UNCLASSIFIED

Questions for the Record Senate Select Committee on Intelligence Foreign Influence Operations and Their Use of Social Media Platforms August 1, 2018

Questions for the Record for Renee DiResta

[From Senator Cotton]

As most people are aware, the most detailed accounting of Russia's past activities is the Mitrokhin Archive. On page 243 of the Mitrokhin Archive, as detailed in The Sword and the Shield, it states,

It was the extreme priority attached by the Centre (KGB Headquarters) to discrediting the policies of the Reagan administration which led Adropov to decree formally on April 12, 1982, as one of the last acts of his fifteen-year term as chairman of the KGB, that is was the duty of all foreign intelligence officers, whatever their "line" or department, to participate in active measures. Ensuring that Reagan did not serve a second term thus became Service A's most important objective.

On February 25, 1983, the Centre instructed its three American residences to being planning actives measures to ensure Reagan's defeat in the presidential election of November 1984. They were ordered to acquire contacts on the staffs of all possible presidential candidates and in both party headquarters...The Centre made clear that any candidate, of either party, would be preferable to Reagan.

Residences around the world were ordered to popularize the slogan "Reagan Means War!" The Centre announced five active measures "theses" to be used...his militarist adventurism; his personal responsibility for accelerating the arms race; his support for repressive regimes around the world; his responsibility for tension with his NATO allies. Active Measures "theses" in domestic policy included Reagan's alleged discrimination against ethnic minorities; corruption in his administration; and Reagan's subservience to the military-industrial complex."

1) So, in 1982, over thirty-five years ago, we had the KGB using active measures in the United Stat to sow racial discord, try to create problems with NATO, discredit our nuclear modernization, undercut military spending, highlight corruptions, and try to encourage the U.S. to retreat from the world stage. Aren't the themes the KGB used in 1982, similar to those we're seeing the Russian Intelligence Services use on social media in 2018?

Thematically there is some overlap between present day and past KGB active measures messaging; wars and corruption figured in prominently in the content on several Internet Research Agency (IRA)-linked sites. However, the themes that the IRA prioritized in 2018 were primarily internal societal struggles designed to create rifts between subsets of Americans. The tensions the IRA sought to exploit included racial discord (which appeared in numerous forms such as black and white militant and separatist content, Confederacy nostalgia, black culture content, police-violence related content), immigration status, cultural differences, religious

freedom, and hot-button political issues such as gun ownership rights and LGBT rights.

2) Isn't this Russian social media campaign really just old wine in new bottles, with perhaps a different distributor?

The difference is, indeed, the distributor -- and this is a critically important difference. Social networks afford an opportunity for speaking directly to people without the intervention of a gatekeeper; older active measures strategies often included the goal of laundering sympathetic content into a respected publication, but now the distrust in mainstream media affords subversive foreign propagandists the ability to simply market their content as "citizen journalism". In addition, the relatively low cost of online publishing, coupled with the availability of fraudulent social media accounts to share and otherwise elevate that content on highly-trafficked social media platforms, provides for nearly limitless experimentation. Adversaries can test market thousands of divisive narratives simultaneously using state-of-the-art tools for measuring already provided to marketers by the social media platforms.

There is an intersection of three factors at work: consolidation of hundreds of millions of users onto a handful of platforms, gameable algorithms, and the ability for precision targeting of content (designed to facilitate targeted ads in support of the advertising business model). The combination of these factors make it possible to distribute computational propaganda across a dense social ecosystem to those most likely to be receptive to it, and the content often receives an algorithmic assist. User-created content is for the most part treated equally; when it achieves a sufficient number of likes or shares, the platform algorithms may begin to promote it as "trending" or "recommended" content. Social platforms are built to drive user engagement; they are made to facilitate virality, and the ease of sharing ensures a velocity of transmission that makes stopping the spread of disinformation a significant challenge.

This new distribution model enables an unprecedented scale for influence operations and serves as an asymmetric advantage to any mildly sophisticated actor intent on pursuing these goals.

We've heard from open testimony before this Committee that the Russians are using active measures to undermine our missile defense deployments, nuclear modernization efforts, and to try and drive a wedge between the U.S. and NATO on these issues. Additionally, we know from Mitrokhin and Bob Gate's memoir "From the Shadows" that this was part of their playbook in the 1980s as well.

3) To what extent have you looked for and seen Russian activity on this front on social media?

With the caveat that attribution is complex, there are ongoing narratives that attempt to discredit NATO being spread and amplified in Kremlin-linked social media communities. This is

not a new phenomenon; there was press coverage in the New York Times (https://www.nytimes.com/2016/08/29/world/europe/russia-sweden-disinformation.html)

and The Guardian (https://www.theguardian.com/world/2017/jan/11/russia-waging-information-war-in-sweden-study-finds) about Swedish audiences being targeted with anti-NATO messaging in 2016 when Sweden was debating a military partnership with the alliance; the Swedish government identified Russia as the source of the false narratives.

There are also potentially abnormal patterns in the creation data of user accounts on Twitter that are focused on the topic. Among accounts currently discussing NATO on Twitter, we have observed an increase in the number of Twitter accounts brought online over time since the Swedish operation began. Further, the English-language accounts that we have observed sharing broader pro-Russia or Russian-origination messaging are also posting negative things about NATO and the US relationship with NATO allies. This is an ongoing investigation, but it appears that Russia and its surrogates are targeting Americans with messaging meant to call our role in NATO into question.

[From Senator Manchin]

4) What modifications would you recommend to the large social media companies that would enable users to identify the source and potential funding of items posted on social media?

There is an information war happening, and multiple types of actors are participating. This includes hostile state and non-state actors, but also includes coordinated attempts to spread disinformation by groups of real American ideologues. The platforms face a challenge in balancing cultural First Amendment expectations and allegations of censorship against the potential damage resulting from the unfettered spread of manipulative narratives designed to cause harm to individuals, society, and businesses alike. We believe that transparency about both the content itself (attribution) and the financial motivations behind it serve the interests of an informed citizenry. However, the ease of anonymous content creation on the internet - anyone can start a blog or make a meme - make identification of the source a significant challenge, and political dark money makes disclosing funding an unwinnable battle in the current legal environment.

What is possible, however, is for social media companies to take dubious distribution patterns into account when deciding what content their algorithms will recommend. Similarly, it is possible to assign quality indicators that factor in past behavior of the accounts sharing it (a ‘spamminess’, or quality quotient) and to the domain the content resides on. There is precedent for this in the effort to mitigate spam. Platforms should look to the history of anti-spam efforts for inspiration on managing computational disinformation as well.

Recognizing that this is an ongoing information war, we do anticipate an evolution in disinformation tactics from simple botnets to far more sophisticated narrative laundering

through authentic American accounts. Therefore, information sharing between third party researchers and technology platforms provides the best framework for ongoing protection of the information ecosystem.

5) Should there be disclaimers on anything other than personal information?

6) Should everything posted on social media have a "tag" that allows users to determine who posted information, even if it was re-posted or shared by another person, so you can always determine the actual source?

To answer both 5 and 6: content provenance is technologically extremely complex to implement at the scale described in the question, and also not terribly difficult for a determined adversary to evade. Visual image memes in particular often evolve slightly as they spread from user to user, so it's unclear what the attribution would or should link to. Repurposing and amplifying existing content is a tactic we have seen used by both the Internet Research Agency and the newly discovered Iranian social media manipulation operation; much of what they shared came from legitimate American news articles and meme pages, so even precise labeling of the content would not have made a significant impact in uncovering the operation.

[From Senator King]

7) At the hearing on August 1, 2018, I asked each witness to submit written policy recommendations to the Committee. Specifically, please provide recommendations on the following topics:

• Technical solutions, such as requirements to label bot activity or identify inauthentic accounts;

• Public initiatives focused on building media literacy;

• Solutions to increase deterrence against foreign manipulation; and

• Any additional policy recommendations.

The technologies underpinning the social media platforms have evolved in such a way that the interplay between three key phenomena -- mass consolidation of audiences onto a handful of platforms, gameable algorithms, and the ability to easily and precisely target people -- have created a problematic information ecosystem.

Legislating technological solutions for feature-level tactics leveraged by the IRA is fighting the last war. The specific features of social networks evolve rapidly; a Facebook ad today looks very little like an ad did a few years ago. Twitter has already diminished the ability for blocs of fully-automated accounts to easily game trending, so requiring that bots be labeled will have much less of an impact in 2018 than it could have had in 2016. The platforms already have policies in place for taking down inauthentic accounts; public pressure and pressure from financial stakeholders has begun to incentivize them to do so much more proactively. We

advocate avoiding the Maginot line of feature-focused legislation and instead prioritizing:

1) Implementation of cross-platform computational propaganda and algorithmic manipulation detection solutions to enable more rapid discovery of the signatures that indicate an emerging influence operation
2) Establishing oversight mechanisms empowered to keep the social media platforms acting in the interest of the public
3) Creating global economic and military deterrence strategies to raise the cost and risk of conducting influence operations for the malign actors involved.

Senator Warner introduced 20 policy proposals in a whitepaper immediately preceding the August 1 hearing that inspired this inquiry. In line with several of his proposals, we advocate for:

- The granting of rulemaking authority to the FTC as a significant step forward in the creation of a system of social platform oversight
- The establishment of an interagency task force, the creation of a formal deterrence strategy, and a re-evaluation of the Information Operations Doctrine
- The establishment of a public-private standing body to support threat information sharing between government, platforms, and researchers.

There is currently no disincentive to dissuade anyone, foreign or domestic, from undertaking a mass manipulation campaign ahead of an election. It is easy, it is inexpensive, and - judging by the fact that Russian and Iranian operations are still ongoing - past consequences have not yet created a perception that attacking the United States in this way will results in severe repercussions.

The United States presently faces extreme difficulties countering influence operations online because of laws such as US Law 50 U.S. Code § 3093(f), which prohibits the government from counter-messaging or engaging out of fear that such activity might violate the provision that prevents action “intended to influence United States political processes, public opinion, policies, or media.” Similarly, there is concern that gathering information, or collecting and analyzing the posts of suspect foreign social media accounts, could potentially violate the 1974 Privacy Act that governs the gathering of information about individuals if an American citizen’s information was also inadvertently gathered. These challenges were identified in 2015 while establishing the Global Engagement Center inside the State Department; engaging with presumed-foreign extremist accounts in anything other than an overt attributed capacity was deemed impossible because of the chance that an American digital bystander might see it, or that the pseudonymous extremist was perhaps themselves an American citizen. Therefore, at the moment, the overseas-partner model of the GEC provides the best option for countering foreign propaganda, and it should be fully staffed and funded.

Within the United States, the responsibility for coordinating investigations and responses to

influence operations is presently fragmented across the intelligence community. The CIA and NSA are constrained, leaving the FBI in charge of investigations. In contrast, several of our allies, including Germany and France, have dedicated cybersecurity organizations devoted to defense against these sophisticated attacks. These organizations are technically skilled agencies that are integrated and share intelligence with the rest of the country's national security entities; they have the technological expertise to engage with tech companies around threat information. The United States needs a similar whole-of-government approach to information operations, and must treat the threat as a cybersecurity issue.

Presently, oversight of threats to American democracy by way of private social platform infrastructure might fall under the purview of the Federal Trade Commission or the Federal Election Commission. The FTC has broad consumer-protection responsibilities but has neither deep expertise in internet manipulation, nor rulemaking authority. And since disinformation on the internet includes malign narratives outside of electoral or political concerns, FEC oversight would likely be insufficient. We need to more clearly assign responsibility and ensure that the agency chosen (or created) has the necessary tools to ensure that social networking companies take responsibility for addressing influence operations on their platform.

Domestic efforts must be complimented by an updated global IO doctrine and international detection and deterrence strategy, with the goal of mitigating foreign influence targeting our allies. We need a clear delegation of responsibility for this activity within the U.S. Government. Empowering law enforcement with updated legal tools to investigate and prosecute sophisticated foreign propaganda is essential; we should consider legislation that defines and criminalizes foreign propaganda that targets not just our political process but also addresses the targeting of commercial industry.

To address the final suggested policy area in the question: public initiatives to build media literacy are worth exploring in the interest of helping American citizens better understand how social media works, from the basics of the fact that there is an algorithmic ranking to the specifics of how misinformation and disinformation spread. We believe that any such program would have to apply not only to younger individuals currently enrolled in formal schooling, but to all Americans. One option for this might be a government-sponsored public service media literacy campaign, perhaps sponsored and disseminated on the social platforms in question.