September 2010.

Jonathan Baron, *Thinking and Deciding*, 4th Edition, Cambridge University Press, Cambridge, United Kingdom, 2007.

Aaron Benjamin et al, "The Mismeasure of Memory: When Retrieval Fluency Is Misleading as a Metamnemonic Index," *Journal of Experimental Psychology: General* 127(1): 55-68, 1998.

Alessandro Bessi and Emilio Ferrara, "Social bots distort the 2016 U.S. Presidential election online discussion," *First Monday*, [S.l.], nov. 2016. ISSN 13960466. <http://journals.uic.edu/ojs/index.php/fm/article/view/7090/5653>.

Adrian Chen, "The Agency," *New York Times Magazine*, June 2, 2015,https://www.nytimes.com/2015/06/07/magazine/the-agency.html

Emilio Ferrara et al, "The Rise of Social Bots," *Communications of the ACM* 59(7): 96-104, July 2016.

Leon Festinger, *A theory of cognitive dissonance*, Evanston, IL, Row & Peterson, 1957.

Melissa Finucane et al, "The Affect Heuristic in Judgments of Risks and Benefits," *Journal of Behavioural Decision Making*, 13:1-17 (2000).

Daniel Gilbert, "How Mental Systems Believe," *American Psychologist* 46(2):107-119, February 1991.

Keir Giles, *Handbook of Russian Information Warfare*, NATO DEFENSE COLLEGE, Rome, November 2016, http://www.ndc.nato.int/news/news.php?icode=995

William Hart et al, "Feeling Validated Versus Being Correct: A Meta-Analysis of Selective Exposure to Information," Psychological Bulletin 135(4): 555-588, 2009, http://psycnet.apa.org/journals/bul/135/4/555.pdf

Adolph Hitler, Chapter 6: War Propaganda, *Mein Kampf*, 1925, http://www.greatwar.nl/books/meinkampf/meinkampf.pdf.

Stanley Ingber, "The Marketplace of Ideas: A Legitimizing Myth," *Duke Law Journal* 1984(1):1-91, 1984, http://scholarship.law.duke.edu/dlj/vol33/iss1/1;

William J. Lynn III, "Defending a New Domain: The Pentagon's Cyberstrategy", *Foreign Affairs*, September/October 2010.

Dan Kahan, "The Expressive Rationality of Inaccurate Perceptions," *Behavioral and Brain Sciences*, forthcoming 2017.

Daniel Kahneman and Amos Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica* 47:263-291, 1979.

Daniel Kahneman, *Thinking Fast and Slow*, Farrar, Straus & Giroux, New York, 2011.

Mein Kampf, Chapter 10

Ziva Kunda, "The Case for Motivated Reasoning," *Psychological Bulletin* 108(3): 480-498, November 1990, http://psycnet.apa.org/psycinfo/1991-06436-001

Howard Lavine et al, "On the Primacy of Affect in the Determination of Attitudes and Behavior: The Moderating Role of Affective-Cognitive Ambivalence," *Journal of Experimental Social Psychology* 34: 398-421, 1998.

Neil MacFarquhar, "A Powerful Russian Weapon: The Spread of False Stories", *New York Times*, August 28, 2016, https://www.nytimes.com/2016/08/29/world/europe/russia-sweden-disinformation.html.

Randal Marlin, *Propaganda and the Ethics of Persuasion*, p 22, Peterborough ,Ontario, Canada, Broadview Press 2002

Steven Metz, *Armed Conflict in the 21st Century: The Information Revolution and Post-Modern Warfare*, March 01, 2000, page 78, Director of Research, Strategic Studies Institute, US Army War College, http://www.strategicstudiesinstitute.army.mil/pubs/download.cfm?q=226.

Richard Nisbett and Lee Ross, *Human Inference: Strategies and Shortcomings of Social Judgment*, Prentice-Hall, Englewood Cliffs, New Jersey, 1980

Statistia, Number of internet users worldwide from 2005 to 2016 (in millions), https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/.

Frans Osinga, Science, *Strategy and War: The Strategic Theory of John Boyd*, Eburon Academic Publishers, Delft, The Netherlands, 2005.

James Poniewozik, "Just Because It's Hacked, Doesn't Mean It's Important," *New York Times*, October 18, 2016, https://www.nytimes.com/2016/10/18/arts/wikileaks-hillary-clinton-hacked.html.

Anthony Pratkanis and Eliot Aronson, Age of Propaganda: The Everyday Use and Abuse of Persuasion, Henry Holt, New York, 2001, p. 11

David Rapp et al, "Reducing reliance on inaccurate information," *Memory and Cognition* 42(1): 11–26, January 2014, http://link.springer.com/article/10.3758%2Fs13421-013-0339-0.

David Rapp, "The Consequences of Reading Inaccurate Information," *Current Directions in Psychological Science* 25(4): 281-285, 2016

David Sanger and Charlie Savage, "U.S. Says Russia Directed Hacks to Influence Elections," *New York Times*, October 7, 2016, https://www.nytimes.com/2016/10/08/us/politics/us-formally-accuses-russia-of-stealing-dnc-emails.html.

Claude Shannon and Warren Weaver, *The Mathematical Theory of Communication*. University of Illinois Press, Urbana and Chicago, 1949.

Vitaly Shevchenko, "Little green men" or "Russian invaders?", *British Broadcasting Company*, March 11, 2014, http://www.bbc.com/news/world-europe-26532154.

Kate Sweeny et al, "Information Avoidance: Who, What, When, and Why," *Review of General Psychology* 14(4): 340-353, 2010, http://psycnet.apa.org/journals/gpr/14/4/340.html.

Kate Sweeny et al, "Information Avoidance: Who, What, When, and Why," *Review of General Psychology* 14(4): 340-353, 2010, http://psycnet.apa.org/journals/gpr/14/4/340.html.

Christopher T. Wonnell, "Truth and the Marketplace of Ideas," *UC Davis Law Review* 19(3): 669-728, Spring 1986;

Charles Taber and Milton Lodge, "Motivated Skepticism in the Evaluation of Political Beliefs," *American Journal of Political Science* 50(3): 755-769, July 2006, http://www.jstor.org/stable/3694247.

Zeynep Tufekci, "WikiLeaks Isn't Whistleblowing", *New York Times,* November 4, 2016, https://www.nytimes.com/2016/11/05/opinion/what-were-missing-while-we-obsess-over-john-podestas-email.html.

Amos Tversky and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science* 185(4157):1124-1131; 27 SEP 1974, http://science.sciencemag.org/content/185/4157/1124.

Sun Tzu, *The Art of War*,

Carl von Clausewitz, Michael Howard, Peter Paret, and Bernard Brodie. 1984. *On War*. Princeton: Princeton University Press

Aldert Vrij, *Detecting Lies and Deceit: Pitfalls and Opportunities*, John Wiley and Sons, West Sussex, England, 2008

U.S. Army, Appendix A, FM 3-05.30, Psychological Operations, Army Field Manual, 2005, https://fas.org/irp/doddir/army/fm3-05-30.pdf.

U.S. Department of Defense, *Joint Publication 3-13, Information Operations*, 27 November 2012, Incorporating Change 1, 20 November 2014, http://www.dtic.mil/doctrine/new_pubs/jp3_13.pdf.

U.S. Department of Defense, *Joint Publication 3-13.2*, *Military Information Support Operations*, 7 January 2010, Incorporating Change 1, 20 December 2011, https://publicintelligence.net/jcs-miso/.

Robert Weissberg, "The Real Marketplace of Ideas," *Critical Review* 10(1): 107-121, 1996, http://dx.doi.org/10.1080/08913819608443411.

Drew Westen, *The Political Brain: The Role of Emotion in Deciding the Fate of the Nation*, Public Affairs, New York, 2007, pp. 103-112.

Mark Zuckerberg, "Building Global Community," https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10103508221158471/.

[1] William J. Lynn III, "Defending a New Domain: The Pentagon's Cyberstrategy", *Foreign Affairs*,

September/October 2010.

[2] Carl von Clausewitz, Michael Howard, Peter Paret, and Bernard Brodie. 1984. *On War*. Princeton: Princeton University Press. p. 90.

[3] Carl von Clausewitz, *On War*, Chapter 1.

[4] Sun Tzu, *The Art of War*, Chapter 3.

[5] This definition is identical to the U.S. Department of Defense definition of the information environment, but the various dimensions of the information environment are somewhat different. See U.S. Department of Defense, *Joint Publication 3-13, Information Operations*, 27 November 2012, Incorporating Change 1, 20 November 2014, http://www.dtic.mil/doctrine/new_pubs/jp3_13.pdf.

[6] The definition of information in this context is the ordinary common-sense meaning of the term: information is "facts provided or learned about something or someone." As such, information is understood to have semantic content, i.e., humanly understood meaning. This definition is different from Shannon information; the latter refers to bit-encoded information and is devoid of semantics (see Claude Shannon and Warren Weaver, *The Mathematical Theory of Communication*. University of Illinois Press, Urbana and Chicago, 1949.

[7] This definition of IIWAM operations is almost identical to the definition of "military information support operations" found in See U.S. Department of Defense, *Joint Publication 3-13.2*, *Military Information Support Operations*, 7 January 2010, Incorporating Change 1, 20 December 2011, https://publicintelligence.net/jcs-miso/..

[8] Appendix A, FM 3-05.30, Psychological Operations, Army Field Manual, 2005, https://fas.org/irp/doddir/army/fm3-05-30.pdf.

[9] The advantages of orienting oneself to ground truth and then making decisions more rapidly than the other side are the foundation of OODA-loop theory, the combat paradigm in which one side in a conflict observes, orients, decides, and acts (and then repeating the cycle). The side that can execute this loop more rapidly usually gains significant advantages over the other side. See Frans Osinga, Science, *Strategy and War: The Strategic Theory of John Boyd*, Eburon Academic Publishers, Delft, The Netherlands, 2005.

[10] Richard Nisbett and Lee Ross, *Human Inference: Strategies and Shortcomings of Social Judgment*, Prentice-Hall, Englewood Cliffs, New Jersey, 1980, pps. 6-7.

[11] Amos Tversky and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science* 185(4157):1124-1131; 27 SEP 1974, http://science.sciencemag.org/content/185/4157/1124. A popularized version can be found in Daniel Kahneman, *Thinking Fast and Slow*, Farrar, Straus & Giroux, New York, 2011. The original Tversky and Kahneman article reports on the availability, representativeness, and anchoring heuristics.

[12] Melissa Finucane et al, "The Affect Heuristic in Judgments of Risks and Benefits," *Journal of Behavioural Decision Making*, 13:1-17 (2000).

[13] Kate Sweeny et al, "Information Avoidance: Who, What, When, and Why," *Review of General Psychology* 14(4): 340-353, 2010, http://psycnet.apa.org/journals/gpr/14/4/340.html.

[14] Daniel Kahneman and Amos Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica* 47:263-291, 1979.

[15] Aaron Benjamin et al, "The Mismeasure of Memory: When Retrieval Fluency Is Misleading as a Metamnemonic Index," *Journal of Experimental Psychology: General* 127(1): 55-68, 1998.

[16] See, for example, Jonathan Baron, *Thinking and Deciding*, 4th Edition, Cambridge University Press, Cambridge, United Kingdom, 2007.

[17] Leon Festinger, *A theory of cognitive dissonance*, Evanston, IL, Row & Peterson, 1957.

[18] See, for example, William Hart et al, "Feeling Validated Versus Being Correct: A Meta-Analysis of Selective Exposure to Information," Psychological Bulletin 135(4): 555-588, 2009, http://psycnet.apa.org/journals/bul/135/4/555.pdf, and Kate Sweeny et al, "Information Avoidance: Who, What, When, and Why," *Review of General Psychology* 14(4): 340-353, 2010, http://psycnet.apa.org/journals/gpr/14/4/340.html.

[19] Ziva Kunda, "The Case for Motivated Reasoning," *Psychological Bulletin* 108(3): 480-498, November 1990, http://psycnet.apa.org/psycinfo/1991-06436-001.

[20] Dan Kahan, "The Expressive Rationality of Inaccurate Perceptions," *Behavioral and Brain Sciences*, forthcoming, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2670981.

[21] Charles Taber and Milton Lodge, "Motivated Skepticism in the Evaluation of Political Beliefs," *American Journal of Political Science* 50(3): 755-769, July 2006, http://www.jstor.org/stable/3694247.

[22] Howard Lavine et al, "On the Primacy of Affect in the Determination of Attitudes and Behavior: The Moderating Role of Affective-Cognitive Ambivalence," *Journal of Experimental Social Psychology* 34: 398-421, 1998.

[23] Drew Westen, *The Political Brain: The Role of Emotion in Deciding the Fate of the Nation*, Public Affairs, New York, 2007, pp. 103-112.

[24] See, for example, Anthony Pratkanis and Eliot Aronson, *Age of Propaganda: The Everyday Use and Abuse of Persuasion*, Henry Holt, New York, 2001, p. 11

[25] Randal Marlin, *Propaganda and the Ethics of Persuasion*, p 22, Peterborough ,Ontario, Canada, Broadview Press 2002

[26] Chapter 6: War Propaganda, Mein Kampf, 1925, http://www.greatwar.nl/books/meinkampf/meinkampf.pdf.

[27] Mein Kampf, Chapter 10.

[28] Adrian Chen, "The Agency," *New York Times Magazine*, June 2, 2015, https://www.nytimes.com/2015/06/07/magazine/the-agency.html

[29] James Poniewozik, "Just Because It's Hacked, Doesn't Mean It's Important," *New York Times*, October 18, 2016, https://www.nytimes.com/2016/10/18/arts/wikileaks-hillary-clinton-hacked.html.

[30] Cf., Zeynep Tufekci, "WikiLeaks Isn't Whistleblowing", *New York Times,* November 4, 2016, https://www.nytimes.com/2016/11/05/opinion/what-were-missing-while-we-obsess-over-john-podestas-email.html.

[31] As digital forgery tools become more effective, the lack of useful "true statements" will become less important—forged documents containing exactly the right information will become available.

[32] Aldert Vrij, *Detecting Lies and Deceit: Pitfalls and Opportunities*, John Wiley and Sons, West Sussex, England, 2008.A second aspect of truth bias is that people are more likely to correctly judge that a truthful statement is true than that a lie is false.

[33]Statistia, "Number of internet users worldwide from 2005 to 2016 (in millions)," https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/.

[34] Alessandro Bessi and Emilio Ferrara, "Social bots distort the 2016 U.S. Presidential election online discussion," *First Monday*, [S.l.], nov. 2016. ISSN 13960466, http://journals.uic.edu/ojs/index.php/fm/article/view/7090/5653>.

[35] See, for example, David E. Sanger and Charlie Savage, "U.S. Says Russia Directed Hacks to Influence Elections," *New York Times*, October 7, 2016, https://www.nytimes.com/2016/10/08/us/politics/us-formally-accuses-russia-of-stealing-dnc-emails.html.

[36] A good single-article press account of Russian activities in IIWAM can be found in Neil MacFarquhar, "A Powerful Russian Weapon: The Spread of False Stories", *New York Times*, August 28, 2016, https://www.nytimes.com/2016/08/29/world/europe/russia-sweden-disinformation.html.

[37] See Keir Giles, *Handbook of Russian Information Warfare*, NATO DEFENSE COLLEGE, Rome, November 2016, http://www.ndc.nato.int/news/news.php?icode=995 and Dmitry (Dima) Adamsky, *Cross-Domain Coercion: The Current Russian Art of Strategy,* Institut Français des Relations Internationales, Paris, France, November 2015, http://www.ifri.org/sites/default/files/atoms/files/pp54adamsky.pdf.

[38] The original Gerasimov article can be found at http://vpk-news.ru/sites/default/files/pdf/VPK_08_476.pdf. A non-authoritative English translation of this article done by Robert Coalson can be found at https://www.facebook.com/notes/robert-coalson/russian-military-doctrine-article-by-general-valery-gerasimov/10152184862563597/.

[39] Dmitry (Dima) Adamsky, *Cross-Domain Coercion: The Current Russian Art of Strategy,* Institut Français des Relations Internationales, Paris, France, November 2015, http://www.ifri.org/sites/default/files/atoms/files/pp54adamsky.pdf.

[40] Vitaly Shevchenko, "Little green men" or "Russian invaders?", *British Broadcasting Company*, March 11, 2014, http://www.bbc.com/news/world-europe-26532154.

[41] Steven Metz, *Armed Conflict in the 21st Century: The Information Revolution and Post-Modern Warfare*, March 01, 2000, page 78, Director of Research, Strategic Studies Institute, US Army War College, http://www.strategicstudiesinstitute.army.mil/pubs/download.cfm?q=226.

[42] Emilio Ferrara et al, "The Rise of Social Bots," *Communications of the ACM* 59(7): 96-104, July 2016.

[43] Daniel Gilbert, "How Mental Systems Believe," *American Psychologist* 46(2):107-119, February 1991.

[44] David Rapp et al, "Reducing reliance on inaccurate information," *Memory and Cognition* 42(1): 11–26, January 2014, http://link.springer.com/article/10.3758%2Fs13421-013-0339-0.

[45] John A. Banas and Stephen A. Rains, "A Meta-Analysis of Research on Inoculation Theory," *Communication Monographs* 77(3): 281-311, September 2010.

[46] The techniques described in this paragraph are taken from David N. Rapp, "The Consequences of Reading Inaccurate Information," *Current Directions in Psychological Science* 25(4): 281-285, 2016; this paper also contains the original citations backing these claims.

[47] Mark Zuckerberg, "Building Global Community," https://www.facebook.com/notes/mark-zuckerberg/building-global-community/10103508221158471/.

[48] See, for example, Stanley Ingber, "The Marketplace of Ideas: A Legitimizing Myth," *Duke Law Journal* 1984(1):1-91, 1984, http://scholarship.law.duke.edu/dlj/vol33/iss1/1; Christopher T. Wonnell, "Truth and the Marketplace of Ideas," *UC Davis Law Review* 19(3): 669-728, Spring 1986; Robert Weissberg, "The Real Marketplace of Ideas," *Critical Review* 10(1): 107-121, 1996, http://dx.doi.org/10.1080/08913819608443411.