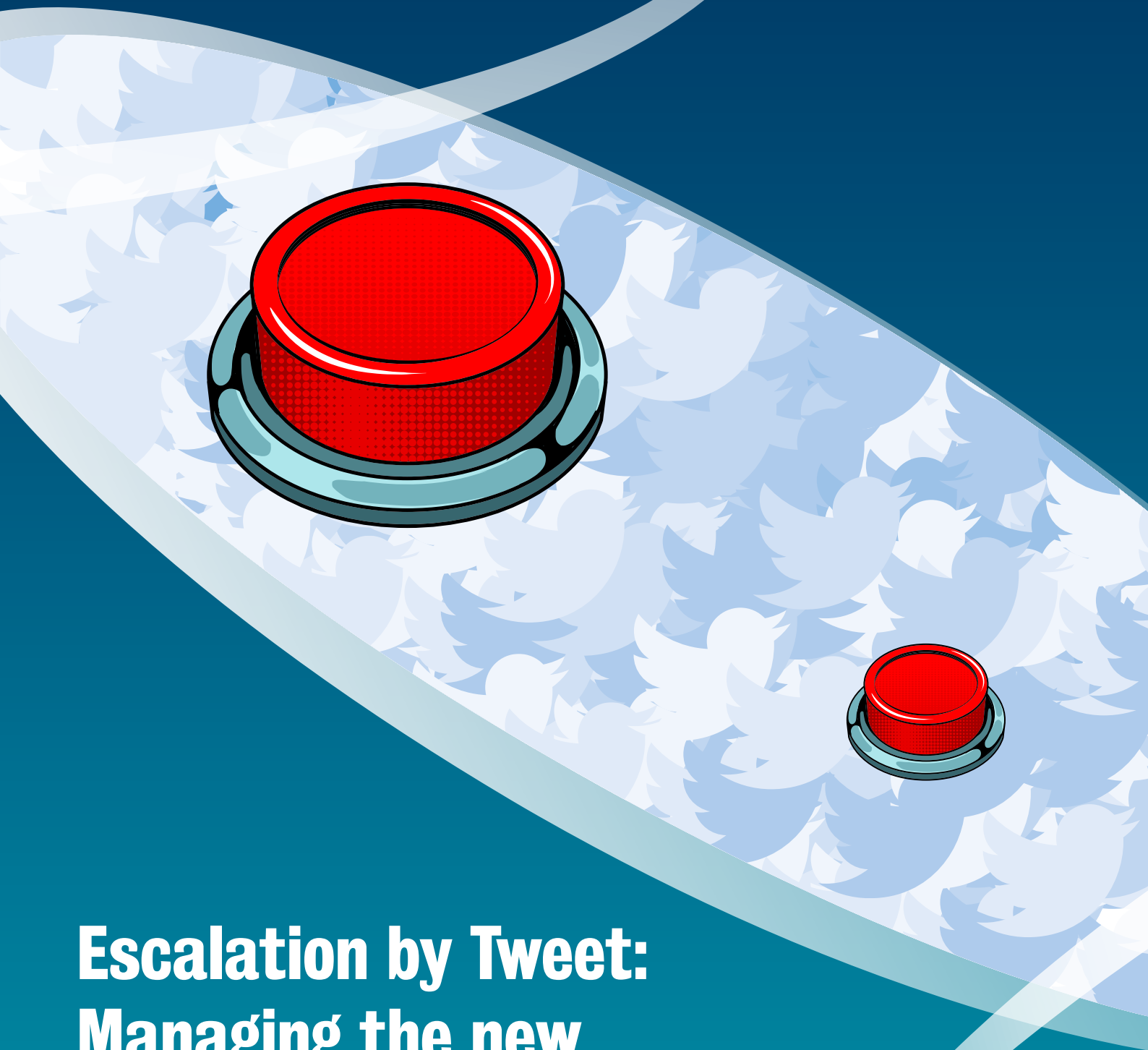
CENTRE FOR SCIENCE
& SECURITY STUDIES

KING'S College LONDON



Escalation by Tweet: Managing the new nuclear diplomacy

Heather Williams and Alexi Drew

JULY 2020

Table of contents

Foreword by Lawrence Freedman **4**

Executive summary **5**

Introduction **6**

Asymmetric tweeting **8**

Ladders, escalators and webs 10
What is new about Twitter? 10

The perfect tweet storm: escalation pathways **14**

Inadvertent escalation 14
Deliberate escalation and disinformation 16
Catalytic escalation 16

Social media and escalation management **17**

Pre-crisis recommendations 17
Crisis management recommendations 17

Conclusion **18**

About the Centre for Science and Security Studies at King's College London

The Centre for Science and Security Studies (CSSS) is a multi-disciplinary research and teaching group at King's College London that brings together scientific experts with specialists in politics, international relations, and history. CSSS forms part of the School of Security Studies at King's and draws on experts from the Department of War Studies and the Department of Defence Studies. Members of the Centre conduct scholarly and policy-relevant research on weapons proliferation, non-proliferation, verification and disarmament, nuclear security, space security and mass effect terrorism including the CBRN (chemical, biological, radiological and nuclear) dimension. In addition to academic staff, CSSS hosts masters and postgraduate research students, as well as visiting fellows and associates drawn from the academic, government, and business sectors. Our educational activities include contributions to the undergraduate and postgraduate offerings in the Department of War Studies, as well as professional development workshops for industry professionals.

Acknowledgements

We would like to thank all the participants from our project workshop in London in 2019 for their valuable contributions, particularly those who also provided feedback on the draft report. Interviewees from government, think tanks and industry were extremely forthcoming with their views on social media and escalation, which helped to inform our findings. We are grateful to Wyn Bowen, Ben Buchanan, Lawrence Freedman Leonie Haiden, Ben Loehrke and Andrew Reddie for their peer reviews and inputs throughout this project. Martin Everett and Felix Ruechardt provided valuable research assistance in the early stages of the project and during the workshop.

This project was made possible by support from the Defense Threat Reduction Agency's (DTRA) Project on Advanced Systems and Concepts for Countering Weapons of Mass Destruction (PASCC). PA# USAFA-DF-2020-255

Foreword by Professor Sir Lawrence Freedman

PROFESSOR SIR LAWRENCE FREEDMAN

Emeritus Professor of War Studies, King's College London

At times of crisis, whether caused by a political conflict, market failures, or a pandemic, there is an enormous appetite for news as people try to establish what is going on and how governments and international organisations intend to act. In the past people waited for radio broadcasts, watched a TV address or bought the latest edition of a newspaper. Now they stare at a screen waiting for snippets of information to appear in a Twitter feed or on a Facebook page. Because crises involve a number of countries, the same information will be sought by many audiences, but it will be interpreted according to their national perspectives. The audiences can include senior policy-makers from both adversaries and allies. For this reason, historically crisis communications have been constructed with care, seeking to make sure that the key messages are clear, showing resolve as necessary and describing ways to de-escalate the crisis.

The strong view coming out of this study is that clear key messages are still the best form of crisis communication, and that social media can help by facilitating their speedy dissemination. But when the tweet constitutes the whole message it may be so compressed that it is subject to misinterpretation and misrepresentation. Of course there are times when too much nuance may get in the way and so a punchy tweet that can convey a key message with urgency and clarity may be helpful. But such tweets need to be used sparingly for maximum effect, so that their special importance can be noted, and kept consistent with other high-level communications, including those sent privately through formal diplomatic channels.

A second important insight from this study is that Twitter is largely a conversation taking place in the English-speaking world, and in particular in the United States, which benefits from the opportunities to share information and engage in a vigorous exchange of opinions. There are also many active users in other countries. However, there are fewer in those countries most antagonistic to the US and its allies, where access to social media is often tightly controlled. This creates a significant asymmetry. Political leaders from these countries can intervene in the American conversation while keeping their own conversations relatively closed to outsiders. This also means the United States is more susceptible to attempts to spread disinformation deliberately, and to the potentially damaging consequences of rumours of unclear origin being taken seriously. Fortunately it seems to be the case that spreading disinformation is not simple, as it may need an established narrative to give it credibility. This takes us back to the need for crisis communications to be full and detailed so that the public can have confidence in what their leaders are saying and are not left trying to make sense of a cacophony of official pronouncements.

Escalation by tweet: Executive summary

DR HEATHER WILLIAMS

Lecturer in the Centre for Science and Security Studies and Defence Studies Department at King's College London. She previously worked at Chatham House and the Institute for Defense Analyses. She can be found on Twitter, @heatherwilly

DR ALEXI DREW

Postdoctoral researcher in the Centre for Science and Security Studies at King's College London. Her work focusses on the nexus between emerging technologies and international security. Her twitter handle is @CyberAlexi.

Social media has quickly become part of the geopolitical landscape, and international leaders and officials are increasingly taking to Twitter during crises. For US decision-makers, however, Twitter presents a bit of a paradox: on the one hand, tweets from government officials may help shape the American public narrative and provide greater insights into US decision-making to reduce misperception by foreign actors. On the other hand, tweets may increase misperception and sow confusion during crises, creating escalation incentives for an adversary. To reconcile this paradox, we examine the use of Twitter by international leaders during crises in recent years, some of which involved nuclear-armed states. In so doing, we explore the changing nature of escalation, which now resembles a complex web more than a ladder, and examine specific escalation pathways involving social media.

Based on this analysis, we find that social media has the potential to be a disruptive technology and exacerbate tensions during crises. To reduce the risk of tweets contributing to escalation in a crisis, we recommend the US Department of Defense:

- lead an interagency effort to develop best practices on the use of social media during crises;
- encourage leaders and officials to refrain from tweeting during crises and instead rely on more traditional means of communication, such as press releases and official statements;
- explore how to build public resilience to disinformation campaigns and provocations via social media during crises, as the American public is asymmetrically vulnerable to these attacks; and
- improve understanding of how various international actors use social media.

Twitter, as a company, and alliances such as NATO, also have a role to play in limiting the negative impact of Twitter during crises. If these findings could be summarised in 280 characters or less, it would be: 'To manage escalation during crises, stop tweeting.'

Introduction

Social media has quickly become part of the geopolitical landscape. On 3 January 2018, Trump famously referred to Kim Jong Un as 'little rocket man' and tweeted, 'North Korean Leader Kim Jong Un just stated that the "Nuclear Button is on his desk at all times." Will someone from his depleted and food starved regime please inform him that I too have a Nuclear Button, but it is a much bigger & more powerful one than his, and my Button works!'[1] And during the January 2020 US-Iran crisis, Iranian Supreme Leader Ali Khamenei posted an Instagram image of US President Donald Trump's face with a handprint, assumedly from being smacked.[2]

One of the primary concerns about social media is its potential impact on conflict escalation. With its 280-character limit, Twitter is not an ideal medium for crafting nuanced diplomatic messaging for purposes such as reassuring allies, deterring adversaries or signalling strategic intentions. Until recently, escalation was still typically conceived as a linear ladder with a clear progression across rungs.[3] But an increasingly complex environment defined by geopolitical and technological uncertainty[4] necessitates revisiting and updating this model and considering how social media could contribute to inadvertent, deliberate and catalytic escalation.[5] Are there generalisable patterns in the way states use social media as a means of public diplomacy? If so, under what conditions do these patterns risk crisis escalation?

Twitter during crises presents a bit of a paradox for US decision-makers. On the one hand, tweets from government officials might help shape the American public narrative and provide greater insights into US decision-making to reduce misperception by foreign actors. On the other hand, because of its lack of nuance, Twitter might equally increase misperception and sow confusion during crises. Based on our analysis of how various international actors use Twitter during crises, we find that social media has the potential to be a disruptive technology that exacerbates tensions during crises. Therefore, we recommend a pre-crisis and crisis management Twitter strategy to include:

- Develop interagency best practices for social media use during crises;
- During crises avoid tweeting by government officials and accounts, except to disseminate information from speeches or formal policy announcements. Tweeting should not be used as an independent signalling tool;
- Improve understanding of how disinformation is spread and consumed via Twitter and build public resilience to disinformation campaigns; and
- Use social media to gather information on potential adversaries, particularly how they use platforms such as Twitter.

There is a role for both governments and Twitter, as a company, in most of these recommendations. In short, to reduce the risks of unintended nuclear escalation, governments and individual officials should refrain from uncoordinated or 'rogue tweeting' during crises.

A goal of this study is to identify what is new about Twitter, compared to other means of diplomatic communication, during a crisis.[6] At the outset, we should highlight a challenge of researching social media: it is methodological Jell-O, spreading out in multiple directions and sliding across platforms, countries, actors and issues. To compensate for this, we focused on Twitter activity by government officials and agencies during a series of global crises over the study period, mid-2018 through early-2020.[7] As many of our findings apply to other social media platforms, the study highlights the need for more research and a developed conceptual framework,[8] particularly within the strategic studies community, on the impact of all social media on international security.[9,10]

Translation of overleaf:

Merely A Slap...

The matter of revenge and such things is another topic. They were slapped last night. What is important in the matter of confronting – military actions in this way do not make up for what they did – is that the corruptive presence of the U.S. in the region must end.[11]

Iman Sayyed Ali Khamenei , January 8th, 2020

فقط
یک
سیلی!
انتقام بحث دیگری است؛
حالا یک سیلی ای دیشب به
اینها زده شد. آنچه در مقام مقابله
مهم است - این کارهای نظامی به
این شکل، کفایت آن قضیه را نمیکند -
این است که بایستی حضور فساد برانگیز
آمریکا در این منطقه تمام بشود.
KHAMENEI.IR

Asymmetric tweeting

Different political actors use Twitter differently. States differ in the formality, regularity and nature of their messaging. Some states, such as Russia, use Twitter to replicate press releases or ministries' policy statements. There are important exceptions to this which we discuss below. Indian Prime Minister Narendra Modi tweets in the first person from a personal account.[12] Some countries go months without tweeting from some official accounts. The Indian Government, for example, expanded its use of Twitter to use by four official Ministries in 2016, but one of these, the Ministry of Railways, has not tweeted since 2018. Some actors use Twitter for aggressive messaging. Others are relatively benign, such as the Norwegian Ministry of Foreign Affairs which uses its English language account to disseminate information for tourists. One of the most important, yet underappreciated, aspects of Twitter diplomacy is that it is used asymmetrically by the United States compared to the rest of the world. Twitter Usage by Country (see Figure 1), shows that the United States has more registered Twitter users than any other country, with 59.35 million compared to 11.45 million users in India.[13]

Though Russia has 9.46 million Twitter users, this is only 6% of the population as Russians tend to prefer national platforms, such as Vkontake. Russian embassies do not have specific guidance from the Kremlin on what to post on social media and can follow the preferences of the ambassador. A very public example followed the 2018 novichok attack in Salisbury, United Kingdom, when the Russian embassy in London took to Twitter to deny accusations of Russian involvement and to accuse the United States of failing to destroy its own chemical weapons supply.[14] Given the relatively small number of Twitter users in Russia, these messages were clearly designed for foreign audiences (see Figure 2, overleaf).

Conversely, for the United States, Twitter is a means of both domestic and international communication. Nearly 20% of Americans use Twitter. In other countries with an active social media community, Twitter is not as useful for other domestic audiences – only 3% use it in Iran and Twitter is blocked in China (see Figure 3, overleaf).

As a primary user of Twitter, this makes the American public and decision-makers more susceptible to Twitter disinformation campaigns and messaging. This also means that Twitter could serve as a useful tool for American adversaries to shape international narratives or influence domestic audiences abroad.

Figure 1: Leading countries based on number of Twitter users as of January 2020 (in millions)[15]

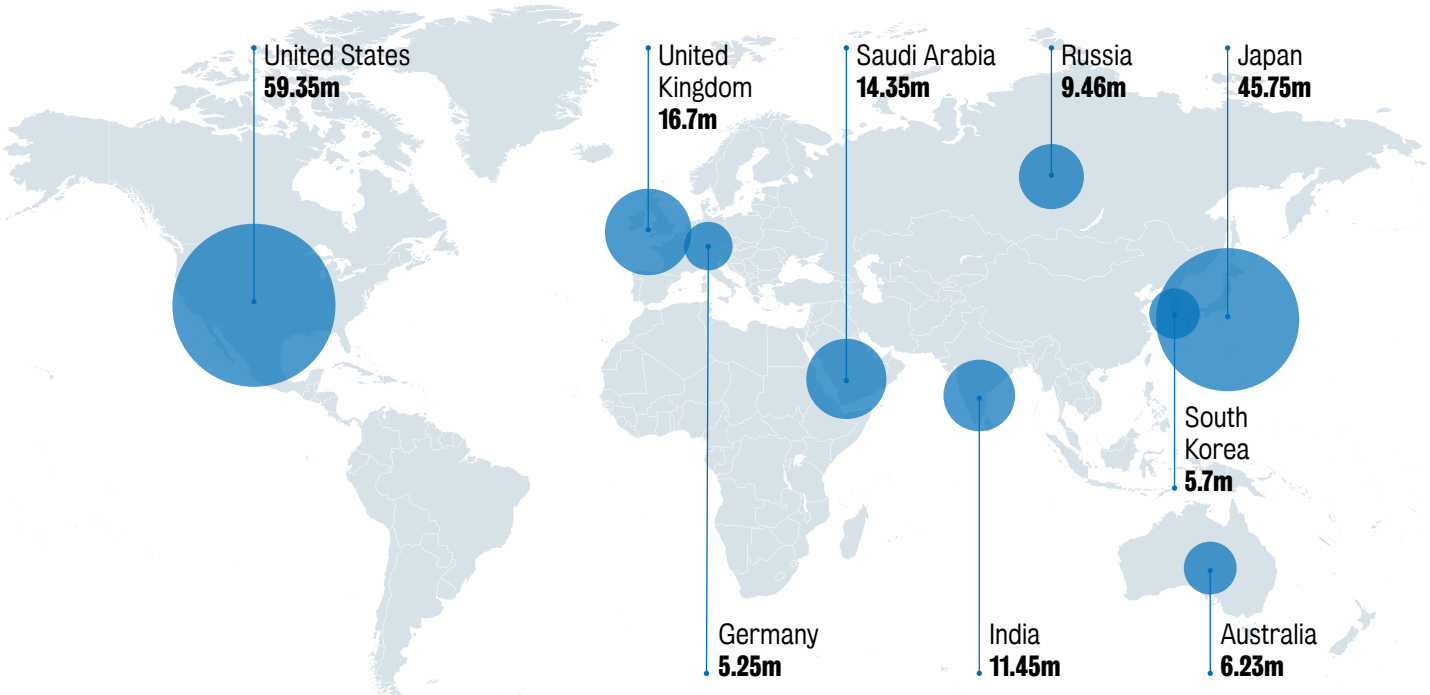
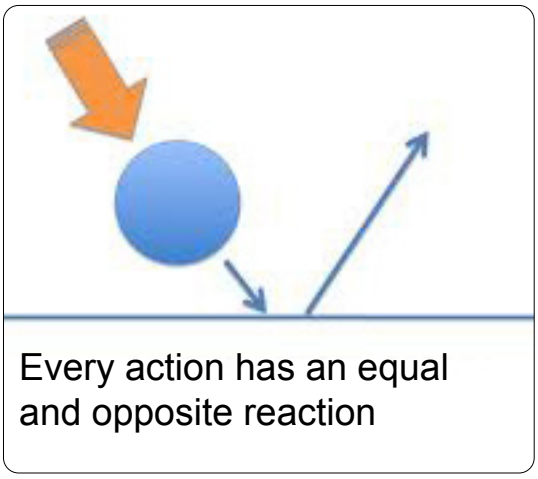


Figure 2: Salisbury exchange

13/03/2018

Russian Embassy, UK
@RussianEmbassy • 13/03/2018 5:25pm

Any threat to take "punitive" measures against Russia will meet with a response. The British side should be aware of that.



579 647 Likes

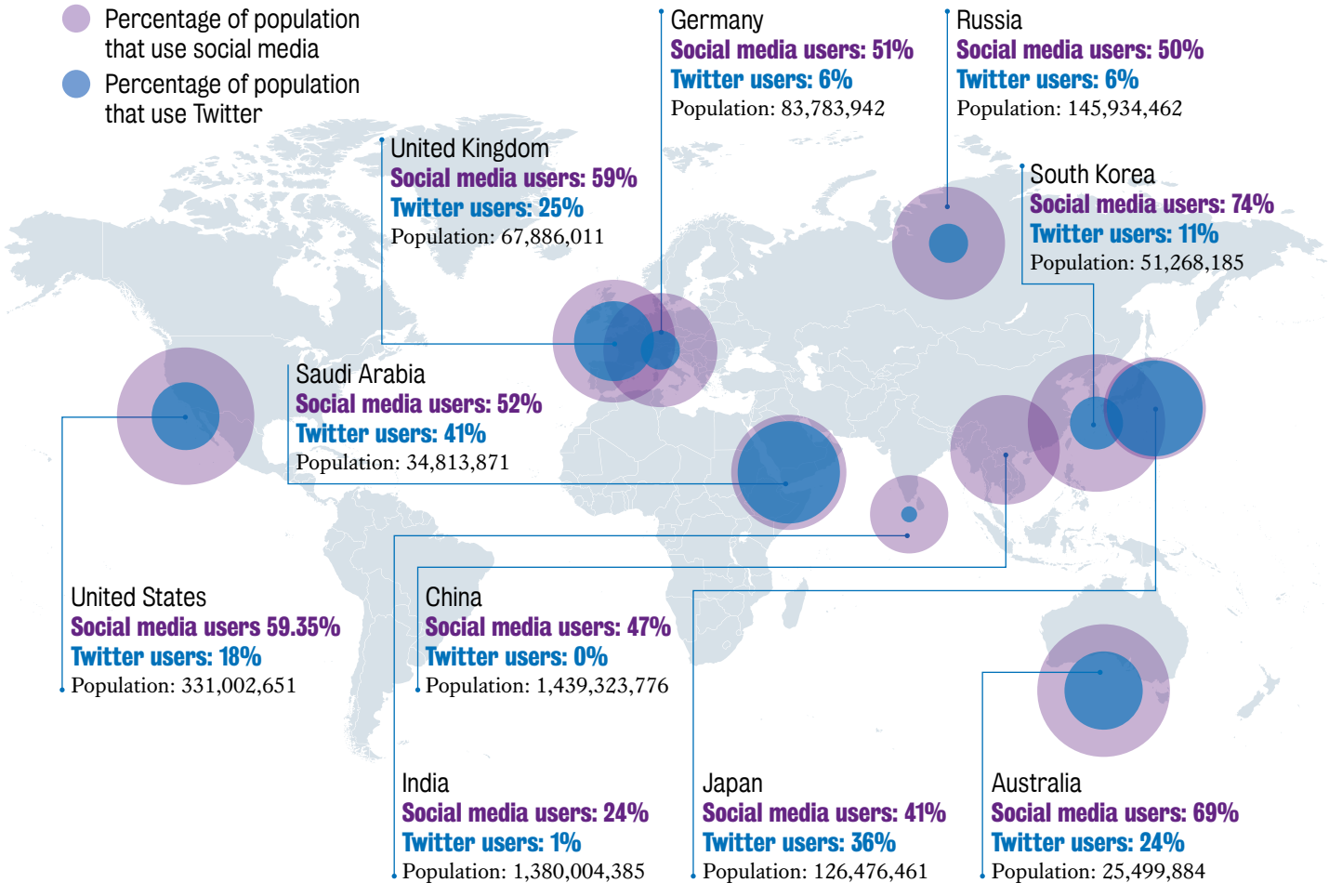
14/03/2018

Russian Embassy, UK
@RussianEmbassy • 14/03/2018 6:44pm

The temperature of 🇷🇺 🇬🇧 relations drops to — 2 3, but we are not afraid of cold weather.

2k 2.9k Likes

Figure 3: Active social network penetration in selected countries as of January 2020[15]



Ladders, escalators and webs

Crises are defined by political uncertainty. Any escalation, defined as 'the sequential expansion of the scope or intensity of conflict'[16] – depends on an intermingling of political factors, such as alliances, the stakes involved, reciprocity of escalation measures, domestic resilience and decision-making processes, along with diplomatic communication, which now includes social media.[17] The Cuban Missile Crisis, for example, did not escalate further because, among other reasons,the stakes weren't high enough for either side as argued by Lawrence Freedman in his 1991 study, 'Escalators and Quagmires.'[18]

Traditional metaphors of escalation have an element of automaticity. This is particularly evident in the classic text on escalation, Herman Kahn's *On escalation: scenarios and metaphors* (see Figure 4 for a condensed version of the original 44-step ladder).

Kahn's ladder is explicitly linear and sequential. But there are limits to this metaphor, particularly in the contemporary conflict and information environment. Whereas during the Cold War escalation was conceptualized as a ladder, we now see it as more of a web across domains with multiple actors and alliances.[19] Escalation risks are shaped by human factors and misperception, uncertainty and new technologies and mediums for communication. The escalation web consists of interconnected political factors and personalities communicating over multiple platforms and military capabilities. Twitter Timeline, the 2020 US-Iran Crisis (see Figure 5 on page 12) demonstrates the alignment of social media activity with crisis flashpoints, and differences in American and Iranian Twitter activity. We will return to this example later in the analysis.

What is new about Twitter?

Diplomatic and other forms of strategic communications have always been a component of escalation management. Three things make Twitter different, perhaps unique: speed, informality and openness. With no intermediaries, tweets can be sent instantaneously and Twitter as a platform rewards frequent and fast tweeting – the more you tweet, the more people see your account, and the more opportunities you have to gain followers. As a result of its speed, Twitter also allows for unprecedented informality: tweets from government officials and accounts do not necessarily have to go through interagency review in the way that other public statements or policies do.[20] And finally, whereas in the past government communications were either highly sanitized for public release or carefully

Figure 4: Kahn's escalation ladder

Civilian central wars

City targeting threshold

Military central wars

Central war threshold

Exemplary central attacks

Central sanctuary threshold

Bizarre crises

No nuclear use threshold

Intense crises

Nuclear war is unthinkable threshold

Traditional crises

Don't rock the boat threshold

Subcrisis Manoeuvring

DISAGREEMENT – COLD WAR

protected, tweets can be read by anyone with an account, and live forever.

All these factors have the potential to both escalate and de-escalate crises. An historical example helps demonstrate the impact of speed on crises. In the War of 1812, it took three weeks for a request for peace to travel by ship from London to the United States, by which time US President James Madison had declared war. In this case, the slow pace of information exchange was an important contribution to war. Even in the 21st century, Russian President Vladimir Putin could not reach President George W. Bush after the 9/11 attacks to understand why the United States had moved to DEFCON 3.[21] The speed of Twitter could potentially have been de-escalatory by providing immediate information and avoiding miscalculation. At the same time speed could have an escalatory effect during a crisis when tensions are high. We have all said things in the 'heat of the moment' which we wish we could take back.

The informal nature of Twitter allows for more personal engagement and frank dialogue, which research has shown can increase trust between international leaders.[22] A recent study, for example, argues that the personal relationship between US Secretary of State John Kerry and Iranian Foreign Minister Mohammed Javad Zarif played an important role in the Iran Nuclear Deal and was facilitated by 'both personal interaction and sustained Twitter communication during P5+1 nuclear negotiations between 2013 and 2015.'[23] This case study demonstrates the potentially positive benefits of Twitter. However it does not show that Twitter will always facilitate positive personal relationships between world leaders or that it will always contribute to rapprochement or de-escalation. Twitter is a hybrid of traditional means of diplomatic communication, such as retweeting official statements by the President, along with informal personal reflections. In the absence of a shared Twitter protocol, messages are often left open to interpretation and shaped by pre-existing views and interactions, not always with a positive effect.

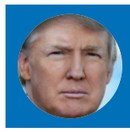
Finally, because tweets can be read by anyone, a tailored message will reach the intended audience where it might have a de-escalatory or calming effect, but also a wider audience that may interpret the tweet very differently. With regards to crisis escalation, there are at least three important Twitter audiences: the domestic public, international audiences and foreign governments. Communications for the purpose of conveying trustworthiness, soft power, and resolve will be interpreted differently across these three audiences.[24] Messages can also be shared, re-broadcast and reframed. This collateral messaging effect, discussed below in greater detail, increases risks of misperception. A tweet intended for a domestic audience can be interpreted as escalatory by an international actor. On the other hand, because tweets are available to everyone, they can provide open source intelligence on an actor's intentions and interests. Increasing available information may de-escalate tensions and build trust. In sum, by increasing available information, tweets are neither escalatory nor de-escalatory.[25] Their impact depends on the strategic interaction and geopolitical context as is the case in any crisis.

ASYMMETRIC TWEETING: STATE OF PLAY

Figure 5: Twitter timeline, the 2020 US-Iran Crisis

Mike Pompeo
@SecPompeo

Department of State
@StateDept



Donald Trump
@realDonaldTrump

Mark Esper
@EsperDoD

The White House
@WhiteHouse

Attack on US Embassy in Baghdad

Day 1: 31/12 | Day 2: 01/01 | Day 3: 02/01 | Day 4: 03/01 | Day 5: 04/01 | Day 6: 05/01 | Day 7: 06/01 | Day 8: 07/01 | Day 9: 08/01 | Day 10: 09/01

General Qasem Soleimani killed in drone strike near Baghdad airport

Iran launches rockets at US bases in Iraq

Iran Foreign Ministry
@IRIMFA_EN

Aytollah Khamenei
@khamenei_ir

Javad Zarif
@Jzarif

Hassan Rouhani
@HassanRouhani

Abbas Araghchi
@araghchi

Key
- Number of American tweets per day
- Number of American retweets per day
- Number of Iranian tweets per day
- Number of Iranian tweets per day

Tweets by actor per day range from 1–11
Reweets by actor per day range from 21–289,319

The perfect tweet storm: escalation pathways

Escalation can happen for at least three reasons: inadvertently or due to accident; intentionally or catalytically.[26] Twitter has the potential to impact all these pathways and we can envision specific escalation pathways involving social media.

Inadvertent escalation

Recent and ongoing studies highlight the potential for social media to exacerbate uncertainty and escalate tension.[27] With regards to nuclear weapons, most scholarship and policy attention focuses on the risk of inadvertent escalation which may or may not have a nuclear component, to include, 'mechanical failure, unauthorized (nuclear) use, or insanity'[28], or when, 'one party deliberately takes actions that it does not believe are escalatory but which are interpreted as escalatory by another party to the conflict.'[29] In short, inadvertent escalation is escalation that neither side necessarily wants but occurs because of mixed messages or changes in the information ecosystem impacting leaders' perceptions, as argued by Kristin ven Bruusgaard and Jackie Kerr.[30] Jeffrey Lewis's novel, *The 2020 Commission Report on the North Korean Attacks against the United States,* crafts a scenario in which a tweet from President Trump prompts Kim Jong Un to launch nuclear weapons. Although fictional, *The 2020 Commission Report* offers a useful thought experiment on the potential for social media to inadvertently escalate crises.

There are at least two ways in which Twitter could facilitate inadvertent escalation: collateral messaging and the fog of war. In collateral messaging, a tweet intended for a domestic audience, perhaps one meant to signal leadership or promote a 'strongman' image, is misinterpreted as aggressive by a foreign audience. This is particularly dangerous when it is at odds with other government messaging. The 2020 US-Iran crisis offered a real world example of social media messaging potentially escalating a crisis among nuclear actors. For the nine days of the crisis, from 31 December until and including 9 January 2020, three key US officials (Trump, US Secretary of State Mike Pompeo and US Secretary of Defense Mark Esper) sent 136 tweets in relation to the crisis, with half of them coming from President Trump. Conversely, three Iranian leaders (Supreme Leader Khamenei, Zarif, and political deputy at the Ministry of Foreign Affairs Abbas Aragchi) tweeted only 35 times about the situation.

During a crisis, tweets from the same government can often be at cross purposes if not carefully coordinated. At the height of the US-Iran crisis on 4 January 2020, President Trump stated in a series of tweets, 'Let this serve as a WARNING that if Iran strikes any Americans, or American assets, we have..... ...targeted 52 Iranian sites (representing the 52 American hostages taken by Iran many years ago), some at a very high level & important to Iran & the Iranian culture, and those targets, and Iran itself, WILL BE HIT VERY FAST AND VERY HARD. The USA wants no more threats!'[31] This was in contrast to numerous tweets by other US officials claiming not to want further escalation. This contrast could have created confusion and misperception among domestic, international and foreign government audiences.

These mixed messages can cause further confusion when they are retweeted and shape a wider narrative. The 2019 downing of a US Global Hawk drone demonstrated how Twitter can become a platform for diplomatic engagement, but also posturing for allies and domestic and international audiences (see Figure 6 on page 15).

President Trump's tweet indicating the United States was 'cocked and loaded to retaliate' led to a flurry of responses from Iranian officials. While the initial tweet may have been intended for domestic audiences to demonstrate America's military resolve, this was seemingly interpreted differently by various audiences.

The second way in which Twitter can lead to inadvertent escalation is through the fog of war, particularly information overload or constant background activity, a risk already identified with regards to cyber escalation.[32] While too little information can foster worst-case thinking, too much information can make it difficult to distinguish offensive from defensive operations.[33] Turning again to the Iran example, on 3-4 January 2020, Pompeo sent 23 tweets mentioning world leaders he had contacted emphasising that the United States wanted peace in the region. We can assume these tweets were intended to signal to Iran that the United States was committed to de-escalation, but they were sent concurrently with tweets by the President and White House that could be interpreted as escalatory, such as, '...We have the best military and the best intelligence anywhere in the world. If Americans are threatened, we are prepared to respond.'[34] From the Iranian perspective, which tweets should they listen to? An additional challenge in this case is that the United States and Iran do not maintain

Figure 6: Lifecycle of a tweet, 2019 drone shootdown

20/06/2019

Iran Military @Iran_Military • 20/06/2019 03:47
BREAKING: Iranian Forces shoot down intruding US Global Hawk HALE UAV over Hormuz Province in Persian Gulf. #Iran #PersianGulf #StraightOfHormuz #نارای
45 **4,404** Followers

Donald J. Trump @realDonaldTrump
20/06/2019 15:15
Iran made a very big mistake!
54,000 **76M** Followers

Iran Foreign Ministry @IRIMFA_EN
20/06/2019 17:06
Iranian Foreign Ministry Spokesman has strongly denounced a US spy drone"s intrusion into the Iranian airspace, warning the aggressors about the consequences of such acts while expressing our strong protest at such acts of aggression and provocative moves,we firmly warn against any aggression and illegal entry into the country"s airspace by any foreign flying object, and any aggression against the territory of the Iran & violation of its frontiers
31 **19,800** Followers

U.S. Central Command @CENTCOM
21/06/2019 00:43
The ISR Flight path and grid plots for the RQ-4A shot down by Iran in the Strait of Hormuz. "This was an unprovoked attack on a U.S. surveillance asset that had not violated Iranian airspace at any time..."
– Lt Gen Joseph Guastella,@USAFCENT
https://pbs.twimg.com/media/D9iuKXhXYAYBnNW?format=jpg&name=900x900
559 **400.5K** Followers

21/06/2019

Donald J. Trump @realDonaldTrump
21/06/2019 14:03
On Monday they shot down an unmanned drone flying in International Waters. We were cocked & loaded to retaliate last night on 3 different sights when I asked, how many will die. 150 people, sir, was the answer from a General. 10 minutes before the strike I stopped it, not proportionate to shooting down an unmanned drone. I am in no hurry, our Military is rebuilt, new, and ready to go, by far the best in the world. Sanctions are biting & more added last night.
32,600 **76M** Followers

23/06/2019

Javad Zarif @Jzarif • 23/06/2019 00:32
More evidence—including encroachment of a MQ9 spy drone on 5/26, speedboat purchases & phone calls planning to attribute ship attacks to Iran—indicate #B_Team was moments away from trapping @realDonaldTrump into a war. Prudence prevented it, but #EconomicTerrorism brings tension.
https://pbs.twimg.com/media/D9vjt9MWwAA_9iX?format=jpg&name=900x900
836 **1.5M** Followers

23/06/2019 ctd

John Bolton @AmbJohnBolton
23/06/2019 00:34
Great meetings today with Israeli PM @Netanyahu and NSA Meir Ben-Shabbat. We re-affirmed our shared priority of confronting Iranian aggression throughout the region by continuing maximum economic pressure and increasing the cost of Iran's malign activity.
671 **928,200** Followers

24/06/2019

Javad Zarif @Jzarif • 24/06/2019 19:23
.@realDonaldTrump is 100% right that the US military has no business in the Persian Gulf. Removal of its forces is fully in line with interests of US and the world. But it's now clear that the #B_Team is not concerned with US interests—they despise diplomacy, and thirst for war.
1,200 **1.5M** Followers

25/06/2019

Donald J. Trump @realDonaldTrump
25/06/2019 15:42
Iran's very ignorant and insulting statement, put out today, only shows that they do not understand reality. Any attack by Iran on anything American will be met with great and overwhelming force. In some areas, overwhelming will mean obliteration.
22,400 **76M** Followers

Iran Foreign Ministry @IRIMFA_EN
25/06/2019 19:14
Americans broke their promise and voided their own signature and stated that the world cannot trust them and they do not have any credit left from them. In this context, how do they expect to negotiate again with Iran. Imposing sanctions on the highest political, social, religious and spiritual leader of a country is a ridiculous act What the United States is doing today is acting against human rights. Today, there is a great confusion and frustration in the US administration.
https://pbs.twimg.com/media/D97S0DvWkAEJZo3?format=jpg&name=small
19,800 Followers

26/06/2019

Khamenei.ir @Khamenei_IR
26/06/2019 10:34
The graceful Iranian nation has been accused & insulted by world's most vicious regime, the U.S., which is a source of wars, conflicts & plunder. Iranian nation won't give up over such insults. Iranians have been wronged by oppressive sanctions but not weakened & remain powerful.
313 **741,600** Followers

embassies in each others' countries so lack a clear communication channel for requesting clarification.

Deliberate escalation, misinformation and disinformation

Why might a state deliberately escalate a crisis? Intentional escalation might occur in order to gain some advantage, to signal resolve if a state is heavily committed to the issue at stake or to avoid defeat.[35] State actors are most likely to use social media as part of an escalation strategy when they intentionally use disinformation to stir up domestic public unrest, shift international opinion or falsely signal escalatory measures. The NATO Strategic Communications Centre of Excellence defines disinformation as, 'false information spread deliberately.'[36] This is in contrast to misinformation, which includes 'false information spread by mistake' and 'junk information' or 'bullshit' (the technical term).[37]

The ability of disinformation campaigns to penetrate target audiences depends on the prevalence of social media platforms, but also on the government and public's ability to discern fact from fiction. Not all countries have independent media or other bodies to provide fact-checking. The use of social media to escalate a crisis can be further qualified as either chronic or acute disinformation campaigns. Chronic disinformation is more likely to have a catalytic effect, as we discuss below. Acute disinformation during a crisis, for a specific purpose, is more likely to have a deliberate escalatory effect.

Catalytic escalation

The scenarios above of unintentional or deliberate escalation involve state-based actors and a more traditional crisis scenario between two states. Catalytic escalation involves a third party who prompts one of the other actors to increase the intensity of a conflict. Given the open nature of Twitter, this may be particularly relevant to escalation by tweet. Other actors that could instigate catalytic escalation include individuals or agencies managing bots, the public and allies.

Bots, whether privately owned or state-sponsored, are increasingly present in the social media environment. Bots are autonomous programmes that might generate Twitter content without human intervention. A NATO study found that between 1 November 2019 and 31 January 2020, for example, English-language bot activity jumped from 12 to 15 per cent in one quarter.[38] This makes it difficult to disaggregate intentional from catalytic escalation because of the challenges of attributing bot behaviour and campaigns on social media to specific states. Examples include deep fakes such as a US Department of Defense press release circulated on Russian social media alleging a US bomber accidentally dropped a 'dummy nuclear bomb' on a Lithuanian building.[39] Many of these reports on Russian social media made their way onto more traditional news platforms, such as *Russia Today* and *Sputnik*, and were subsequently circulated worldwide via Twitter, as mapped in a recent study by Kate Starbird at the University of Washington.[40]

In addition to malicious individuals and/or corporations manipulating bots, there is a third actor here that is unavoidable with regards to Twitter: the American public. Specifically, Twitter is an ideal tool for manipulating American public opinion due to asymmetric use of the platform in the United States, whether by spreading inflammatory stories and/or disinformation.[41] Research into non-nuclear crisis scenarios demonstrates that whether or not the public believes a rumour on social media, 'is determined by pre-existing cognitive schemes and attitudes rather than simply by credulity or gullibility.'[42] For the most part, the public uses social media during a crisis to reduce uncertainty by verifying information and disproving false rumours,[43] foster a shared 'keynote' narrative of events and receive government guidance.[44] Whether or not this would also be the case in a geopolitical crisis and how governments use Twitter in a crisis, however, requires further research.

For the United States, tweets could also have a catalytic escalatory effect on allies. For example, if tweets indicated waning commitment to a common security agreement, allies could be prompted to build up their own defences and potentially create a new security dilemma. A particular concern is how adversaries might use social media to undermine alliance cohesion by targeting allies' domestic audiences. In the introduction to a recent research paper, the NATO Strategic Communications Centre of Excellence summarized its position with regards to the risks of social media and other digital platforms as, 'Malicious use of digital information poses a threat to armed forces by potentially compromising the confidentiality of information concerning geolocation, capabilities, tactics, and the future intent of friendly forces, or enabling and supporting an adversary's influence activities.'[45]

Social media and escalation management

Given the challenges of increasingly complex crises with multiple actors and information platforms, how can states manage escalation in an age of social media? Can tweeting strengthen deterrence messaging and de-escalate a crisis while also providing reassurance to domestic audiences? Or will tweets increase misperception and inadvertently escalate a crisis? US Strategic Command (STRATCOM) has already started grappling with these questions, following a major misstep: on 31 December 2018, STRATCOM tweeted that it was 'ready to drop something much, much bigger' than the ball in Times Square, presumably nuclear weapons. The tweet was heavily criticised both on social media and traditional media as 'inappropriate and unamusing' and subsequently deleted.[46] Since then, however, STRATCOM has developed detailed social media policies that highlight the risk of using unfiltered media, which, 'might be misinterpreted more frequently due to the inability to convey tone of voice and body language.'[47] STRATCOM sees its media strategy as contributing to the deterrence message by 'proactively conveying facts about command activities.'[48] In a recent example, STRATCOM tweeted about Global Lightning[49], an annual command and control battle staff exercise to assess joint operational readiness; however, its 'Myth Monday' series has drawn recent criticism.[50]

NATO does not have an alliance social media policy. Such policies are left to individual member states. Nonetheless, NATO has been a leader in highlighting how social media can contribute to shared understandings of risks, can target audiences with fact-driven messages to combat disinformation campaigns and can serve as an education platform. NATO-affiliated groups are working to counteract damaging narratives that might encourage escalation in domestic audiences among member states.[51] The impact of diplomatic communications during crises are often largely informed by pre-crisis perceptions and attitudes as discussed above; therefore, using Twitter as a tool for escalation management requires not only considering the impact of tweets during a crisis but also before it begins.

Pre-crisis recommendations

Pre-crisis management requires a three-pronged strategy. First, the US Government should develop interagency best practices for the use of social media during crises. This effort could be led by the Department of Defense and State Department, building on STRATCOM's existing policies. This might include a list of possible scenarios that would prompt a no-tweeting policy by key government officials.

Secondly, given America's asymmetric use of Twitter and the public's vulnerability to disinformation campaigns via Twitter, the Department of Defense should form an Interagency Working Group to consider the effects of disinformation campaigns by foreign actors and how to build public resilience. This working group could include input from the Defense Science Board or the Defense Policy Board. Meanwhile private industry, other experts, and Twitter as a company should not wait for the government to act on this recommendation. Twitter should be a key contributor to efforts to identify disinformation and how it is spread. Multidisciplinary approaches by academics and non-government experts can help identify how to build public resilience to disinformation drawing on the fields of psychology and crisis management. A team at the University of Washington, for example, has produced original data on how disinformation is spread, which offers an important starting point for such efforts.[52]

Thirdly, as part of intelligence collection and analysis, the United States should understand the different social media platforms used across the world; but more importantly, it should ask how countries will use social media to send signals and diplomatic communications during a crisis. This can prevent mirror-imaging and avoid assumptions that all countries use social media in the same way as the United States.

Crisis management recommendations

For governments, the best crisis management strategy is to refrain from tweeting from personal accounts and to instead rely on officially coordinated messaging. In the case studies we examined, official accounts for various branches of the US military typically demonstrated this policy already exists and is in practice by limiting tweets to concise and factual information. Following the shoot down of an Iranian drone in 2019, for example, Central Command tweeted: 'CENTCOM confirms that a U.S. Navy drone was shot down by an Iranian surface-to-air missile system while operating in international airspace over the Strait of Hormuz at approximately 11:35 p.m. GMT on June 19, 2019.'[53]

The case studies also point to areas of confusion, however, that could have an amplifying affect during crises. For example, the flurry of tweets between 3 and 4 January 2020 from numerous US government-affiliated accounts, including Secretary of State Pompeo and President Trump, had potential to increase the fog of war and lead to misunderstanding about US intentions towards Iran following the assassination of General Qasem Soleimani. Therefore, the interagency should have a pre-existing Twitter protocol for various crisis levels.

Responsibility also lies with Twitter to some extent to prevent tweets from exacerbating a crisis. We recommend that Twitter enforce policies designed to limit platform manipulation and disinformation; actively remove content and accounts engaged in influence operations targeting ongoing crises; continue to draft and enforce neutral policies aimed at discouraging platform manipulation and influence operations during crisis; and increase transparency regarding historic instances of platform manipulation that have led to or might lead to conflict.

Conclusion

The nature of escalation is inherently uncertain, and Twitter is but one of many factors that will make escalation management an increasingly complex task. Based on the above theoretical treatment across different types of escalation, we conclude tweets are unlikely to independently start a crisis, move up the 'escalation ladder' or expand the escalation web. There is a risk, however, that tweets can enable or accelerate an ongoing crisis, and that American audiences will be disproportionately at risk to manipulation because of their asymmetric use of Twitter.

At the outset we introduced the paradox Twitter presents to US decision-makers: on the one hand, the platform provides a fast and open way to communicate with the American public and potentially to signal resolve to adversaries during a crisis. On the other hand, these same traits mean tweets are often misinterpreted and can feed into pre-existing biases or fears. Given that Twitter is less popular in all other countries, US decision-makers could understandably assume that their tweets will not be read by foreign audiences and therefore tailor tweets during crises to domestic audiences. As we have demonstrated, however, while foreign *publics* may not be reading Twitter, foreign *governments* are increasingly active on Twitter and use it not only to gain insights into US decision-making but also to send their own signals and shape international narratives about the United States. Resolving this paradox, therefore requires extreme caution.

Much of the work for mitigating the escalatory effects of a tweet is required *before* a crisis begins, such as developing interagency best practices and a plan to coordinate tweets during a crisis. Governments are not the only ones who can play a role in Twitter escalation management. Twitter as a company can help identify disinformation campaigns and risks, particularly during crises. And finally, and most importantly, the best way to ensure tweets do not escalate a crisis is for US officials to refrain from tweeting at times of heightened tension. Unfiltered and uncoordinated messages can increase misperception through collateral messaging or thicken the fog of war. More traditional media, such as press releases and official statements, avoid many of the risks associated with tweeting during crises. These coordinated statements can be retweeted and circulated via social media, but ultimately 280-characters rarely allows enough space for the necessary care and nuance required during a crisis.

At the time of the Cuban Missile Crisis, Nikita Khrushchev warned John F. Kennedy in a letter that they were at risk of becoming blind moles clashing if they did not show wisdom.[54] Social media ostensibly increases uncertainties and reduces opportunities for such wisdom associated with escalation: decision-makers are no longer simply 'blind moles', but blind moles with smartphones.

References

1. Donald Trump Twitter Account, 3 January 2018. Available at: https://twitter.com/realdonaldtrump/status/948355557022420992?s=21.
2. This message came following the assassination of General Qasem Soleimani by the United States, followed by an Iranian response. The events surrounding the crisis are discussed below in greater detail. Ayatollah Ali Khamenei Twitter Account, 9 January 2019. Available at: https://twitter.com/Khamenei_fa/status/1215301620990468096?s=20.
3. Herman Kahn, *On Escalation Metaphors and Scenarios* (Oxford: Routledge, 2009), p.39
4. US Department of Defense, *Nuclear Posture Review February 2018* (Office of the Secretary of Defense: Washington DC, 2018).
5. Herbert Lin, 'Escalation Dynamics and Conflict Termination in Cyberspace,' *Strategic Studies Quarterly 6* no. 3 (2012), p.53
6. Data presented later in this paper demonstrates the breadth of social media platforms and the diversity of their uptake across national boundaries. Twitter was chosen for the basis of this study due to the ability to collect historical and current data, and because it is a platform which enjoys variable but consistent usage across the widest number of states.
7. We did not include social media activity during the COVID-19 pandemic.
8. For a particularly significant contribution to the field, see: Harold Trinkunas, Herbert Lin, and Benjamin Loehrke, *Three Tweets to Midnight: Effects of the Global Information Ecosystem on the Risk of Nuclear Conflict* (Stanford: Hoover Institution Press, 2020).
9. Kate Starbird, 'Information Operations and Online Activism within NATO Discourse', in Harold Trinkunas, Herbert Lin, and Benjamin Loehrke, *Three Tweets to Midnight: Effects of the Global Information Ecosystem on the Risk of Nuclear Conflict* (Stanford: Hoover Institution Press, 2020).
10. An additional justification for focusing on government officials and accounts is that these are more likely to be retweeted and play an asymmetrically influential role in the information space. See, for example, Starbird.
11. For English translation of Instagram image, see, Ayatollah Khamenei, Twitter Account, 9 January 2020. Available at: https://twitter.com/Khamenei_fa/status/1215301620990468096?s=20
12. Narendra Modi, Twitter Account: https://twitter.com/narendramodi
13. With a population of 1.34 billion people, this means that less than 1% of the population in India uses Twitter, compared to 18% of the United States and 6.5% of Russia.
14. Russian Embassy UK Twitter Account, 26 January 2018. Available at: https://twitter.com/russianembassy/status/956892230266118145
15. *Figure 1 source:* https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries. Published by J. Clement, 14 February 2020. Statista.com *Figure 3 source:* Active social network penetration in selected countries as of January 2020. https://www.statista.com/statistics/282846/regular-social-networking-usage-penetration-worldwide-by-country/ J. Clement, 14 February 2020. Statista.com. Social media users as a percentage of the total population Australia 2015-2018. https://www.statista.com/statistics/680201/australia-social-media-penetration/ Published by C. Hughes, 9 October, 2019, Statista.com. Social network user penetration in Saudi Arabia from 2017 to 2023. https://www.statista.com/statistics/553656/mobile-messaging-user-reach-saudi-arabia/. Published by A. Puri-Mirza, 20 February, 2019. Statista.com
16. Robert Osgood and Robert Tucker, *Force, order, and justice* (Baltimore: John Hopkins University Press, 1967), p.17; cited in Daniel S. Geller, *Nuclear Weapons, Deterrence, and Crisis Escalation,* Journal of Conflict Resolution 34 no.2 (1990): 292
17. Lawrence Freedman, 'Escalators and quagmires: expectations and the use of force', *International Affairs 67,* no.1 (1991)
18. Ibid., p.23.
19. See, for example, Vincent Manzo, 'Deterrence and Escalation in Cross-Domain Operations: Where Do Space and Cyberspace Fit?' *Strategic Forum,* no. 272 (December 2011). Rodney Jones, 'Nuclear Escalation Ladders in South Asia', Defence Threat Reduction Agency / Policy Architects International, Report number ASCO 2011 007, April 2011.
20. This is largely based on anecdotal evidence from government officials; however, many U.S. government offices seem to harvest their tweets from language in interagency approved documents.
21. William M. Arkin and Robert Windrem, 'Secrets of 9/11: New Details of Chaos, Nukes Emerge', *NBC News,* 11 September 2016. Available at: https://www.nbcnews.com/storyline/9-11-anniversary/secrets-9-11-new-details-chaos-nukes-emerge-n645711
22. Nicholas Wheeler, *Trusting Enemies: Interpersonal Relationships in International Conflict* (Oxford: Oxford University Press, 2018)
23. Constance Duncombe, 'Twitter and transformative diplomacy: social media and Iran-US relations', *International Affairs,* 93:3 (2017), p.546.
24. We are extremely grateful to Leonie Haiden at the King's College London Centre for Strategic Communications for this point and helping to tease out the different audiences.
25. On the role of information and trust-building, see, for example, Andrew Kydd, *Trust and Mistrust in International Relations* (Princeton: Princeton University Press, 2007).
26. See, for example, Lin.
27. 'Three Tweets to Midnight: Nuclear Crisis Stability and the Information Ecosystem', Policy Memo, Stanley Foundation, 6 November 2017.
28. Barry R. Posen, 'Inadvertent Nuclear War?: Escalation and NATO's Northern Flank', *International Security,* 7:2 (Fall, 1982), p.29.
29. Lin, p.52
30. Kristin Ven Bruusgaard and Jaclyn A. Kerr, 'Crisis Stability and the Impact of the Information Ecosystem', in Harold Trinkunas, Herbert Lin, and Benjamin Loehrke, *Three Tweets to Midnight: Effects of the Global Information Ecosystem on the Risk of Nuclear Conflict* (Stanford: Hoover Institution Press, 2020), pp.137-158.
31. Donald Trump Twitter Account, 4 January 2020. Available at: https://twitter.com/realDonaldTrump/status/1213593965838163968?s=20
32. Lin, p.57
33. Barry R. Posen, 'Inadvertent Nuclear War?: Escalation and NATO's Northern Flank', *International Security,* 7:2 (Fall, 1982), p.34.
34. The White House Twitter Account, January 2020. Available at: https://twitter.com/WhiteHouse/status/1213287607091781632?s=20
35. Lin, p.52
36. Jente Althuis and Leonie Haiden, (eds), Fake News: A Roadmap, (Riga: The NATO StratCom Centre of Excellence, 2018).
37. See, for example, UK Government Communication Service, *RESIST: Counter-disinformation toolkit, 2019.*
38. NATO Strategic Communications Centre of Excellence, 'Robotrolling', 2020.
39. DFRLab, '#FakeNews: American bomber did not drop a bomb on a house in Lithuania', 16 June 2017, Available at: https://medium.com/dfrlab/fakenews-american-bomber-did-not-drop-a-bomb-on-a-house-in-lithuania-6ae64241fe9e
40. Starbird.
41. G. A. Fine, 'Rumor matters: An introductory essay', in G. A. Fine and V. C. Heath (Eds.), *Rumor mills: The social impact of rumor and legend* (New Brunswick, NJ: Transaction, 2005).
42. K. Hazel Kwon, C. Chris Bang, Michael Egnoto and H. Raghav Rao, "Social media rumors as improvised public opinion: semantic network analyses of twitter discourses during Korean saber rattling 2013," *Asian Journal of Communication* (2019), p.5.
43. Tanja Schreiner, 'Information, Opinion, or Rumour? The role of Twitter During the Post-Electoral Crisis in Cote d'Ivoire', *Social Media + Society* (2018).
44. Panos Panagiotopoulos, Julie Barnett, Alinaghi Ziaee Bigdeli, and Steven Sams, 'Social Media in Emergency Management: Twitter as a Tool for Communicating Risks to the Public', *Technological Forecasting and Social Change* (July 2016).
45. Sebastian Bay et al., *Camouflage for the Digital Domain,* (Riga: NATO Strategic Communications Centre of Excellence, 2020), p.4
46. See, for example, Matt Stevens and Thomas Gibbons-Neff, 'Military Deletes New Year's Eve Tweet Saying It's 'Ready to Drop Something'', *New York Times,* 31 December, 2018.
47. Office of Public Affairs, United States Strategic Command: Public Affairs Communications Strategy (2018): Annex D, p.2
48 Ibid., page 3
49. US STRATCOM Twitter Account, January 23, 2020. Available at: https://twitter.com/US_Stratcom/status/1220351599811399680
50. See, for example, Matt Korda, Twitter Account, 13 July 2020, Available at: https://twitter.com/mattkorda/status/1282705725073104900
51. James Pamment et al., The Role of Communicators in Countering the Malicious Use of Social Media (Riga: NATO Strategic Communications Centre of Excellence, 2018).
52. See, for example, Starbird.
53. US Central Command, Twitter Account, 20 June 2019. Available at: https://twitter.com/CENTCOM/status/1141703097397534720?s=20
54. As cited in Freedman, Robert Kennedy, *Thirteen days: the Cuban missile crisis of October 1962* (London: Macmillan, 1969), p.80.

Centre for Science and Security Studies
Department of War Studies
King's College London
Strand
London WC2R 2LS
United Kingdom

www.kcl.ac.uk/csss
@KCL_CSSS

© 2020 King's College London

Designed by: RF Design (July 2020)
Approved by: brand@kcl.ac.uk