On Cyber-Enabled IWIO (Information Warfare and Influence Operations)

DECADAL SURVEY OF SOCIAL AND BEHAVIORAL SCIENCES FOR APPLICATIONS TO NATIONAL SECURITY

Board on Behavioral, Cognitive, and Sensory Sciences

National Academies

October 11, 2017

Cyberwar to date

- Usually: destruction, degradation, or disruption of important information technology artifacts, such as networked information systems and networks or computing devices embedded in weapons, might be central to attacks on society or its vital national interests.
  - High level "cyberwar"
    - cripple society as a whole
    - attack critical infrastructure
    - destroy weapons systems, attack the military using this weapon
  - Low level ~~"cyberwar"~~ (better understood as cybercrime)
    - drug dealing, child porn
    - Hacktivism
    - Credit card fraud
    - Theft of IP
  - Note: no one has any idea what "cyber war" means.
- Flaws in information technology (of design, of implementation) enable adversaries to prosecute cyberwar by using IT in ways it was never intended to be used.

A framing question for this briefing

What would Hitler have been able to do with the Internet?

Information Warfare and Influence Operations

- Information warfare and influence operations (IWIO): the deliberate use of information to confuse, mislead, and affect the choices and decisions that the adversary makes.
- IWIO, especially cyber-enabled IWIO, is hostile, but it is not warfare in any sense recognized under the United Nations Charter or the laws of armed conflict.
- Information warfare is as old as the history of conflict
- Note juxtaposition between "information" and "warfare"
  - "warfare": connotations of hard power, armed conflict, shooting war, kinetic weapons, death and destruction, transition between war and peace, Clausewitz
  - "information and influence": connotations of soft power, propaganda, persuasion, culture, social forces, sub-threshold, Sun-Tzu
    - "The supreme art of war is to subdue the enemy without fighting."

- IWIO has its own battlespace, strategy/theory of victory, operational art, operators
  - Battlespace of kinetic warfare is the 3-D environment (air, land, sea, space)
  - Battlespace of cyberwarfare is cyberspace.
  - Battlespace of IWIO is the Information environment also has three dimensions: physical, informational, and mental.
    - Physical: C2 systems and infrastructure that enable the creation of effects.
    - Informational: where and how information is collected, represented, processed, stored, disseminated, and protected.
    - Mental: the minds (cognitive) and hearts (emotional) of those who transmit, receive, and respond to or act on information.

Strategy and a Theory of Victory in IWIO

- Victory is when the adversary's political goals become aligned with those of the victor.
  - But NOT because of "capitulation" or loss of the ability to resist—on the contrary, the adversary is openly willing.
- Ultimate goal of IWIO: change adversary perceptions in the cognitive dimension of the information environment.
  - IWIO damages knowledge, truth, and confidence, rather than physical or digital artifacts (brainspace not 3-D space or cyberspace)
  - IWIO injects fear, anger, anxiety, uncertainty, and doubt into the adversary's decision making processes.
- Targets of IWIO: government agencies, political leadership and segments of society
  - No "noncombatants" that enjoy immunity from IWIO attack.
  - IWIO targets legitimizing institutions, e.g., government and other institutions that promote societal cohesion (e.g., schools, news media)

- Sowing of chaos and confusion in an adversary without apparent purpose is useful
  - Confused government and population unlikely to take decisive action quickly ➔ more freedom of action for IWIO user.
  - IWIO actions shape that make the information environment more favorable for actual operations if necessary.
    - Operational preparation of the information battlefield
    - May reveal targets of opportunity that can be exploited.
  - Confusion and discord in a nation damage its reputation on world stage.
- Words and images do not have the same kind of effect on a society as do kinetic weapons or even cyber weapons.
  - IWIO persuades, informs, misleads, and deceives to negate adversary military capabilities—adversary has capabilities but does not use them.
    - No known military response to Russian interference in election
  - Below “act of war” thresholds, hence more usable.
  - Successful IWIO operations of actor X are able to persuade large segments of the targeted society that X is not their adversary.
  - IWIO *can* provoke kinetic action, though it need not.

IWIO operations

- IWIO operations are mostly conducted outside the explicit context of military operations (e.g., when traditional military operations are not going on).
- IWIO operations are white, gray, black.
- The impact of IWIO operations can be significantly increased when they are used:
  - To channel or influence other preexisting forces in society. (e.g., economic forces, cultural forces, social forces, psychological forces, organizational or bureaucratic forces).
  - In an atmosphere of uncertainty and doubt, enabling faster OODA loops.
  - To exacerbate and deepen existing societal fractures.
- IWIO is not likely to be a supremely powerful instrument of conflict in the same sense as nuclear weapons. IWIO is decisive primarily when small margins matter a lot (e.g., in electoral contests)

A Typology of IWIO operations

- Propaganda operations
  - Convey often false information or true/false mix to large audiences to influence opinion, attitudes, and emotion in ways that help the originator.
  - attract broad public attention, provide the most simple formulations, appeal to emotions rather than reasoning, repeat continually. (Mein Kampf)
- Chaos-producing operations
  - e.g., trolls posting fake disaster messages without apparent purpose
  - Need not be consistent, thus high volume and rapid response possible. Serve primarily to disorient.
- Leak operations
  - Breaching secrecy of embarrassing or compromising information
  - Breach of secrecy enhances notoriety, draws public attention

IWIO is old as, but cyber-enabled IWIO is new

- **What cyber brings**
  - High connectivity.
  - Low latency.
  - Anonymity.
  - inexpensive production and consumption of information.
  - democratized access to publishing capabilities.
  - Many-to-many bi-directional communications.
  - Disintermediation.
  - Insensitivity to distance and national borders.
  - High availability of personal information.
  - Information insecurity.

- What the results are
  - high tempo of IWIO operations
    - Fast response to real-world events (no lawyers)
    - First mover advantages
    - Suppress adversary messages by drowning out
  - Large megaphones to small players
    - Fringe players no longer isolated
  - Lack of accountability
  - Intimidation
  - Echo chambers via social media
  - Regulatory arbitrage across national borders
  - Automated chatbots
  - Leaks of sensitive info spread far and wide

Some cyber-enabled IWIO weapons effects

- Creation of filter bubbles (e.g., automated Twitter accounts to amplify one-sided messages).
- Enables previously marginalized communities to find like-minded compatriots and gives them megaphones that are disproportionately powerful.
- Communication with large populations at low cost without accountability.
- Tailoring of political messages in a manner highly customized to narrow audiences.

Note well --

IWIO is not sophisticated cyberwar. IWIO takes advantages of the advertised features of information technology, rather than the flaws in information technology.

Why does cyber-enabled IWIO work?

- Answer: same psychology that underlies the transformation of neoclassical economics to behavioral economics – that is, cognitive and emotional biases in human beings
- Cognitive biases
  - human use of intuitive reasoning strategies vs analytical strategies (fast vs slow, System 1 vs System 2).
  - Consider heuristics: substitution of simple judgments for complex inferential tasks
    - Fast but more often wrong
  - Fluency bias: ease of processing information predisposes individual for positive response. Enables simplistic messaging (140 char Tweets w/o nuance).
  - Confirmation bias: preference for seeking and interpreting new information in ways that are consistent with their prior beliefs and decisions, and avoiding inconsistent information.
  - Illusory truth bias: perception of greater truth for statements that are heard more often. Drives repetition of simplistic messages.

- Emotional (aka motivated) biases: emotional investment denies benefits of rational consideration. Example: cognitive dissonance
  - avoid exposure to challenging information
  - seek confirming information.
  - Less rigor for preferred arguments, more critical of disliked arguments
  - Emotional stance towards political candidate often more important than his or her view about candidate's policies or the facts known about the candidate.

  IWIO seeks to stimulate the emergence of strong emotion (fear, ethnocentrism, and pride), immunizes targets to real information and rational consideration.

Policy significance of cyber-enabled IWIO

- Interest stimulated by recent reports of Russian interference in various democratic elections.
- Main consequence:
  - NOT that it influenced the election outcome (hard to prove in any case)
  - **Resulting amplification of political polarization,** which would have resulted regardless of which side won any of these elections.
- Easier to destroy than to create:
  - NOT to create a rationally coherent alternative worldview
  - Destroy the foundation for *any* coherent worldview.
    - An obvious weapon for Russia: our society may suck, but yours does too, and you're no better than us.
    - Note well: messaging contradiction and inconsistency are desirable rather than undesirable.
      - Destroys coherence
      - Easily delivered at high speed in large volume, overwhelming attempts to process it rationally.

U.S. vulnerabilities to cyber-enabled IWIO

- High degrees of polarization: hyperpartisanship and enemy-of-my-enemy-is-my-friend thinking.
- Porous cyber defensive posture
- First Amendment and belief in value of free speech
  - US policy constrains USG information operations that might mislead Americans.
- Societal belief in sharp lines between war and peace
- Ambivalence about engaging in conflict
- Denigration of “soft power” by professional military
- Inability to see ourselves as the rest of the world sees us.
- Professed commitments to “fairness” and seeing both sides.

Many of these vulnerabilities generalize to some degree to all of the liberal democracies.

Responding to cyber-enabled IWIO

- Identifying IWIO as It Occurs
  - Recognizing who has something to gain from IWIO
    - Russia
    - China
    - Islamic State
    - Extremist movements in Europe/US/elsewhere?
  - Identifying targets of IWIO and determining if these targets provide legitimacy and reliable information, societal stability and continuity. Patterns of attack may identify IWIO attacks in progress.
  - Detecting automated IWIO weapons in use, e.g., chatbots.

- Countering cyber-enabled IWIO
  - Recognizing what won't help
    - Traditional institutions that require coordination
    - "Smarter and better educated" people
  - Some things that might help a little bit
    - Drown out bad guy messages
    - Promote truth rather than refute falsehood
    - Increase tempo of operations with gray operations
    - Encourage private sector to address problems of fake news
    - Better cybersecurity
- Also - use of cyber-enabled IWIO difficult against authoritarian regimes:
  - Greater degrees of information control
  - Less access and means of communication
  - Relative disconnect between leadership and population
- Grand bargain possible?
  - Are we willing to stop doing things that they want stopped in return for their cessation of IWIO?

Bottom line: we need more and better counters to cyber-enabled IWIO.

For more information…

Herb Lin
Stanford University
650-497-8600 office
202-841-0525 cell
herblin@stanford.edu