EasyChair Preprint

№ 1809

Disinformation in the Cyber Domain: Detection, Impact, and Counter-Strategies

Ritu Gill, Judith van de Kuijt, Magnus Rossell and Ronnie Johansson

October 31, 2019

24th International Command and Control Research and Technology Symposium

'Managing Cyber Risk to Mission'

***Disinformation in the Cyber Domain: Detection, Impact, and Counter-Strategies*[1]**

Ritu Gill, PhD
Defence R&D Canada
ritu.gill@drdc-rddc.gc.ca

Judith van de Kuijt, MA
Netherlands Organization-Applied Scientific Research
Judith.vandekuijt@tno.nl

Magnus Rosell, PhD
Swedish Defence Research Agency
magnus.rosell@foi.se

Ronnie Johannson, PhD
Swedish Defence Research Agency
ronnie.johansson@foi.se

Abstract

The authors examined disinformation via social media and its impact on target audiences by conducting interviews with Canadian Armed Forces and Royal Netherlands Army subject matter experts. Given the pervasiveness and effectiveness of disinformation employed by adversaries, particularly during major national events such as elections, the EU-Ukraine Association Agreement, and the Malaysian Airlines Flight 17, this study assessed several aspects of disinformation including i) how target audiences are vulnerable to disinformation, ii) which activities are affected by disinformation, iii) what are the indicators of disinformation, and iv) how to foster resilience to disinformation in the military and society. Qualitative analyses of results indicated that in order to effectively counter disinformation the focus needs to be on identifying the military's core strategic narrative and reinforcing the larger narrative in all communications rather than continuously allocating valuable resources to actively refute all disinformation. Tactical messages that are disseminated should be focused on supporting the larger strategic narrative. In order to foster resilience to disinformation for target audiences, inoculation is key; inoculation can be attained through education as part of pre-deployment training for military, as well as public service announcements via traditional formats and through social media for the public, particularly during critical events such as national elections. Manually working with identified indicators of disinformation to monitor ongoing disinformation campaigns is a tedious and resource intensive task in the presence of fast flowing information in multiple social media channels. The authors discuss how such indicators can be leveraged for automated detection of disinformation.

[1] Gill, R., van de Kuijt, J., Rosell, M. & Johansson, R. (2019). *Disinformation in the Cyber Domain: Detection, Impact and Counter-Strategies*. Peer reviewed conference paper. To appear in Conference Proceedings of the 24th International Command and Control Research and Technology Symposium, Laurel, Maryland., October 2019.

1. Introduction

The information environment (IE) has become increasingly noise polluted, largely due to the development of the internet, more specifically, social media platforms. In this saturated IE and in the age of the Internet and social media, we live in a hyper-connected world, which has several implications; in particular, every individual can be a journalist without the necessary credentials or credibility. In other words, all individuals are sources of information, and conversely, all are consumers of information. Increasingly, the accuracy of information is becoming less relevant but it is the headline with the catchiest or most inflammatory title that garners more attention, or more 'clicks'.

The Internet provides a medium or platform to access and disseminate information and messages in near real time, expediting opportunities to coordinate humanitarian assistance or draw attention to injustices, such as #BlackLivesMatter. While there are numerous benefits of using the internet, adversaries have effectively weaponized the internet, and weaponized social media, for the purposes of information warfare. NATO's supreme allied Commander General Philip Breedlove referred to a country as having 'the most amazing information warfare we have ever seen in the history of information warfare' (White, 2016). Some countries have a coordinated, integrated, and well maintained information campaign as part of their broader military strategy (Paul, Clarke, Schwille, Hlavaka, Brown, Davenport, Porche & Harding, 2018). A well-known example of the weaponization of social media is the US Election (2016) in which adversaries interfered in several ways. Using sock puppets and bots, material was posted on divisive topics to foster discord, chaos, and confusion among US citizens. Ultimately, the purpose was to flood social media with messaging that would benefit the adversary's choice of US Presidential candidate. Trolls (i.e., individual intentionally spreading disinformation), bots (i.e., automated accounts spreading messages), and bot networks inundated social media with right-wing propaganda, while concurrently undermining the historically supportive Black vote for opposing Democratic Presidential candidate. Trolls planted claims on social media that the Democratic candidate received money from the Ku Klux Klan, alienating the Black vote. In addition, a Facebook page called 'Blacktivist' was created, attracting 4.6 million followers, highlighting that the democratic candidate did not value Black lives, pushing the Black vote to the Green party (Swaine, 2018).

In February 2018, Special Counsel Robert Mueller indicted Russian individuals and Russian organizations for interfering with the US elections in 2016 (Permanent Select Committee on Intelligence, 2018). Specifically, Mueller commented on his findings regarding the activities of the Internet Research Agency impersonating US citizens on the internet:

"Defendants, posing as U.S. persons and creating false U.S. personas, operated social media pages and groups designed to attract U.S. audiences. These groups and pages, which addressed divisive U.S. political and social issues, falsely claimed to be controlled by U.S. activists when, in fact, they were controlled by Defendants... a strategic goal [was] to sow discord in the U.S. political system, including the 2016 U.S. presidential election. Defendants posted derogatory information about a number of candidates, and by early to mid-2016, Defendants' operations included supporting the presidential campaign of then-candidate Donald J. Trump ("Trump Campaign") and disparaging Hillary Clinton…Defendants also staged political rallies inside the

United States, and while posing as U.S. grassroots entities and U.S. persons, and without revealing their Russian identities and ORGANIZATION affiliation, solicited and compensated real U.S. persons to promote or disparage candidates. Some Defendants, posing as U.S. persons and without revealing their Russian association, communicated with unwitting individuals associated with the Trump Campaign and with other political activists to seek to coordinate political activities."

Adversaries were able to accomplish powerful online influence operations in the US given the availability of manpower and technology required to conduct this type of information warfare, as well as being unrestricted by legal, ethical, and/or public policies and accountability. Hundreds of employees are able to post propaganda online using sock puppets, bots and bot networks, and work two twelve hour consecutive days (two days off in between), with pre-determined goals set per day of posting content on the internet: five political posts, ten non-political posts, and 150-200 comments on others workers' posts (Abrams, 2016). Not only has the US election been interfered with, but other Western democracies including France, UK, Germany, and the Netherlands. Given that this is Canada's election year (2019) the anticipation is that there will also be interference to create chaos, confusion, and tension within the Canadian population. Similar to the tactics employed for the US election, taking on personas via social media that look like and sound like Canadians, developing connections with other Canadians online, and then disseminating narratives within Canada are all possibilities. As observed in previous influence and information warfare, Canada may expect conservative, right-valued Canadian audiences will be targeted with nationalistic, anti-immigration and anti-Trudeau content, while liberals or left-valued Canadians will be sent messages and narratives that focus on anti-government conspiracies, corruption and indigenous rights.

In information warfare several deception and manipulation tactics are employed to achieve goals; such tactics include the use of disinformation (Paul et al., 2018). Disinformation refers to a deception technique that is based on the dissemination of untrue information with the *intention* to deceive, manipulate, and mislead (versus misinformation which the disseminator is unaware that the information is untrue and has *no intention* to deceive) (Pamment, Nothhaft, Agardh-Twetman & Fjallhed, 2018). The goal of disinformation is to divide, create discord, sow doubt, fear, and demotivate people. Disinformation is not new; it is part of an old playbook of active measures. The goal of active measures such as disinformation is not intelligence collection, but subversion (Pamment et al., 2018). The development of the internet, more specifically social media platforms, has created a new medium to disseminate messages and has expedited the ability to propagate disinformation. The internet provides fertile ground for disinformation and the manipulation of perceptions and attitudes (NATO Strategic Communications, 2016). Adversaries effectively use disinformation as a tool in information warfare to undermine the credibility of operational missions (Joint Air Power Competence Centre, 2017).

2. Disinformation as an Active Measure

Disinformation is widely and expertly employed by adversaries through organized, integrated, and coordinated approaches. Specifically, Russia has a top down approach; disinformation campaigns are coordinated and controlled by the highest levels of Kremlin leadership (Abrams, 2016). Disinformation dates back to the cold war, and was a widely practiced Soviet active

measure tactic, and included the manipulation of global media via planted stories (Abrams, 2016). Given the advent of the internet and social media, it is easier to propagate disinformation, for it to go viral, and dominate the information environment. Disinformation is effective with audiences as it has the "ability to inspire fear, disgust and surprise" (Pamment et al., 2018, p. 44); in other words, it is more penetrating than true stories, giving greater potential for going viral. Similarly, it is relatively easier to spread disinformation on the internet and social media given the interactive and engaging features (e.g., sharing, likes, commenting, reposting) compared to traditional forms of information avenues, such as newsprint (Pamment et al., 2018).

Weaponizing the internet and social media for disinformation purposes, adversaries have an extensive network of internet trolls, bots, and sock puppets to generate and spread disinformation, permeating all areas of the internet most likely to exert influence over people and their perceptions, specifically via Facebook and Twitter. Disinformation is not just planted on the internet, but takes a more sophisticated approach by the Russian Internet Research Agency, or troll factories; for instance, disinformation planted in an online news story is corroborated by adding in hyperlinks to 'experts', forged documents, and fake photos and videos, all in an effort to reinforce the believability of the planted disinformation. Even if countered and disinformation is shown to be false, the damage has frequently already been done in terms of planting doubt or creating confusion, giving rumours traction and making its way into public consciousness (Abrams, 2016).

Disinformation appears in various forms, including fabrication, manipulation, misappropriation, propaganda, satire, parody, and advertising (Pamment et al., 2018); in all cases, the intention is to purposefully deceive audiences. Fabrication refers to news that has no factual basis and is presented in a style to make audiences believe it is real. In order for fabrication to be believable it has to be based on some pre-existing narratives and or have the appearance of being legitimate. Building on fabrication is manipulation, which is more than disinformation that is presented in text as it includes the manipulation of video, photos, and audio which can be doctored or completely fake to deceive and support a false narrative, also known as deep fakes. Misappropriation is a form of disinformation that uses misleading context and false connections; for instance, using different aspects of information to contextualize an issue or individual to fit a specific narrative, or referencing sources that do not contain the actual information. Propaganda refers to "information created with the purpose to influence public perception or public opinions to benefit a public figure, an organization, or a government" (Pamment et al., 2018). Propaganda differs from other types of disinformation in that it is more overt in its purpose and focuses on larger, strategic narratives. Lastly, satire, parody, and advertising are meant to be entertaining ways to propagate disinformation via humour and exaggeration, making it a challenge for audiences to discern what is real and what is disinformation within the context of humour and parody.

Ben Nimmo, a UK based analyst and writer on European security issues and senior fellow at the Atlantic Council's Digital Forensic Research Laboratory, highlighted how Russia's disinformation propaganda relies on four tactics, which he refers to as 4D: dismiss, distort, distract, and dismay (Nimmo, 2015). Dismiss the facts and critics by denying or denigrating accusers, distort the facts to serve the Russian narrative by launching accusations elsewhere and turning attention away from Russia, distract with alternative narratives, and dismay the audience

to intimidate opponents. The Malaysian Airlines flight MH17, shot down from the sky by a Russian missile, killing all 298 people onboard provides an excellent example of Russia's distort and distract strategies. In order to cast doubt and confusion over any potential Russian involvement, the Russian government went further than simple denial; as per the 4D strategy, they engaged in the distortion and distraction strategy by promulgating stories such as the airline already had dead bodies prior to departure and dropped over Eastern Ukraine in an effort to frame Russia, or that the actual target was an airplane with Putin in it flying near MH17 (Ukrinform, 2018).

The goal of disinformation is to create social and political divisions, foster conflict between allies, discredit foreign governments and militaries as well as undermine societies confidence in those organizations, create confusion to delay government and military responses, foster mistrust of news sources, and distract audiences. Adversaries are experts at propagating online disinformation through various methods; imitating or replicating fake media in which real media is fully imitated but with subtlety changed content to contain disinformation. This approach is very believable as it is so close to truth but with minor, almost imperceptible variations. For instance, replicating the online news source website for The Guardian, a false news article on MI6 was implanted. The website style and text was nearly identical to that of The Guardian and it also contained a very similar domain name, changing the 'I' in Guardian to the Turkish character 'I' (Pamment et al., 2018). As a result, unsuspecting readers would believe they are reading the reputable news source they always read, but are in fact reading a very close imitation version of their news source.

Employing narratives with strong emotional content is also an avenue used to propagate disinformation. In order to create hostility toward visiting military's or NATO troops and undermine military alliances, the Russian government created fake rape stories across Germany, Lithuania, and the Ukraine highlighting how visiting military soldiers abducted and raped an underage girl, Lisa, also known as the soldier rape narrative (Andriukaitis, 2018). Disinformation that contains charged emotional details or 'shock value' tends to do more damage than disinformation that is not emotionally charged (Andriukaitis, 2018). Using the 'shock value' tactic in disinformation, the Kremlin also purposefully initiates negative events to spin the event to serve a specific narrative. For instance, in 2017 Russian television crews tried to bribe Swedish teenagers to torch cars and start a riot. This bribe occurred shortly after Trump publicly stated that Sweden's immigration policy was ineffective.

To be able to effectively counter disinformation, disseminated deceptive messages need to be identified, analyzed and possibly responded to. A useful tool for this purpose is formulating indicators, or clues to the message that is characteristic of disinformation. For instance, explicit reference to a certain dubious news outlet could be an indicator for disinformation, repetition of a well-known adversarial narrative could be another indicator. It is not required that an indicator unambiguously identifies a deceptive message, but it is meant to assist the analyst in discovering disinformation.

3. Present Study

Given the pervasiveness and effectiveness of disinformation employed by adversaries on target audiences, particularly during major national events such as elections, the EU-Ukraine Association Agreement, or the Malaysian Airlines Flight 17 (MH17), the current qualitative study assessed several aspects of disinformation including i) how target audiences are vulnerable to disinformation; ii) which activities are affected by disinformation; iii) what are the indicators of disinformation; iv) what are the techniques or methods to counter disinformation; v) how to foster resilience to disinformation in the military and society; and last, vi) final thoughts to reiterate or share any other insights not already raised regarding disinformation. Semi-structured interviews with subject matter experts in the Canadian Armed Forces (CAF) and Royal Netherlands Army (RNA) subject matter experts were conducted to gain a deeper understanding of disinformation.

3.1 Participants

Six CAF and twelve RNA participants (N=18), from various units and departments[2], ranging from practioners, analysts, and advisors to senior level leadership participated in this study, all of whom have subject matter expert knowledge and experience in the area of disinformation.

3.2 Materials

Participants were asked i) how are target audiences vulnerable to disinformation; ii) which activities are affected by disinformation; iii) what are the indicators of disinformation; iv) what are the techniques or methods to counter disinformation; v) how to foster resilience to disinformation in the military and society; and last, vi) final thoughts to reiterate or share any other insights not already raised regarding disinformation.

3.3 Procedure & Analyses

All participants were interviewed one-on-one, and gave permission to be audio-recorded for the interview. Prior to initiating the interview participants were asked to read an information letter describing the study. Audio recordings were transcribed and were stripped of any identifying information. Interviews lasted from 45 to 120 minutes. All study materials and methods used in this research were reviewed and approved by the DRDC Human Research Ethics Committee. Analyses of interviews were based on identifying the recurring responses or themes and are presented in the results section.

4. Results

i) How Target Audiences are Vulnerable to Disinformation[3] (RNA perspective)

Different viewpoints exist on who in the RNA is most vulnerable to disinformation efforts. Some of the participants argue that the older generation may be especially vulnerable. In general, this

[2] To protect the anonymity of participants' further details will not be divulged.
[3] Responses to this question for CAF participants are classified.

audience is less familiar and experienced with the use of social media and its information overload potential. As a result, they may be less able to assess what information is real or not. However, others argue that the most vulnerable group is the younger generation as social media strongly gained traction within this audience. They derive their daily news from online platforms such as Facebook and Twitter and use social media on a daily basis. According to some RNA members it is this group that is easily manipulated through social media. They are not only the greatest consumers of social media, but also lack the ability of critical thinking and verification of the content.

In addition, the distinction between strategic versus tactical-level vulnerabilities was also highlighted, concluding that it is the political level that sets the parameters of where, when, and how the Defense organisation operates and, therefore, a potential target of disinformation campaigns.[4] Specific military units and organizations that are involved within hybrid warfare (e.g. marines, Special Forces (SOF)) or with analysing data (e.g. Open Source Intelligence (OSINT)/ Social Media Intelligence (SOCMINT)) or gatekeepers (e.g. policy directorate) are also considered as more vulnerable to disinformation.

ii) Which Military Activities are affected by Disinformation

Disinformation was noted to affect several areas of military activities, including decision making, time and resources, and public perception. Disinformation can delay planning, responding, and ultimately decision making, creating decisional paralysis, typically at the strategic level. A well-known example of disinformation delaying planning and ultimately decision making is investigating the downing of Malaysian Airlines Flight 17 (MH17). Since the July 17, 2014 downing of MH17, Russian officials and state-sponsored media outlets have disseminated a wide variety of alternative theories attempting to avert blame (Ukrinform, 2018). For example, every time on the eve of an announcement or event from the joint investigation team the amount of disinformation activities increased. Debunked narratives promoted by Russian government officials and state-sponsored media outlets caused delays in the investigation. The contradicting and conflicting narratives did not correspond with the evidence found on the location of the incident which frustrated the progress of the investigation.

If the focus is on debunking a narrative that contains disinformation, time is spent on responding, planning, and ultimately having senior leadership spend effort on decision making that does not require their attention. This has been the case in the context of the downing of MH17. On July 21 2014, Frans Timmermans, a Dutch politician serving as First Vice-President of the European Commission, spoke to the UN Security Council and developed a counter-narrative to the Russian disinformation, that may not have produced the desired effect, potentially wasting time and resources.

Rather than focusing on how to debunk narratives forwarded by adversaries, the focus should be on militaries having their own compelling strategic narrative, reinforcing their credibility, which erodes the credibility of the adversary. The RNA needs to invest more in Strategic

[4] Next to discrediting foreign governments, other goals of disinformation campaigns that were highlighted by the RNA participants include: undermine society's confidence in their governments and militaries, disrupt relations between NATO and EU nations, and create social and political division, influence public opinion.

Communication (STRATCOM), allowing for a coherent strategy coordinated by all critical departments on strategic level. Identifying a military narrative and reinforce this in all communications is not being applied enough yet. When disinformation is detected, a decision must be made as to what to do about it. This automatically disrupts normal operations (e.g., the planning or conduct of operations), one of the goals of disinformation. In the CAF, a decision chart was developed by the Canadian Joint Operations Command, and informed by Strategic Joint Staff. This chart helps decision makers focus on whether disinformation should be responded to, optimizing the conservation of time, energy, and resources.

Similarly, disinformation can occupy time and resources of analysts and decision makers, wasting leadership's time. The more time leadership spends on reacting to disinformation, such as fake pictures of Russell Williams in women's lingerie implying what to expect from CAF personnel visiting Latvia, or news articles of Dutch officers (e.g. deputy battalion cdr.) visiting prostitutes demoralising Dutch units in Lithuania, the less time is spent on accomplishing tasks in theatre of operations. When suspected disinformation is published online, such as Sputnik or RT, it would take 16-24 man hours for analysts to discover if the story was legitimate or disinformation and what the potential impact of the disinformation, resulting in a significant amount of analysts' time and resources spent on verifying and debunking a story instead of doing their main task.

The impact of disinformation on public perception was also highlighted as key terrain for the CAF. Specifically, if disinformation campaigns damage public perception of the military among host and domestic audiences, then the battle is lost as the adversary is controlling what the public thinks and respects. Outside of leadership, analysts are not focused on the public image of the military, as they are more focused on specific tasks. However, disinformation campaigns on social media do impact public perception of the military, which in turn impacts perceptions of politicians. The politicians are the ones who set the parameters of where, when, and how the military operates. If disinformation campaigns are successful and a negative public perception of the military results, then parameters and restrictions will likely be set, determining what the military can and cannot do. Ultimately, disinformation campaigns that undermine the credibility and legitimacy of the military in host nations and the eyes of the public, and are not effectively countered can damage public perception of the military, potentially ending missions.

The same holds for the RNA; many examples of disinformation campaigns come from Eastern Europe. These campaigns aim to strengthen the dichotomy between different camps by negatively influencing the perception of the RNA among host and domestic audiences and, in turn, the Dutch political decision making process. The most significant example of disinformation campaigns directed at targeting the public image of the RNA is the enhanced Forward Presence (eFP) mission. The stories about attacks and rapes have had implications for the Dutch contribution to the mission. As a result parameters and restrictions have been imposed to determine what Dutch soldiers are allowed to do. For instance, Dutch soldiers are not allowed to consume alcohol, are made aware of so called honey traps, and are taught to be defensive in attacks. Moreover they are stimulated to conduct pro-active 'damage control'. These precautionary measures are largely a reaction to Russian information operations. They are taken to avoid any risk of unwanted media coverage of misbehaving soldiers (real or not) and

consequently undermine potential disinformation campaigns resulting in a negative public perception.

However, restricting the freedom of movement of NATO countries does not at all prevent the Russian government from releasing false stories about misconduct (Kamphuis, 2018). By restricting the movement and visibility of personnel, NATO countries are possibly more or less alienating their units from their environment. This development, in turn, might make it easier for Russia to continue its stream of negative coverage regarding NATO in the very same countries, because people have a tendency to fear or distrust anyone they do not know. Alienating NATO soldiers from their environment by imposing restrictions on freedom of movement could be exactly the outcome Russia has been aiming for all the time.

iii) What are the Indicators of Disinformation

Future disinformation activities can be anticipated. In line with this, indicators in the military and security domain are derived from the concept of warning intelligence. Indicators in this tradition help to detect disinformation activities to observe an immediate threat. The most significant indicator of disinformation is the source; specifically, the website and the individual writing the article. For instance, during Operation Reassurance, both Sputnik and RT published a story about Canadian soldiers being killed Ukraine, specifically, that Canadian NATO servicemen were active on the front lines and their car exploded after hitting a landmine. The first indication was the source – both Sputnik and RT are outlets for Kremlin propaganda, so their credibility for publishing legitimate information is limited. However, the disinformation was translated and picked up by Western media, and was shared over 4000 times, predominately on Facebook.

Notably, with disinformation campaigns for those who are well versed in detecting disinformation, or even being aware of disinformation, it is obvious what the indicators are. While for those with limited knowledge of disinformation and its indicators, headline news is often taken at face value.

Style of writing is also an indicator of disinformation, and requires closer examination to detect. Frequently, written articles will start with factual information as a way of gaining credibility with readers. Falsities and rumours will be weaved into the story, ending with factual information. Embedded within facts is false information that sways readers' reality and opinion, without realizing it. Similarly, 'experts' will be used to reinforce the credibility of disinformation, and in turn, the believability; for instance, citing a professor from an obscure university in Europe that no one has heard of. The use of hyperlinks is also frequently employed to gain credibility of disinformation. Frequently the hyperlinks lead to fake websites, or are dead links simply inserted to increase the believability of the disinformation. Again, the more experienced one is in disinformation the less likely the citing of experts is to influence opinion, however, those less aware of disinformation and its indicators are likely to take the citing of such experts at face value.

Humour is another tactic employed to propagate disinformation, particularly through the use of memes resulting in viral media warfare. For instance, in response to allegations from the British of Russian involvement in the Skripal poisoning, the Russians employed humour via literary

figures to debunk the allegations; for instance, stating that given the lack of evidence on Russia's involvement, Hercules Poirot or Sherlock Holmes are needed on the case. Similarly, memes can be used to counter political narratives in an effort to disseminate the truth or provide an alternative perspective on narratives adversaries are propagating. Ultimately, for adversaries, the goal is to introduce misleading narratives, create doubt, and confusion via humour. Deepfakes using photographs and videos may also indicate disinformation, and as technology advances in developing deepfakes, detection becomes more imperative. The Kremlin frequently doctor old photographs to spread disinformation to support a new narrative.

iv) What are the Techniques or Methods to Counter Disinformation

The optimal way to counter disinformation is to inoculate audiences, including a robust and reliable media system that can help spread the importance of inoculation. Media platforms have a responsibility to their users to raise awareness of disinformation conducted on their platform (Pamment et al., 2018). In order to effectively inoculate, it is critical for a vulnerable or target audience to be identified, and then it should be communicated that there will be disinformation campaigns appearing in one's Facebook feed and/or twitter. In such cases, if the disinformation appears extreme, it is likely to be false. However, if target audiences are still unsure, they should confirm with verified and trusted sources of information. In Latvia, for instance, defence journalists were briefed on new trends in the information space and were asked to be vigilant of disinformation campaigns or attacks. If journalists suspected disinformation they were asked to verify with the CAF before retweeting or reposting to audiences. As part of inoculation, teaching and encouraging effective information practices among audiences are critical, not only at schools, but also through, for example, public awareness campaigns. The Dutch Defence Communication Organisation (DCO) provides a manual on how to use social media, for instance on operational security. However, the use of social media is the responsibility of the brigades themselves and, as such, monitoring any threats within their own brigades. Furthermore, sharing the indicators of disinformation, unreliable websites and the importance of verifying information with trusted sources, especially during critical events such as national elections can help foster inoculation against disinformation. Employing 'influencers' or 'digital warriors' to help spread inoculation of disinformation and solid information practices can help audiences pause and reflect on processing online information. Such practices include increasing the awareness of echo chambers and their effect on the believability of online information.

Another effective response to disinformation campaigns is to ignore it; in fact, most of the time this would be the most effective action. Indeed, one of the strategies of DCO is to ignore disinformation instead of trying to debunk narratives forwarded by adversaries. It is critical to avoid the 'back and forth' responding. Such tactics tend to backfire, yielding opposite effects, providing the disinformation more attention, traction, and credibility. Given one of the goals of disinformation is to elicit a response from a target(s) by losing control and becoming distracted from one's own information campaign, ignoring can be a very effective response. If the military is reacting and responding to every piece of disinformation, vulnerability could be revealed. Adversaries would know they are having an effect and disrupting military planning efforts. However, if disinformation involves attacking the military's narrative, then communication should be focused on underscoring the core strategic narrative. At the strategic level, future conflict will likely be focused on a battle of narratives, and influencing the information space

with events occurring in an area of operation. For instance, what is the overall perception as to why NATO is in the Baltics? Is it because NATO is aggressive, or is it because NATO is reassuring people who have been threatened by Russian action? These opposing perceptions and narratives that are put forth by NATO and Kremlin is where the real battle resides, in the information domain. The tactical space of who said what becomes less meaningful when the bigger picture or the larger strategic narrative is framed. All other actions will be defined through that strategic lens.

v) How to Foster Resilience to Disinformation in the Military and Society

In order to foster resilience to disinformation among target audiences, and more broadly society, it is critical to inoculate. Educating people to be savvy information consumers, aware of the effects of echo chambers, and when disinformation campaigns are responded to the emphasis should be on the military's core strategic narrative. Furthermore, avoiding the pull into the tactical 'back and forth' which can distract from the focus of the larger strategic narrative is critical.

Military members should also be educated on disinformation via formal training. Personnel should also have the ability and mechanisms in place to capture a piece of disinformation and report it up the chain of command, especially if the disinformation goes against the core strategic narrative. In Lithuania, for example, Dutch soldiers were briefed on how to use social media and their smartphones in the Enhance Forward Presence (eFP) mission. Moreover, they were briefed to report the capture of any hybrid activity, such as disinformation, to avoid negative public opinion. However, formal training on the effects of disinformation and the methods of how disinformation can be used, for instance during pre-deployment, is not yet embedded in the Royal Netherlands Army and should be more invested in. Part of this education should be focused on internal and external communication to raise awareness of these issues. During deployment the military should identify the target audience and communicate from the outset that there will be disinformation campaigns appearing in one's Facebook feed or twitter.

As part of education, during pre-deployment training for military personnel include a series of lectures outlining the potential effects of disinformation, how soldiers can be a target of disinformation, the potential impact on his/her family, and how to protect themselves against disinformation. For instance, if personnel take their smart phones on deployment, it will likely be compromised. Part of inoculation involves being aware of the methods of how disinformation can be used against oneself, even if the situation appears to be innocuous it can be spun to appear differently. For instance, if there is a fight at the local McDonald's in an area of operation and military personnel were not involved but were present, they may be represented in the news as visiting military personnel are fighting with local populations. It is also important to set up precautionary measures, such as parameters and restrictions, determining what permissible on and offline activity in order to avoid unwanted media coverage.

Having control over military messaging is also important to foster resilience. By being proactive and disseminating messages, such as NATO has conducted with their eFP in Latvia to emphasize themes of reassurance and deterrence prior to adversaries gaining traction with their messaging, audiences are likely to experience the 'recency effect'. This occurs when people support messages and narratives that they tend to hear first and the most. By being proactive, controlling,

and disseminating messages prior to adversaries gaining traction, militaries can gain the advantage. It is important to not be seen as just responding to the adversary's messages or disinformation, but to focus on one's own message and reiterate it as much as possible.

vi) Final Thoughts to Reiterate or Share Regarding Disinformation

It is important to recognize that there are limits to what the military can know with respect to online activity, as it is impossible to monitor everything online. This negatively impacts the military's awareness, and in turn, responses, although militaries can frequently act effectively in an ad hoc manner. For instance, the Color Revolution in Northern Africa; nobody was aware that it was coming but it built up very quickly. Citizens do not just show up in the tens of thousands in the squares in capitals cities overnight without significant coordination, conversation, and organization. If this was not seen as coming, what else is happening that militaries do not see coming.

In terms of how long it takes for disinformation to be online before damage is done depends upon the context, specifically whether it is attacking one's core narrative. To estimate, 90% of disinformation has a very short lifespan and is relatively meaningless by itself, but there is also a cumulative effect where it can slowly erode the larger narrative and critical ground. Militaries need to be mindful that never responding and not reinforcing your own narrative can result in a cumulative effect, while concurrently not falling into the pitfall of overreacting and thinking that one disinformation attack necessitates a counter response.

As mentioned earlier, education and inoculation are important pieces of safeguarding and increasing resilience against disinformation. The Canadian public in general, more so the younger generation, who live on social media, are important to educate. The public needs to develop critical thinking skills and analysis and use it when information is consumed online. How is the best accomplished? It is beyond having a press information session in the House of Commons; while it may be impactful for politicians, it is less impactful and meaningful to the broader Canadian public. How can social media or other online platforms be used to disseminate information on education and inoculation to reach the public is the key question. Also for Dutch military personnel, if it is true that the younger generation is the most vulnerable group to disinformation than it is this group that needs to be educated and inoculated to increase their resilience against disinformation.

In addition, the focus should not be on the kinetic, but the non-kinetic, and allocate resources accordingly; NATO's presence in Latvia demonstrates this. Information is powerful and sometimes the reach, influence, and impact of information are underestimated. Similarly, align words with actions to ensure the consistency and continuity of the larger strategic narrative, an area that public affairs can address and contribute to the effort for targeting for strategic effects.

5. Discussion

The goal of disinformation is to create chaos, discord, confusion, and division. As part of Russian active measures, the Kremlin are experts at disinformation and adapting it to new forms of technology via social media platforms. If anything, the advent of the internet and social media platforms have expedited the ability for the Kremlin to propagate disinformation and foster confusion, discord and chaos, irrespective of geographical boundaries.

Subject matter experts have identified the implications disinformation can have on several military activities, including decision making, time and resources, and public perception. Indeed, part of Russian disinformation is to confuse militaries and governments to foster decision paralysis and create distraction from their actual intentions and goals. Similarly, by employing disinformation and impairing the military's decision making cycle, time and resources are dedicated to areas that do not need to be focused on. Ben Nimmo clearly describes how Russia's disinformation propaganda relies on four tactics, which he refers to as 4D: dismiss, distort, distract, and dismay, all of which serve to undermine the military's decision making, time and resources, and public perception. However, as highlighted by participants in this study, the focus should be on reinforcing the military's strategic narrative, which in turn serves to reinforce the military's credibility and erode the credibility of the adversary. Responding to or countering disinformation should be conducted with the core strategic narrative at the forefront; in other words, if disinformation does not undermine the larger strategic narrative, then countering is not necessary.

Future disinformation activities can be anticipated. In line with this, indicators in the military and security domain are derived from the concept of warning intelligence. Indicators in this tradition help to detect disinformation activities to observe an immediate threat. Indicators of disinformation highlight not only what militaries should examine, but also what the public should be aware of as consumers of information. Source, style, humour/memes and deepfakes emerged as significant indicators of disinformation from participants. Consumers of information need to be aware of such indicators in order to inoculate themselves to disinformation. As technology advances, techniques associated with disinformation also advance. For instance, online articles are designed to look like they came from reputable news sources, such as Le Soir (Belgium) or the Guardian. Unsuspecting consumers are likely to digest and propagate disinformation presented in this type of credible and legitimate looking imitation. Similarly, adversaries developed a fake British Broadcasting Corporation (BBC) website to spread disinformation on the Charlie Hebdo attacks in Paris in 2015. The original UK BBC website (http://www.bbc.co.uk/) was so subtlety altered that users would not observe the change (http://www.bbc-news.co.uk).

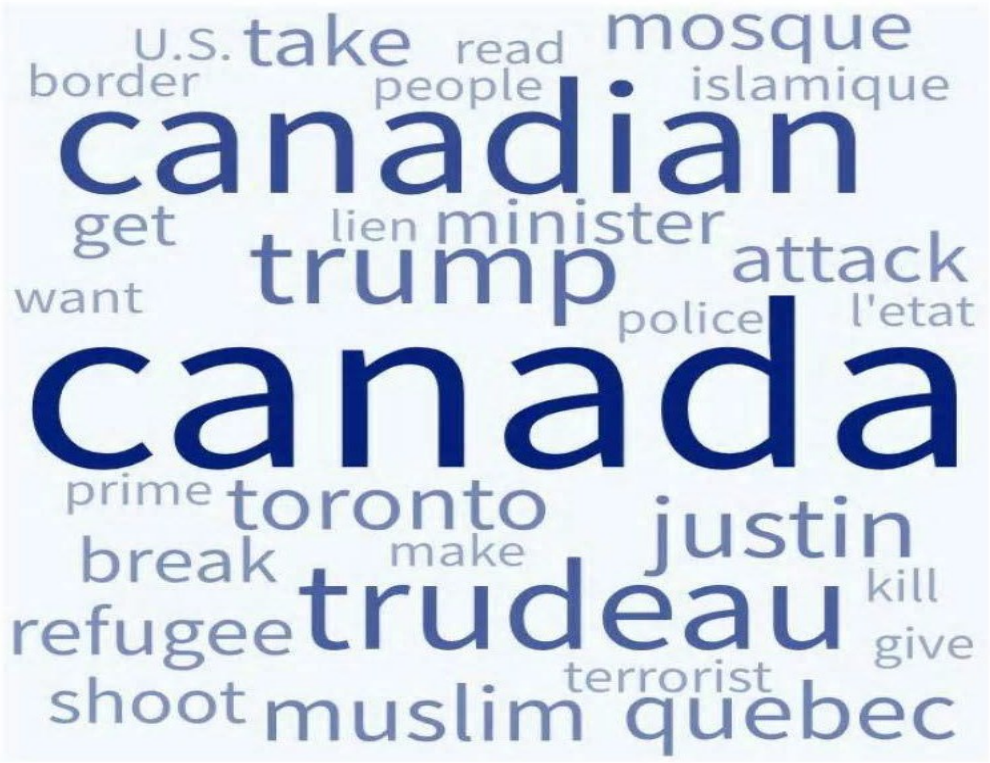
Manually working with identified indicators to monitor ongoing disinformation campaigns is a tedious and resource-intensive task in the presence of fast flowing information in multiple (social) media channels. Computational methods can be used to automate at least part of this work. There are many types of information available in the content and metadata of for instance social media which can be extracted using computational methods (Franke & Rosell, 2014). If the analyst has a toolbox of such indicators at hand the individual can use these to filter out the information that is important and/or interesting for a certain case.

The previous example with Operation Reassurance can serve as an example to explain some of the possibilities. In this case the analysts identified certain sources as being of special interest (e.g. limited credibility). A very simple computational method hence is to filter out only data that has been published by a list of known less credible sources. As the analysts investigate data they can add new sources they find (for instance from mentions of known sources) to the list. More advanced automatic methods for identifying interesting sources could also be applied.

Filtering data by the sources is one example of using metadata as an indicator. The actual content can also be handled directly using computational methods. The simplest possible content filter in our example is obviously the keyword “Operation Reassurance”. Texts containing that particular keyword are more likely to be interesting to analysts. There are many more advanced methods of filtering out what may be interesting textual information, for instance using methods for synonymous or related words. Regardless of which methods are used to filter data by both sources and content, the two can be combined to give analysts a smaller set of data that is more manageable to handle manually. The simplest being texts containing the keyword “Operation Reassurance” that are published by a list of interesting sources. With many different indicators and filters, huge amounts of data can be handled semi-automatically in this way.

Given the past history of the adversary’s interference in democratic elections, including the US, France, and Netherlands, foreign interference is anticipated in the 2019 Canadian elections. In fact, in late 2018 a former Russian troll warned Canada that they should be on guard for Kremlin interference in the upcoming Canadian elections (Semple, 2018). Indeed, in late 2018, the Canadian Centre for Cyber Security warned the Canadian public that state-sponsored actors can conduct influence operations employing sock puppets, or acting as Canadian citizens (The Canadian Press, 2018), similar to the sophisticated influence operations observed during the US election (2016) and France’s election (2017). Indeed, at the time of the writing of this report, over 9 million troll tweets from suspected foreign influence campaigns on divisive issues, such as pipelines and immigration in Canada, were posted (Rocha & Yates, 2019). The troll accounts were deleted by twitter, and are suspected to have originated from Russia, Iran, and Venezuela. Figure 4 highlights the 30 most frequent words used in troll tweets, with larger words indicating more frequent use. Based on Figure 1 it is clear that the troll tweets targeted divisive issues within Canada in an attempt to foster chaos, discord, and confusion. In anticipation of foreign influence on the 2019 Canadian election, the Canadian government is establishing a new ‘critical election incident of public protocol’ group to alert the public to potential threats during the campaign period (Tunney, 2019). In addition, a new security task force involving Canada’s intelligence agencies entitled ‘Security and Intelligence Threats to Elections’ (SITE) Task Force will collaborate to identify foreign threats and influence on Canada’s electoral process and aid the government in responding (Tunney, 2019).

Figure 1: Word Cloud Depicting Most Commonly Used Words in Troll Tweets



The results of the interviews also indicate that it is key to inoculate target audiences to disinformation, and participants suggest several methods to inoculate. Indeed, the EU and other Western governments are hoping that by alerting the public to disinformation, citizens will become inoculated (Emmott, de Carbonnel, Humphries, 2019). The Dutch government increased her investments in countering disinformation activities by launching online campaigns and broadcast campaigns against fake news. Finland is actively fighting disinformation via media literacy and critical thinking by developing a critical thinking curriculum in a Finnish high school. The high school partnered with a Finnish fact checking organization, Faktabbari (Fact Bar) to develop a digital literacy toolkit. Students are taught critical thinking skills specific to social media; for instance, prior to liking or sharing information, consider the source, where it was published, can the information be verified (Mackintosh, 2019). While democracies cannot teach the broader society the same curriculum as a classroom, the focus on fostering the ability of citizen to engage in critical thinking and media literacy is key to inoculate against disinformation. Similarly, educating older generations on media literacy is imperative as they were raised in an era where print media and other unidirectional media (such as radio and television) was the main source of trusted information.

A significant challenge in combatting disinformation is the illusory truth effect, or the reiteration effect, which in the psychology literature refers to repeated statements are more likely to be perceived as true (Fazio, Brashier, Payne & Marsh, 2015). This effect applies not only to political propaganda, but also consumer marketing and rumours. The illusory effect occurs even

when individuals know better (Fazio et al., 2015). A repeated statement becomes more believable than a new statement (Unkelbach, 2007; Unkelbach & Stahl, 2009). Disinformation propagated by adversaries indeed follows this approach of repetition. When disinformation is repeated on various social media platforms, it is reinforced as being true even when users may have prior knowledge to the contrary. To help inoculation against disinformation, by repeating the facts (*not* repeating the disinformation which may have the opposite effect of increasing its believability) democracies can in effect increase the chances of target audiences believing the facts, utilizing the illusory truth effect to their advantage. Similarly, as discussed in the results, adversaries effectively use images and memes to propagate disinformation, which are more impactful on target audiences than text (Hameleers, Powell, Van der Meer & Bos 2019). Allies can also leverage this information by employing the same approach to disseminating facts in order to help inoculate target audiences and empower them with the truth.

6. Conclusion

The current research results suggest that in order to effectively counter disinformation the focus needs to be on identifying the military's core strategic narrative and reinforcing the larger narrative in all communications. Tactical messages that are disseminated should be focused on supporting that larger strategic narrative. Furthermore, in order to foster resilience to disinformation for target audiences, inoculation is key; inoculation can be attained through education as part of pre-deployment training for military, as well as via robust and reliable media systems and public service announcements, particularly during critical events such as national elections.

References

Abrams, S. (2016). Beyond propaganda: Soviet active measures in Putin's Russia. *Connections: The Quarterly Journal, 15(1),* pp. 5-31.

Andriukaitis, L. (2018). Russia uses fake rape stories to create hostility to NATO troops. https://www.stopfake.org/en/russia-uses-fake-rape-stories-to-create-hostility-to-nato-troops. Accessed on February 27, 2019.

Baron, & Hershey (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology, 54(4),* 569-579.

Emmott, R., de Carbonnel, A., & Humphries, C. (2019). Who burned Notre Dame? Brussels goes after fake news as EU election nears. https://www.reuters.com/article/us-eu-election-disinformation/who-burned-notre-dame-brussels-goes-after-fake-news-as-eu-election-nears-idUSKCN1SM0LQ. Accessed on May 22, 2019.

Fazio, L.K., Brashier, N.M., Payne, K.B. & Marsh, E.J. (2015). Knowledge does not protect against illusory truth. *Journal of Experimental Psychology: General, 144(5),* 993-1002.

Franke, U. & Rosell, M. (2014). Prospects for detecting deception on Twitter. In *Proceedings of the first international workshop on social network analysis, management and security* (SNAMS 2014) held in conjunction with the 2nd international conference on future internet of things and

cloud (FiCloud-2014), 27-29 August 2014, Barcelona, Spain.

Hameleers, M., Powell, T.E., Van der Meer, G.L.A., & Bos, L. (2019). A picture paints a thousand lies? The effects and mechanisms of multimodal disinformation and rebuttals via social media. *Manuscript under review.*

Joint Air Power Competence Centre (2017). Mitigating disinformation campaigns against air power.

Kamphuis, C. (2018). Reflexive Control. The relevance of a 50 year old Russian theory regarding perception control. *Militaire Spectator*, 187 (6).

Mackintosh, E. (2019). Finland is winning the war on fake news. What it's learned may be crucial to Western democracy. https://edition.cnn.com/interactive/2019/05/europe/finland-fake-news-intl/ Accessed on May 23, 2019.

NATO Strategic Communications (2016). Social Media as a tool of hybrid warfare. Nato Strategic Communications Centre of Excellence. Romerstrabe, Germany.

Nimmo, B. (2015). Anatomy of an Info-War: How Russia's Propaganda Machine Works, and How to Counter It. https://www.stopfake.org/en/anatomy-of-an-info-war-how-russia-s-propaganda-machine-works-and-how-to-counter-it/ Accessed on May 20, 2019.

Nissen, T.E. (2015). The weaponization of social media: Characteristics of contemporary conflicts. Royal Danish Defence College.

Pamment, J., Nothhaft, H., Agardh-Twetman & Fjallhed, A. (2018). Countering information influence activities: The state of the art. Department of Strategic Communication, Lund University.

Paul, C., Clarke, C.P., Schwille, M., Hlavaka, J.P., Brown, M.A., Davenport, S.S., Porche, I.R. & Harding, J. (2018). Lessons from others for future U.S. Army operations in and through the information environment: Case studies. Research ANd Development (RAND).

Permanent Select Committee on Intelligence (2018). Exposing Russia's effort to sow discord online: The internet research agency and advertisements. https://intelligence.house.gov/social-media-content. Accessed on January 23, 2019.

Rocha, R. & Yates, J. (2019). Twitter trolls stokes debates about immigrants and pipelines in Canada, data show. https://www.cbc.ca/news/canada/twitter-troll-pipeline-immigrant-russia-iran-1.5014750. Accessed on February 19, 2019.

Semple, J. (2018). Is Canada a target? A former Russian troll warns Canadians to be vigilant. https://globalnews.ca/news/4518563/canada-target-russian-troll-warns-vigilant. Accessed on February 19, 2019.

Swaine, J. (2018). Russian propagandists targeted African Americans to influence 2016 US election. *The Guardian.* https://www.theguardian.com/us-news/2018/dec/17/russian-

propagandists-targeted-african-americans-2016-election. Accessed on March 1, 2019.

Tunney, C. (2019). Ottawa setting up new team to warn Canadians of potential election interference. https://www.cbc.ca/news/politics/election-interference-panel-1.4998409. Accessed on March 3, 2019.

White, J. (2016). Dismiss, distort, distract, and dismay: Continuity and change in Russian disinformation, Issue 13. Institute for European Studies.

Ukrinform (2018). Fake news: Accepting the challenge. https://www.ukrinform.net/rubric-society/2593397-fake-news-accepting-the-challenge.html. Accessed on March 2, 2019.

Unkelbach, C. (2007). Reversing the truth effect: Learning the interpretation of processing fluency in judgment of truth. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 219-230.

Unkelbach, C. & Stahl, C. (2009). A multinomial modeling approach to dissociate different components of the truth effect. *Consciousness and Cognition, 18,* 22-38.