Report of the Select Committee on
Deliberate Online Falsehoods - Causes,
Consequences and Countermeasures
Parl. 15 of 2018
Presented to Parliament pursuant to
Standing Order 105.
Ordered by Parliament to lie upon the Table:
19 September 2018

THIRTEENTH PARLIAMENT OF SINGAPORE

Second Session

REPORT OF THE SELECT COMMITTEE ON DELIBERATE ONLINE FALSEHOODS – CAUSES, CONSEQUENCES AND COUNTERMEASURES

Parl 15 of 2018

**Presented to Parliament on
19 September 2018**

PART A

MAIN REPORT

COMPOSITION OF THE SELECT COMMITTEE
(*as at 11 September 2018*)

The Select Committee was appointed by resolution of Parliament passed on 10 January 2018 and comprised the following Members:

Mr Deputy Speaker (Mr Charles Chong) (Punggol East) (*Chairman*)

Ms Chia Yong Yong (Nominated Member)

Dr Janil Puthucheary (Pasir Ris-Punggol), Senior Minister of State, Ministry of Transport and Ministry of Communications and Information

Mr Desmond Lee (Jurong), Minister for Social and Family Development and Second Minister for National Development and Deputy Leader of the House

Mr Pritam Singh (Aljunied)

Ms Rahayu Mahzam (Jurong)

Mr Seah Kian Peng (Marine Parade)

Mr K Shanmugam (Nee Soon), Minister for Home Affairs and Minister for Law

Ms Sun Xueling (Pasir Ris-Punggol), Senior Parliamentary Secretary, Ministry of Home Affairs and Ministry of National Development

Mr Edwin Tong Chun Fai (Marine Parade), Senior Minister of State, Ministry of Law and Ministry of Health

---

MASTER CONTENTS OF PARTS A, B AND C

REPORT OF THE SELECT COMMITTEE

CONTENTS OF PART A

REPORT OF THE SELECT COMMITTEE ON DELIBERATE ONLINE FALSEHOODS – CAUSES, CONSEQUENCES AND COUNTERMEASURES

The Select Committee on Deliberate Online Falsehoods – Causes, Consequences and Countermeasures ("Committee"), constituted pursuant to resolution of Parliament, has agreed to the following Report: -

Introduction

1. On 5 January 2018, the Ministry of Law announced that it would ask Parliament to appoint a Select Committee to study the problem of deliberate online falsehoods, and to recommend how Singapore should respond.

2. The reasons for appointing the Committee were set out in a Green Paper submitted to Parliament by the Ministry of Communications and Information and the Ministry of Law titled "Deliberate Online Falsehoods: Challenges and Implications." The Green Paper outlined the real and serious challenges posed by deliberate online falsehoods and called for a wide-ranging conversation on what Singapore's response should be as a country and as a society.

3. On 10 January 2018, Parliament resolved –

"That this House appoints a Select Committee to examine and report on:

(a) the phenomenon of using digital technology to deliberately spread falsehoods online;

(b) the motivations and reasons for the spreading of such falsehoods, and the types of individuals and entities, both local and foreign, which engage in such activity;

(c) the consequences that the spread of online falsehoods can have on Singapore society, including to our institutions and democratic processes; and

(d) how Singapore can prevent and combat online falsehoods, including:
    (i) the principles that should guide Singapore's response; and
    (ii) any specific measures, including legislation, that should be taken."

4. Parliament also resolved that the Committee would comprise Deputy Speaker Charles Chong as Chairman; and seven Members of Parliament from the Government benches, one Member of Parliament from the Opposition Benches and one Nominated Member of Parliament, to be nominated by the Committee of Selection. The default position would be for the Committee to have members from the Government and Opposition benches. The Government proposed including a

Nominated Member in place of one Member from the Government benches, to have more diversity.

5. The resolution to appoint the Committee and on the composition of the Committee was unanimously adopted by Parliament.

(I) MEMBERSHIP AND MEETINGS OF THE SELECT COMMITTEE

6. The Committee comprised Deputy Speaker Charles Chong as Chairman and the following Members who were nominated to the Committee by the Committee of Selection:

(i) Ms Chia Yong Yong
(ii) Dr Janil Puthucheary
(iii) Mr Desmond Lee
(iv) Mr Pritam Singh
(v) Ms Rahayu Mahzam
(vi) Mr Seah Kian Peng
(vii) Mr K Shanmugam
(viii) Ms Sun Xueling
(ix) Mr Edwin Tong Chun Fai

7. The Committee held 16 meetings, the minutes of which are at Appendix I.

8. The Committee deliberated extensively and went through numerous suggestions and formulations before agreeing on the final version of the report. Arising from this, all decisions made by the Committee were unanimous and consensual. It reflects the Members' shared understanding of the problem and of what Singapore needs to do to counter it.

(II) INVITATION TO THE PUBLIC TO SUBMIT WRITTEN REPRESENTATIONS

9. On 16 January 2018, the Committee issued a press release inviting the general public to submit written representations on any matter falling within the Committee's Terms of Reference. Written representations could be submitted in English, Chinese, Malay, or Tamil.

10. The Committee encouraged a wide range of views from the public. Members reached out to and engaged experts and other stakeholders who could add useful perspectives to the Committee's work.

11. The closing date for submissions was originally 28 February 2018. In response to requests, the Committee extended the closing date by one week, to 7 March 2018. A press release on the extension was issued on 27 February 2018.

(III) WRITTEN REPRESENTATIONS RECEIVED

12. In total, the Committee received 170 written representations. These included six late written representations that the Committee decided to accept. On 9 April 2018, the Committee published 167 of the 170 written representations.

13. Appendix III of this Report reproduces these 167 written representations.

14. The Committee decided not to publish the written representations by Dr Damien Cheong and Dr Gulizar Haciyakupoglu, as they addressed matters with foreign sensitivities. In place of their written representations in Appendix III are summaries of their oral evidence, which was heard in private sessions.

15. The Committee decided not to publish the written representation by Mr Alex Tan, being of the opinion that it was not made in good faith. It contained personal insults, irrelevant comments and sarcastic proposals.

(IV) PUBLIC HEARING

16. After due consideration of the written representations, the Committee heard oral evidence from 65 individuals and organisations. The oral evidence was heard in public, over eight days, on 14-16, 22-23 and 27-29 March 2018. The hearing lasted approximately 50 hours. Written representations of the oral representors and video recordings of their sessions were made publicly available on Parliament's website on the same day.

17. The verbatim Minutes of Evidence are set out in Appendix IV.

18. The Committee should also refer to two parties who were invited but eventually did not come to give evidence. The first is Human Rights Watch ("HRW"). HRW initially accepted the invitation. However, two working days after requesting to be heard on a specific date, they informed the Committee that their representative had made travel plans that "could not be changed". They were then offered a number of dates and also told that if these dates were not suitable, they could avail themselves of the alternative option of video-conferencing at any time between 15 and 29 March 2018. HRW did not take up the offer. HRW's attendance would have enabled an examination of HRW's views in respect of the measures to combat deliberate online falsehoods. It was clear to the Committee that HRW's excuses for non-attendance were contrived.

19. The second is Reporters Without Borders (also known as Reporters Sans Frontieres ("RSF")), who was invited to give oral evidence, including on its publications on Singapore. RSF initially expressed interest in giving oral evidence, and proposed to attend the hearings. However, it eventually declined to attend, citing "organisational reasons". They were then offered a number of dates

and also told that if these dates were not suitable, they could avail themselves of the alternative option of video-conferencing at any time between 20 and 29 March 2018. RSF did not take up the offer. It was clear to the Committee that RSF's excuses for non-attendance were contrived.

20. The facts of what transpired in relation to HRW and RSF are set out in the Addendum.

(V) OUTCOMES OF THE PROCESS

21. The Committee received representations and heard from a broad cross-section of society. This included:

    a. local and foreign academics in relevant fields across various educational and research institutions;
    b. experts who shared other countries' experiences for our reference;
    c. technology and media giants;
    d. operators of social media platforms;
    e. religious leaders and community groups;
    f. civil society activists; and
    g. students and other members of the public.

22. The youngest representor was a 15-year-old while the oldest was aged 80. The representors included 26 school students.

23. Several of the written representations were substantive, providing an in-depth and yet varied appreciation of the issue.

Findings and Views of the Committee

24. The Green Paper sets out some of the experiences of other countries, and helped frame the issues on which further evidence should be gathered. Its purpose was to serve as a starting point for the Committee's work. The evidence since considered by the Committee has addressed in greater depth and detail the issues raised in the Green Paper and in the Committee's Terms of Reference.

(I) UNDERSTANDING THE PHENOMENON

(A) The Nature and Use of Deliberate Online Falsehoods

25. The Committee received substantial and in-depth evidence on the phenomenon of deliberate online falsehoods. The evidence received by the Committee showed that deliberate online falsehoods are often created or spread by different actors, through various types of digital technologies; they can impact society severely in many ways; and are very difficult to combat. To better understand the nature and use of deliberate online falsehoods, Part I(A) of this Report will address the following four issues:

a. the actors behind online falsehoods and the objectives that the falsehoods are designed to achieve;
b. the use of digital technologies to spread online falsehoods;
c. the types of impact that deliberate online falsehoods have had, namely, on national security, public institutions, individuals, and businesses; and
d. the difficulties in combatting online falsehoods, in light of how people are influenced by falsehoods, and social changes caused by the digital revolution.

26. Several representors made statements in respect of foreign countries in the context of disinformation campaigns or information operations. As stated by the Committee during the hearing, the Committee is not in a position to draw any conclusions in favour of or against any of the other countries mentioned. Statements set out below concerning any country should be regarded as statements made by representors. These statements, as they relate to the actions of the other countries, do not reflect the Committee's views.

(1) Actors Who Use Falsehoods and Their Objectives

27. The Committee heard evidence concerning the actors who are behind the spread of deliberate online falsehoods. These actors may be foreign or local, States or civilians. They may be motivated by politics, prejudice, or ideology. Individuals, both local and foreign, may also be motivated by profit, mischief or social connection.

28. The falsehoods spread by these actors are designed to achieve a myriad of objectives. These objectives may be part of broader ideological and political agendas. The objectives of the different actors may sometimes align, despite differences in their underlying motivations. When they do so, the threat they pose is greater. The evidence received by the Committee on actors who use falsehoods and their objectives is set out more comprehensively in **Annex A**.

a. Foreign State actors

29. The objectives of falsehoods spread by foreign State actors include advancing or undermining a particular policy, discrediting public institutions, influencing election outcomes, sowing discord among communities and groups, and fracturing society's shared sense of reality. These objectives ultimately work to further broader geopolitical interests. Several examples were given and the evidence suggested a broad targeted attack on several institutions and countries. A few examples are set out below.

30. *Advance or undermine a domestic policy.* One representor observed generally how the use of falsehoods by foreign State actors can make it almost impossible for European governments to develop constructive policies to deal with issues such as migration. One example of such a falsehood was the "Lisa" case,[1] which involved false allegations that the German police were covering up the rape of a girl in Germany by a group of refugees.[2] The girl had claimed to be kidnapped and assaulted by men of Middle-Eastern descent. The German police found that her claims had been fabricated. However, media from a foreign country continued to publicise the girl's claims without reference to the findings of the German police. This triggered anti-immigrant demonstrations and a campaign to "expose" the German government's attempts to cover up crimes perpetrated by refugees and immigrants.

31. In the Czech Republic, false narratives were spread online about how the United States (US) was responsible for the influx of Syrian refugees into Europe, and that the US and the North Atlantic Treaty Organisation (NATO) were ultimately responsible for the conflict in Ukraine. The objective of such narratives was reportedly to increase domestic support for the Czech Republic to leave the European Union (EU) and NATO.[3]

32. *Advance or undermine a foreign policy.* Foreign State actors have allegedly used online falsehoods to influence the foreign policy of European countries towards Ukraine. In Sweden, a forged official letter (purportedly from Sweden's Ministry of Justice) circulated online suggested that Ukraine had sought to improperly

[1] Jakub Janda, "Full-Scale Democratic Response to Hostile Disinformation Operations", *European Values Think-Tank* (20 June 2016), p 1.
[2] Ben Nimmo, Written evidence submitted to UK Digital Culture, Media and Sport Committee "Fake News" inquiry (19 April 2017), para 17.
[3] Jakub Janda and Ondfej Kundra, "Mechanisms of Influence of the Russian Federation into Internal Affairs of the Czech Republic", *European Values Think-Tank* (4 September 2016), p 1.

influence a war crimes case before the Swedish courts. This, together with a series of other forged letters, was allegedly part of an attempt by a foreign country to undermine the support among the Swedish public for Ukraine.[4]

33. *Discredit public institutions and leaders.* Several examples were given of online falsehoods designed to discredit public institutions and leaders. In the "Lisa" case described above at [30], the foreign media continued publishing the girl's false allegations of rape without reference to the findings of the German police, and even alleged that the German police were part of a cover up. Similar one-sided reporting of claims seeking to discredit governments were also found to be a problem in the United Kingdom (UK).[5]

34. In Ukraine, a foreign country reportedly spread the false narrative that the Ukrainian government was fascist and corrupt, and that foreign military intervention to save Ukrainians was necessary. This false narrative was built up using falsehoods, such as a false online video interview claiming that Ukrainian soldiers had crucified a child. The objective was to sow distrust in the Ukrainian government, and galvanise support for foreign military intervention in parts of Ukraine. Surveys and other evidence cited by Mr Jakub Janda from the Czech Republic (Head, Kremlin Watch Program; and Director, European Values Think-Tank in Prague, Czech Republic) indicated that opinions of the Ukrainian government were indeed negatively influenced by such falsehoods, as elaborated on at **Annex E**.

35. *Achieve an election outcome.* During the 2016 US Presidential Election, a foreign disinformation campaign allegedly sought to denigrate one political candidate in favour of another, and influence the outcome of an election.

36. *Sow discord.* In the US, online falsehoods reportedly spread by foreign disinformation agents sought to polarise political discourse and stir up tensions in society, in order to advance the broader geopolitical aim of diminishing the US' international influence.[6] They did so by targeting already divisive issues, such as race, LGBT rights, gun control, and immigration. They did not promote a particular policy position, but instead played on all sides of the political spectrum, turning groups against each other.

37. One example given was of a video which inaccurately claimed to show an African-American woman being shot by a policeman in Atlanta, Georgia. The video was

[4] "Fake letter 'likely to be part of wider campaign'", *Radio Sweden* (14 September 2015); N MacFarquhar, "A Powerful Russian Weapon: The Spread of False Stories", *New York Times* (28 August 2016).
[5] Ben Nimmo, Written evidence submitted to UK Digital Culture, Media and Sport Committee "Fake News" inquiry (19 April 2017), para 9. For example, a news outlet of a foreign country was found by the UK's communications regulatory authority to have violated the obligation to preserve due impartiality by publishing one-sided interviews of people who made grave accusations against certain governments, without providing adequate comment from those governments.
[6] "Assessing Russian Activities and Intentions in Recent US elections", *Intelligence Community Assessment* (6 January 2017), p. ii.

spread by a group of accounts with the hashtag #shockingmurderinatlanta.[7] According to Mr Ben Nimmo from the United Kingdom (Senior Fellow, Information Defense Digital Forensic Research Lab), the video was fake and spread by a foreign troll factory, and its purpose was to widen the divide between the African-American community and the police, as well as to undermine the police as an institution.

38. Foreign troll accounts which attempted to influence the 2016 US Presidential Election also sought to sow discord by "using vaccination as a political wedge issue",[8] according to a recent study published in the American Journal of Public Health. The study found that while there was general consensus regarding the efficacy of vaccines in the American population, the discussions on Twitter gave a different impression and suggested that there was a lot of debate about the issue.[9] Tweets containing false information about vaccines were posted by what were most probably inauthentic accounts.[10] Examples included: "Did you know #vaccines caused autism?" and "#vaccines contain mercury! Deadly poison!".[11] From the pro-vaccine camp, tweets such as "You can't fix stupidity. Let them die from measles, and I'm for #vaccination" and "#vaccines are a parent's choice. Choice of a color of a little coffin" were posted.[12] The researchers who conducted the study commented that these foreign troll accounts used polarising language and linked vaccination to controversial statements about race, class and government legitimacy[13] in a bid to sow discord.

39. *Fracture society's shared reality.* According to disinformation experts, the consistent stream of foreign disinformation in Ukraine had the objective of sowing doubt and confusion over the truth, so as to undermine reality-based politics, meaningful civic discourse, and consequently, democratic stability.[14] The broader strategic aim was to weaken the country's resistance to foreign influence and aggression.

[7] "Cop shooting, Ebola scare in Atlanta invented by Russians: Report", *AJC* (3 June 2015); Adrian Chen, "The Agency", *New York Times* (2 June 2015); Andrew Prokop, "The new Mueller indictments tell us a lot about Russian trolls", *Vox* (16 February 2018).
[8] David Broniatowski et al, "Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate", *American Journal of Public Health* (23 August 2018), p 2.
[9] David Broniatowski et al, "Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate", *American Journal of Public Health* (23 August 2018), p 6.
[10] David Broniatowski et al, "Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate", *American Journal of Public Health* (23 August 2018), p 5.
[11] David Broniatowski et al, "Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate", *American Journal of Public Health* (23 August 2018), p 6.
[12] David Broniatowski et al, "Weaponized Health Communication: Twitter Bots and Russian Trolls Amplify the Vaccine Debate", *American Journal of Public Health* (23 August 2018), p 6.
[13] "Russia trolls 'spreading vaccination misinformation' to create discord", *BBC* (24 August 2018).
[14] Monika Richter, "The Kremlin's Platform for 'Useful Idiots' in the West: An Overview of RT's Editorial Strategy and Evidence of Impact", *European Values Think-Tank* (18 September 2017); Peter Pomerantsev, "Russia and the Menace of Unreality", *The Atlantic* (9 September 2014).

b. Local actors

40. Local actors may spread falsehoods for political and ideological objectives, such as to achieve specific election outcomes, attack politicians, turn groups against one another, and promote or oppose policies. They may also do so to gain financial benefit, or create mischief, among other objectives.

41. *Achieve an election outcome.* During the 2016 US Presidential Election campaign, the domestic alt-right reportedly drove a number of major false narratives to harm the Clinton campaign and boost that of Trump.

42. In Indonesia, domestic, politically-motivated actors have used online hoaxes that play on ethnic and religious sentiments to undermine election candidates. Representors familiar with Indonesia observed how such hoaxes have been rife in Indonesian elections since 2012, and have increased in intensity over the years, a trend which is expected to continue.

43. *Attack politicians.* Independent of elections, online hoaxes driven by domestic groups in Indonesia have reportedly sought to undermine Indonesian President Jokowi. One common tactic has been to falsely claim that President Jokowi has communist affiliations, thereby tapping on entrenched anti-communist sentiments.

44. *Turn one group against another.* Anti-Muslim falsehoods have been spread in the US and the UK by domestic far-right groups. For example, in the aftermath of the terrorist attack in Paris in 2017, a far-right political leader in the UK posted a video on Twitter, and described it as showing Muslims celebrating the attack.[15] It was in fact a video of people celebrating a cricket match victory in Pakistan. The video gained nearly 500,000 views in a matter of hours.

45. In France, anti-immigration falsehoods have identified migrants as a threat to the French way of life. One example given was a false report by one of the most influential French far-right opinion websites that the Breton lighthouse in Paris would be demolished to provide housing for migrants.

46. Indonesian authorities have uncovered an extensive and politically well-connected network known as the Muslim Cyber Army. This local network has spread falsehoods and hate speech online to inflame sentiments against gay men and lesbians, alleged communists, Chinese, and the government, and promoted a hard-line Islamist stance. It was reportedly coordinated through a central WhatsApp group, and used bot armies to amplify falsehoods.

47. *Promote or oppose policies or ideological beliefs*. Online falsehoods promoting or opposing policies or ideologies may be spread to manipulate those who do not

[15] Matt Novak, "This Video of 'Muslims Celebrating the Paris Terror Attack' Is Totally Fake", *Gizmodo* (21 April 2017).

hold the same political beliefs, and to get more people to concur with one's own ideological beliefs. Examples include the following:[16]

a. Representors from Germany and the Czech Republic testified that online falsehoods by domestic politically-motivated groups had made it difficult for their countries to make constructive policies on migration.

b. In the US, after the February 2018 school shooting in Parkland, Florida, gun rights advocates spread false stories that survivors interviewed by the media were in fact actors, and that the shooting had never happened.[17] Their objective was to shore up support for gun rights.

48. *Financial gain.* Digital advertising models have allowed website owners to earn advertising revenue based on the level of user engagement with the advertisements placed on or linked to their websites. Several representors described how digital advertising models incentivised online content producers to compromise the truth in order to attract "clicks" and generate advertising revenue. This was said to be because "the economics of social media favour gossip, novelty, speed and 'shareability' and not truth", and that "an altered reality" tends to make stories more interesting.

49. Representors identified a range of actors who were to various degrees incentivised by digital advertising revenue, from those motivated purely by financial gain, whom Mr Nimmo termed "fake news merchants", to citizen journalists, and the media industry.

50. While financially-motivated actors may have no political agenda, they may have political impact. For example, American Paul Horner, who claimed that he hated Trump, wrote false stories attacking Clinton and promoting Trump during the 2016 US Presidential Election, raking in an alleged US$10,000 a month as a result.[18] In an interview, he expressed regret that, with hindsight, he may have helped rather than hurt Trump's campaign.

51. Companies may spread falsehoods to shore up the survival of their businesses. An example given by a group of students from the Singapore Management University

[16] In the UK, the 2012 Leveson Inquiry into the culture, practices, and ethics of the press found numerous examples of misleading news articles, including online articles, that prioritised political agendas over accuracy. For example, several news outlets reported that new criminal sentencing guidelines in the UK would allow drug suppliers to avoid custodial sentences. In fact, the new sentencing guidelines made no change at all to the sentencing approach for drug suppliers. Fact-checking organisation, Full Fact, noted that such misreporting was part of a trend to portray the judiciary as lax on crime. It advanced an established agenda of resisting any perceived "softening" on criminal sentencing. See: "An Inquiry into the Culture, Practices and Ethics of the Press – Report", *The Leveson Inquiry* (November 2012), Volume II, para 9.48.

[17] Issie Lapowsky, "Parkland Conspiracies Overwhelm the Internet's Broken Trending Tools", *WIRED* (21 February 2018).

[18] Jeremy Stahl, "Purveyor of fake news says he targeted Trump supporters, influenced election", *Slate* (17 November 2016); Sally French, "This person makes $10,000 a month writing fake news", *Marketwatch.com* (18 November 2016).

(SMU) Law School was of large companies who funded research of dubious accuracy to make consumers believe their products are not harmful. Tobacco companies are known to have financed research challenging whether smoking causes lung cancer; similarly, fossil fuel manufacturers have reportedly sought to attribute climate change to natural causes. Representatives from TrendMicro submitted that it was not uncommon for companies to seek to undermine their competition using hoaxes and smear campaigns.

52. *Mischief*. Besides political agendas, conspiracy theories may find continued life because of the desire to create mischief. This was demonstrated by an example given by Mr Nimmo of a forged letter purporting to expose the spying by Britain on then-candidate Mr Donald Trump at the request of then-President Barack Obama. This fed conspiracy theorists who suggested that Mr Trump was the victim of an international "deep state" conspiracy aimed at undermining his presidency. Despite being repeatedly and easily exposed as a fake by netizens, some nevertheless suggested sending the letter to news broadcasters "for the lulz" (*i.e.* for entertainment). In the same vein, 15-year-old student Mr Zubin Jain shared how his own motivation for having posted falsehoods in the past was to alleviate boredom, and that it was not unusual for his peers to spread online falsehoods for the attention or profit.

53. Falsehoods may be created for the sheer thrill of being able to influence people. One representor shared how her review of online conversations on a spam website revealed that the creators of spam sometimes sought to "show off their ingenuity".[19]

c. Foreign non-State actors

54. Online falsehoods from overseas may emanate from private persons too. Foreign private individuals and organisations may spread online falsehoods targeting a particular country, or to achieve political, ideological or financial objectives that surpass national boundaries.

55. *Achieve an election outcome.* During the 2017 German Federal Election, the alt-right from the US was said to be involved in disinformation campaigns on Twitter that supported the election agenda of German alt-right politicians.[20] Elections in several Latin American countries were the subject of online influence campaigns run by Colombian Andres Sepulveda. Although he was paid to do so, Sepulveda has said that he was primarily motivated by right-wing ideology, and sought to remove dictatorial and socialist governments.[21]

[19] Yvonne Wong, Appendix III: Written Representations, Paper No. 11, page B23.
[20] Simon Hegelich, "Who is trolling the German election? Russia, AltRight or both?", *Political Data Science* (14 September 2017), available at <http://politicaldatascience.blogspot.sg/2017/09/who-is-trolling-german-election-russia.html>.
[21] Jordan Robertson et al, "How to hack an election", *Bloomberg Businessweek* (31 March 2016).

56. *Promote or oppose policies.* An Asian country reportedly has an online "army" of content creators, whose role is to promote the government's policies and attack criticisms of those policies, both within and outside that country. This "army" is said to comprise individual netizens and non-governmental institutions, most of whom are volunteers. Similarly, in another foreign country, individuals are said to carry out troll activities in other countries, not necessarily because they were paid to do so, but because of a strong ideological impetus.

57. *De-legitimise a government.* Foreign non-State actors such as NGOs and media organisations may use falsehoods aimed at de-legitimising a government. According to Mr Ruslan Deynychenko from Ukraine (Co-founder, StopFake.org), news organisations from a foreign country had spread falsehoods about how the Ukrainian government had persecuted its own citizens, and sought to demonise and de-legitimise the Ukrainian Government. This included false reports of how Ukrainian citizens were being tortured, raped and murdered.

58. *Turn one group against another.* Racist and other such prejudiced agendas are often not limited by national borders. Falsehoods supporting such agendas may be published online for worldwide consumption. For example, an anti-Muslim falsehood posted on US website InfoWars was found by the UK authorities to be among the extremist material read by UK citizen Darren Osborne before he committed a violent anti-Muslim act.[22]

59. *Radicalise.* Terrorist organisations, such as ISIL, have used online disinformation to radicalise people around the world. In 2017, ISIL released a video featuring a radicalised Singaporean fighter, who called on viewers to join ISIL's fight.[23] The Islamic Religious Council of Singapore, MUIS, subsequently released a media statement noting that the video was "full of distortions and falsehood" and "deliberately designed to mislead Muslim viewers into sympathising with ISIS".[24]

60. *Financial gain.* The "fake news" industry in Macedonia was responsible for a proportion of the fictional and hyper-partisan stories that proliferated during the 2016 US Presidential Election. Here at home, a false story claiming that Singapore's Minister of Foreign Affairs had collapsed at an international event was published by an overseas website. While the exact motivation for the publication of this falsehood was not reported, one representor surmised that it was very likely to be to generate digital advertising revenue, as the website in question had a practice of creating such falsehoods in the past.

[22] Kevin Rawlinson, "Finsbury Park-accused trawled far-right groups online, court told", *The Guardian* (23 January 2018).
[23] "MUIS condemns ISIS video featuring Singaporean", *The Straits Times* (28 September 2017).
[24] "Media Statement – MUIS Statement on ISIS Video", *MUIS* (27 September 2017), para 2.

d. Alignment of different actors

61. The objectives of these different types of actors may overlap. Mr Janda highlighted how the interests of local actors can align with the geopolitical interests of foreign State actors and impact on the State. According to Mr Janda, an example of this in the Czech Republic was where there was alignment of a foreign country's geopolitical interest, with local Czech actors who supported the foreign country's geopolitical interest and local actors who published disinformation simply for economic gain.

62. Such alignment tends to cause a falsehood to appear more credible and be amplified further. The alignment may be deliberate, or unwitting.

63. A July 2018 report published by the University of Oxford Computational Propaganda Research Project ("CPRP 2018 Report") analysed the trends and strategies of organised media manipulation by State actors. It found that there was evidence in several countries around the world that State actors have formally coordinated with other actors in society. These other actors included private industry, civil society organisations, Internet subcultures, youth groups, hacker collectives, fringe movements, social media influencers and volunteers who ideologically support the cause.[25]

64. The alleged foreign disinformation campaign during the 2016 US Presidential Election was said to have capitalised on falsehoods created by the unwitting US-based alt-right, and foreign and local profit-driven "click-bait" writers. An example of this was given by Dr Claire Wardle, currently based in the US (Executive Director, First Draft; Research Fellow, Shorenstein Center for Media, Politics and Public Policy, Harvard Kennedy School), using a false article titled "Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement."[26] According to research cited by Dr Wardle, the article was created by an unidentified person, and published on a website known as WTOE5News in July 2016. WTOE5News was later found by journalists to be part of a network of 43 fake news sites, which earned digital advertising income by generating readership. The article was shared on Facebook by someone working for this fake news network. It was then re-shared by different groups of people, namely, (i) those who sought to amplify the reach of the article to make profit, (ii) Trump supporters, (iii) other forces who had an interest in Trump winning, *e.g.* trolls from a foreign country, and (iv) Clinton supporters, to show how easily Trump supporters could be fooled.

[25] Samantha Bradshaw and Philip Howard, "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation", *University of Oxford, Computational Propaganda Research Project* (2018), pp 9-10.

[26] See Claire Wardle and Hossein Derakshan, "Information Disorder: Toward an interdisciplinary framework for research and policy making", *Council of Europe report* (27 September 2017), pp 23-25; and Craig Silverman and Jeremy Singer-Vine, "The True Story Behind the Biggest Fake News Hit of the Election", *BuzzFeed* (17 December 2016).

(2) Use of Digital Technologies to Spread Online Falsehoods

65. The deliberate spread of falsehoods is not new. However, considerable evidence was given showing how modern digital technology has made the creation and dissemination of falsehoods easier, cheaper and more profitable, transforming it into what experts regard as a new global phenomenon.[27] Its ease, speed, scale and impact are unprecedented. Tools and services are easily and cheaply available in the market. Further, the technology, tools and services available to malicious actors are continuously improving. Anyone, not just well-resourced States, can carry out impactful disinformation campaigns. The evidence received by the Committee on the use of digital technologies to spread falsehoods is set out more comprehensively in **Annex B**.

a. Amplification and targeting of online falsehoods

66. The Internet has made spreading information near instantaneous. Using everyday social media functions, almost anyone can spread information to a wide audience almost immediately. Online falsehoods may be organically amplified, or artificially amplified through coordinated methods and social media tools. Through targeted online advertising, falsehoods may be spread to influence people based on their known preferences. The algorithms of social media platforms then provide a further boost, by automatically promoting the visibility of popular posts to users. Social media platforms are a strategically attractive option for foreign States to spread disinformation. Online falsehoods often spread in cascades across multiple platforms, including but not limited to social media.

67. *Easy amplification.* Falsehoods may be spread further and faster using basic, everyday social media functions, such as posting, "sharing," "liking", re-tweeting, hyper-linking and hash-tagging. On Facebook, an individual can share a public post with up to 5,000 people with just one free click. In a full WhatsApp group, one can send a message to 256 people instantaneously. The borderless nature of the Internet means one can reach anyone anywhere in the world. The sheer size of some social media platforms provides a huge potential audience. On Facebook alone, the number of active monthly users was over 2 billion as at late 2017.

68. Falsehoods may be amplified when like-minded people with ideological motivations act in concert. For example, a false story about election fraud in the 2017 German Federal Election was amplified over Twitter in a "Twitter storm" by supporters of the German far-right political party, using re-tweeting and a hashtag.[28] Some may share a falsehood regardless of whether they believe it. In the 2016 US Presidential Election, it was said that many of those sharing attacks

[27] See Claire Wardle and Hossein Derakshan, "Information Disorder: Toward an interdisciplinary framework for research and policy making", *Council of Europe report* (27 September 2017), p 4.
[28] Ben Nimmo and Maks Czuperski, "#ElectionWatch: Final Hours Fake News Hype in Germany", *Digital Forensic Research Lab* (24 September 2017).

on Hillary Clinton based on falsehoods did not believe them, but hoped others would.

69. *False amplification.* Inauthentic social media accounts may be used to artificially amplify online falsehoods. Fake social media accounts are easily created, due to either lax or non-existent verification requirements. They usually seek to attract followers, to boost the size of their social network and audience.

70. Such efforts can be extremely successful. In the US, for example, one troll account on Twitter belonging to a fictitious "Jenna Abrams" had at one point over 70,000 followers, and was quoted by the New York Times, The Washington Post, Breitbart, and other high-profile media outlets. Another troll account that impersonated the Tennessee Republican Party had over 150,000 followers, and was re-tweeted by a Presidential candidate and senior members of his campaign. It attracted a much larger following than the Tennessee Republican Party's real Twitter account, which had 13,800 followers.[29]

71. Fake social media accounts may be run either by humans, known as "trolls", or by algorithms, known as "bots". The CPRP 2018 Report found that there was evidence of fake accounts used to create, disseminate and share disinformation online in almost all of the 48 countries surveyed.[30] Human "trolls" work in a coordinated manner to rapidly amplify a particular online falsehood. According to US authorities, a foreign troll factory was behind at least 3,814 fake Twitter troll accounts and at least 470 fake Facebook troll accounts that targeted the 2016 US Presidential Election.

72. Bots, on the other hand, are automated social media accounts that present as real users and post content without human intervention. They can play a range of roles, both useful and harmful. An example of generally useful bots are those that automatically aggregate content from different sources to produce news feeds.[31] Harmful bots may be used to spread online falsehoods in different ways. They can do so by manipulating social media algorithms through the strategic posting of certain keywords and causing certain content to trend, or by flooding social media hashtags with automated messages.[32] They may also spread online falsehoods by repeatedly amplifying selected accounts or other signals that they are designed to pick up. Mr Nimmo described this as "the digital equivalent of rushing in the same direction and bleating loudly."[33] For example, one Twitter bot posted 294 tweets

[29] "Sean Edgett's Answers to Questions for the Record", *Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism Hearing on Extremist Content and Russian Disinformation Online: Working to Find Solutions, October 31, 2017* (19 January 2018), pp 16 – 17.
[30] Samantha Bradshaw and Philip Howard, "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation", *University of Oxford, Computational Propaganda Research Project* (2018), p 11.
[31] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 46.
[32] Samantha Bradshaw and Philip Howard, "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation", *University of Oxford, Computational Propaganda Research Project* (2018), p 6.
[33] Ben Nimmo, "Why Bot Makers Dream of Electric Sheep", *Digital Forensic Research Lab* (27 June 2017).

on the Macron campaign leaks in three-and-a-half hours. Bots may also be managed by humans. In the 2016 US Presidential Election, a foreign troll factory that allegedly engaged in a disinformation campaign to influence the outcome of the election was said to have managed at least 50,258 bot accounts, in addition to thousands of troll accounts.

73. "Botnets" are another method to artificially amplify content. A "botnet" is a large number of accounts, usually numbering in the thousands, created to re-share the same post once each. Large botnets, counting thousands of accounts, were especially active during the 2016 US Presidential Election, pushing divisive, partisan and false content. Mr Nimmo also uncovered botnets that amplified "click-bait" content to attract users of a particular profile and steer them towards a money-making, "pay-per-click" advertisement site.

74. Notably, bot armies have been found not only in the US and Europe, but in countries elsewhere in the world, from Mexico and Venezuela, to the Middle East, South Africa, and Indonesia[34].

75. Bots can be difficult to detect, according to several expert representors. They have a short life-span, and new bots emerge quickly. There are "cyborg" accounts, which are bots that occasionally make their own posts appear more human. These elements of genuine human interaction make it even more difficult for such accounts to be detected and shut down.[35] Some sophisticated bots do not use codes in common with other bots, which makes them even more difficult to detect.

76. Troll and bot accounts can work together to achieve massive amplification of content. During the 2017 French Presidential Election, the #Macronleaks hashtag was used to guide Twitter users to false claims that the emails showed evidence of his offshore accounts, tax evasion and a slew of other nefarious activities.[36] The hashtag was amplified through a network of trolls and bots driven by the alt-right in the US.[37] It reached 47,000 tweets in just three and a half hours after the initial tweet.

77. The use of bots and false avatars by disinformation agents was described in detail in a recent case study analysis by ASERO Worldwide. According to the report, bots and avatars are often used to flood the social media profiles of targeted individuals with identical posts and messages originating from fake accounts. This is coupled with the technique of "feeding", which is a term referring to the use of social media functions such as "sharing", "liking" and "reacting" with the aim of manipulating a social media platform's algorithm in order to boost viewership of

[34] Kate Lamb, "Muslim Cyber Army: a 'fake news' operation designed to derail Indonesia's leader", *The Guardian* (13 March 2018).
[35] Samantha Bradshaw and Philip Howard, "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation", *University of Oxford, Computational Propaganda Research Project* (2018), p 12.
[36] Ben Nimmo, et. al, "Hashtag Campaign: #MacronLeaks", *Digital Forensic Research Lab* (5 May 2017).
[37] Ben Nimmo, "Why Bot Makers Dream of Electric Sheep", *Digital Forensic Research Lab* (27 June 2017).

the post. Such techniques are combined with VPN3 services in an attempt to mask their source.[38]

78. *Targeted advertising.* Online platforms such as Google and Facebook offer easy-to-use and cheap targeted advertising tools that anyone can use to send advertisements to specific users based on their known preferences. This is done by selecting targeting options that are provided by the advertising platform on its advertising interface, such as demographics, location, interests, and recent purchasing behaviour.[39] This is a form of "micro-targeting". As explained by strategic communications consultant Mr Nicholas Fang, online micro-targeting uses artificial intelligence programmes to obtain data on users' personal tendencies and characteristics. The data is then used to determine how to target different groups of people with tailored messaging.

79. Targeted advertising can be an influential and effective amplification tool. A study by network theorists showed that when falsehoods are initially aimed at those predisposed to believe them, they spread further.[40] According to the campaign of US Senator Toomey, their strategy of using Facebook Ads to customise messages to individual voter groups "significantly shifted" the intent of voters, and contributed to the senator's re-election.[41] Micro-targeting was identified by Mr Fang as a potential future threat generally.

80. Targeted advertising was a key tool of a foreign disinformation campaign during the 2016 US Presidential Election. Using US$100,000, a foreign troll factory was able to spread Facebook advertisements to 126 million Americans, including ones targeted at specific profiles.[42]

81. *Social media algorithms.* When falsehoods artificially amplified by these methods and tools gain popularity, they are then given a further boost by the algorithms of social media platforms, which are designed to automatically promote popular posts. All this enables the viral spread of falsehoods online.

82. A notable example given was of a conspiracy video that circulated after a shooting at a school in Parkland, Florida in the US in February 2018. The conspiracy video falsely claimed that a 17-year-old survivor of the shooting was not a genuine victim but an actor. The video was briefly pushed to the top of YouTube's Trending section, significantly increasing its visibility online.

---

[38] "Case Study Analysis: Fake News and Disinformation Campaign against a Leading Journalist", *ASERO Worldwide,* pp 2-4.
[39] "Facebook Advertising Targeting Options", *Facebook Business*; "Targeting your ads – AdWords Help" *Google Support*; "Ad targeting best practices for Twitter", *Twitter for Business*.
[40] "Why Fake News Spreads So Fast on Facebook: Ad Technology has weaponised disinformation", *Bloomberg (Op-Ed)* (31 August 2017).
[41] Adam Pasick, "Facebook says it can sway elections after all – for a price", *Quartz* (1 March 2017).
[42] Kate Conger and Dell Cameron, "Here are 14 Russian ads that ran on Facebook during the 2016 Election", *Gizmodo* (11 October 2017).

83. *Social media platforms an attractive option.* Social media platforms are a strategically attractive option for purported foreign disinformation outlets to reach their audiences, according to Dr Kevin Limonier from France (Associate Professor, French Institute of Geopolitics; Associate Researcher, Castex Chair of Cyberstrategy). Dr Limonier explained that media outlets from a foreign country were using techniques initially used to generate digital advertising revenue. They published "click-bait" articles that apparently had little to do with the news they usually carried. This allowed them to attract greater user engagement to their online platforms, which boosted their visibility in social media feeds and grew their audience due to social media algorithms.

84. In the same vein, the CPRP 2018 Report observed that the ability of social media platforms to directly reach large numbers of people, while simultaneously micro-targeting individuals with personalised messages, is what has caused social media platforms to be so attractive to foreign adversaries.[43]

85. *Online falsehoods cascade over different platforms*. Although social media platforms have been a key vector for the spread of online falsehoods, other online platforms have also been important. In the earlier example of the false article titled "Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement,"[44] the story spread through a network of "fake news" websites as well as on Facebook. In another example, Mr Nimmo shared how a false claim that the latest Russian technology could wipe out the entire US Navy was first posted on a Russian television station's website, then picked up by two British tabloids. Within hours, the story started to trend on social media, and quickly spread across the websites of a significant number of news outlets, comprising both mainstream and alternative media.

86. Closed messaging platforms such as WhatsApp were also identified by representors as playing an important role in the spread of online falsehoods, including in Singapore. The CPRP 2018 Report found that there is growing evidence of disinformation campaigns taking place over chat applications such as WhatsApp, Telegram and WeChat.[45]

87. Online social networks can enable falsehoods to be spread among diverse audiences. This was shown by a preliminary mapping by Dr Limonier of the "galaxy" of Twitter users who relayed content from two foreign newspapers in France alleged to be propagating foreign disinformation and propaganda. The mapping showed that the content of these foreign news sources was able to spread through different actors to reach a politically varied audience, comprising not only

[43] Samantha Bradshaw and Philip Howard, "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation", *University of Oxford, Computational Propaganda Research Project* (2018), p 4.
[44] See Craig Silverman and Jeremy Singer-Vine, "The True Story Behind the Biggest Fake News Hit of the Election", *BuzzFeed* (17 December 2016).
[45] Samantha Bradshaw and Philip Howard, "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation", *University of Oxford, Computational Propaganda Research Project* (2018), pp 3, 6 and 13.

the French nationalist far-right, but also users sharing different political opinions and of different political leanings.

b. Creation of low cost and high impact online falsehoods

88. Creating believable online falsehoods is much easier and costs much less than on traditional media. This is so for several reasons.

89. First, on social media, information is often shared amongst peers without verification of content or source. An online falsehood can be created simply by typing out some text online, or swapping the caption of a video or photograph. It could then easily find a believing audience on social media. Fabricated articles or misleading headlines may also take advantage of how information appears to Internet users. Even satire may be more difficult to identify when read off a social media feed, according to Dr Wardle. During the 2017 French Presidential Election, CrossCheck, a fact-checking project, found that people were disseminating falsehoods masquerading as satire in order to avoid fact-checks.

90. Second, consumer-friendly tools for creating audio-visual online content are readily available. Such tools have allowed relatively unskilled users to manipulate and distort visual media in ways that are very difficult to detect, according to various representors, including computer scientist Dr Hany Farid from the US (Professor & Chair, Computer Science, Dartmouth College).

91. For example, representors drew attention to free artificial intelligence tools that can convincingly simulate actual people to deliver messages that are not from the apparent sender, as well as easy-to-use software for editing and creating audio. There are already applications which allow users to feed a computer image and audio of a person to teach it to imitate that person's voice. There are video tutorials online to teach one how to use such applications.[46] Such software can make it relatively easy to transpose a picture of one person on an existing video to create a fake video (known as a "deepfake"). A Financial Times article described how such "deepfakes" can be easily used to put words and expressions on the face and mouth of a politician and influence elections.[47] One New York Times reporter said that creating a "deepfake" cost him less than US$100.

92. Third, online platforms such as websites and blogs can be created at relatively low cost. Purveyors of falsehoods can easily masquerade as genuine reporting outlets. For example, a website was created to mimic a genuine South African news site, and spread the false claim that South African President Jacob Zuma had resigned. This triggered a brief spike in the value of the South African rand. In Singapore, a student created a fake copy of a government website, and posted the false

[46] Roula Khalaf, "If you thought fake news was a problem, just wait for 'deepfakes'", *Financial Times* (25 July 2018).
[47] Roula Khalaf, "If you thought fake news was a problem, just wait for 'deepfakes'", *Financial Times* (25 July 2018).

announcement that Mr Lee Kuan Yew had passed away. Established international news outlets fell for the hoax and reported it to an international audience.

93. These low-cost and user-friendly methods can rival or exceed the influence of traditional media. A simple splicing edit to a video of then-incumbent Jakarta governor Basuki Tjahaja Purnama (popularly known as "Ahok") made it seem that he had committed blasphemy.[48] This fuelled rallies involving hundreds of thousands of people, and protests that turned violent.[49] In the US, doctored photographs were used to accuse the police of setting fire to a protestors' campsite, inflaming sentiments against the police.

c. Market for online disinformation tools and services

94. Online disinformation campaigns are now a profitable industry. Digital tools such as bots and botnets may be bought or hired for easy "plug and play". Individual services that require more manpower and skill are likewise available for a price. At the higher end of the scale are "hired guns" who offer package deals for the online manipulation of public opinion and voting outcomes. These online influence tools and services cost significantly less than conventional advertising and marketing, and achieve the same, if not greater, reach. With access to these markets, not only highly equipped States, but also ordinary people, can engage in sophisticated online disinformation campaigns.

95. <u>*Tools.*</u> Fake social media accounts are commonly used to spread falsehoods. These include accounts that have over some time, years even, been cultivated into convincing personas. These can include bots. There are commercial "bot herders" that hire out bots they create, some on a scale of thousands or tens of thousands of accounts. Mr Nimmo found that political posts in the lead-up to elections of leadership of the African National Congress were amplified by bots purchased from a commercial bot seller in the US.

96. <u>*Services.*</u> Ms Myla Pilao (Director, Core Technology Marketing, TrendMicro) gave evidence of the services available on the market. One example is "click farms", which comprise a large number of low-paid workers who click on links or posts. "Click farms" allow "click farm masters" to sell things like video views, "likes" and even votes. One can buy one million Instagram "likes" for only US$18, 1,000 WeChat "likes" for US$0.19, and 500 re-tweets for US$2. There are also content marketing services, which offer fake news articles for as little as US$15 to US$30 for 500 to 1,500 words.

97. More sophisticated services include "public opinion monitoring systems", which survey, research, and influence opinions in online forums and social media networks for between US$1,850 and US$4,175. Fake content can be made to

[48] Erwida Maulia, "Fake news charges emotionally driven Jakarta election", *Nikkei Asian Review* (13 February 2017), p 2.
[49] "Mass prayer rally in Jakarta against governor 'Ahok'", *BBC* (2 December 2016), p 2.

appear on legitimate news sites without appearing as paid content, although this costs a premium of more than US$20,000. TrendMicro estimated that one could use online propaganda to instigate a street protest in the US for US$200,000.

98. *"Hired guns".* The demand for online public manipulation has spawned syndicates such as the Saracen Cyber Team in Indonesia. This organisation was paid to spread falsehoods on social media to further the political agendas of their clients. According to the Indonesian authorities, Saracen is only one among many organisations profiteering in online falsehoods.

99. Some of these "hired guns" specialise in election interference to achieve their clients' desired election result, according to Dr Shashi Jayakumar (Head, Centre of Excellence for National Security, S. Rajaratnam School of International Studies (RSIS)). Dr Shashi referred to the case of Andres Sepulveda, a Colombian who had (since around 2006) rigged elections in Latin America using cyber methods.[50] Sepulveda bought and managed thousands of fake social media accounts, which he used to spread falsehoods on key domestic issues and policies, and create false impressions of public support (*i.e.* astro-turfing). He also used cloned websites to falsely smear members of rival campaigns. In several cases, the election outcome was that desired by his clients.

100. "Hired guns" may take less sinister though no less influential forms. Ms Jennifer Yang Hui (Associate Research Fellow, RSIS) described how online influencers, comprising "buzzers" and "micro-celebrities", were paid by politicians to promote messages that benefited their financiers, even at the expense of the facts. "Buzzers" were Twitter users with more than 2,000 followers who were paid to send short and personalised messages to potential customers, while "micro-celebrities" were social media celebrities who used online platforms to attract attention.

d. Digital technologies are improving continuously

101. The digital technologies available to malicious actors are improving continuously, which makes combatting the problem all the more difficult. Before the US Congress, Facebook CEO Mark Zuckerberg explained how there is an "ongoing arms race" with foreign actors who are constantly seeking to exploit online systems, and are only going to get better at doing so.[51]

102. In July 2018, Facebook announced that it had removed 32 accounts and pages from Facebook and Instagram on the basis that these accounts and pages were involved in "coordinated inauthentic behaviour" seeking to influence the 2018 mid-term US Congressional elections.[52] Facebook did not attribute the activity to

[50] See Jordan Robertson et al, "How to hack an election", *Bloomberg Businessweek* (31 March 2016).
[51] "Zuckerberg: Facebook is in 'arms race' with Russia", *BBC* (11 April 2018).
[52] "Removing Bad Actors on Facebook", *Facebook newsroom* (31 July 2018); "Facebook bans pages aimed at US election interference", *BBC* (31 July 2018).

any one group, but noted that these actors had "better operational security", "improved capabilities" and had been more careful to cover their tracks, as compared to the actors responsible for spreading disinformation during the 2016 US Presidential Election. Some methods employed in this instance included the use of VPNs and internet phone services, as well as paying third parties to run ads on their behalf. Facebook's chief security officer, Alex Stamos, commented on how offensive organisations would always seek to improve their techniques, once they have been uncovered.

(3) Impact of Online Falsehoods

103. Online falsehoods can have both short-term and long-term impact. They can cause different types of harm, to (a) national security, (b) public institutions, (c) individuals, and (d) businesses. The evidence received by the Committee on the impact of online falsehoods is set out more comprehensively in **Annex C**.

a. Immediate and "slow drip" effects

104. Online falsehoods can take effect over both the short term and long term, as highlighted by expert representors. An example of a falsehood that took effect immediately is the fake tweet about a bomb attack on the White House in 2013. This triggered a temporary crash on the stock market, wiping about US$136.5 billion off Standard & Poor's 500 Index.[53]

105. In contrast, "slow drip" falsehoods do not always cause an immediate impact on society; it may take a longer period of time to see their effect. Mr Nimmo described how they could "gradually inflame tensions and hollow out the political centre."[54] These falsehoods often promote or attack a particular point of view over time, and can change the views of individuals and society gradually. According to Dr Elmie Nekmat (Assistant Professor of Communications and New Media, National University of Singapore), exposure over time to falsehoods mixed with extremist or partisan views on social media can skew world views. To demonstrate this, Mr Nimmo referred to the 2017 case of a man, Darren Osborne, who drove a van into a crowd outside a London mosque. A UK court found that Osborne had been radicalised over the Internet by online hate speech against Muslims.[55] Police investigations found that Osborne had been researching material from far-right conspiracy theory websites and fake news websites in the weeks prior to the incident.[56]

106. Falsehoods may play on existing "slow burn" issues, such as simmering communal tensions, to create more serious crises in the long run. For example, in

[53] Peter Foster, "'Bogus' AP tweet about explosion at the White House wipes billions off US markets", *The Telegraph* (23 April 2013).
[54] Ben Nimmo, Appendix III: Written Representations, Paper No. 36, page B145, para 46.
[55] *Sentencing remarks of Mrs Justice Cheema-Grubb: R v Darren Osborne*, (2 February 2018), para 7(b).
[56] Kevin Rawlinson, "Finsbury Park-accused trawled far-right groups online, court told", *The Guardian* (23 January 2018).

Sri Lanka, fear was aroused by rumours that were circulating of a Muslim plot to sterilise and destroy Sri Lanka's Sinhalese majority. Against this backdrop, a Muslim restaurant owner, due to his unfamiliarity with the Sinhalese language, mistakenly admitted to putting sterilisation medicine in the food he served. Communal violence erupted as a result of this incident. The restaurant owner was beaten, his shop destroyed, and a local mosque was set on fire. The restaurant owner's "confession" was recorded and uploaded to Facebook, where it went viral, resulting in mobs in several towns burning mosques, shops, and homes owned by Muslims.[57]

107. In Myanmar, falsehoods spread on Facebook have stoked ethnic violence between Buddhists and Muslims. In 2014, an online claim that a Buddhist woman had been raped by one or more Muslim men provoked deadly mob violence in Myanmar, killing two people.[58] Falsehoods are reportedly stirring up fatal violence against the Muslim ethnic minority known as the Rohingya.[59] Indonesia and India have also seen the use of falsehoods to increase communal animosity and trigger or worsen serious crises, as described at [116] to [117] below.

108. Representors warned that falsehoods can progressively erode the harmony and cohesion between different communities, and can be used to undermine the credibility and trust in institutions, including the media.

b. Threats to national security

109. Online falsehoods by foreign States can harm national security when they seek to undermine a nation's sovereignty. Such falsehoods may interfere in a country's elections and domestic and foreign policies, or weaken the country's government and the resilience of the people to pave the way for the foreign State to gain control. Foreign disinformation campaigns are discussed in greater detail in Part I(B) of the report.

110. Whether or not a foreign State is behind it, online falsehoods may harm national security if they undermine social cohesion, incite public unrest or violence, or cause public alarm.

111. *Undermining of social cohesion.* Several representors highlighted how falsehoods have divided and polarised society. Dr Carol Soon (Senior Research Fellow, Institute of Policy Studies, Lee Kuan Yew School of Public Policy, National University of Singapore) and Mr Shawn Goh (Research Assistant, Institute of Policy Studies, LKY SPP, National University of Singapore) explained that, from their research, "deliberate online falsehoods often mirror the cracks and fissures

[57] Janet Guyon, "In Sri Lanka, Facebook is like the ministry of truth", *Quartz* (22 April 2018).
[58] Amina Waheed, "Rape Used as a Weapon in Myanmar to Ignite Fear", *Al Jazeera English* (28 October 2015); "Why is There Communal Violence in Myanmar?", *BBC* (3 July 2014).
[59] "In Myanmar, fake news spread on Facebook stokes ethnic violence", *Public Radio International* (1 November 2017).

that pervade each country", and "[exploit] the pain points found in political systems and societies, and capitalise on people's anxieties, doubts, fears and insecurities".[60] Falsehoods that target societal fault lines are like "throwing gasoline on fire", as a New York Times reporter observed.[61] These fault lines may be (i) political, (ii) economic or (iii) identity-based.

112. Political fault lines are often found between political party camps, and where there are controversial policy issues. For example, in the US, from 2015 to 2017, foreign disinformation agents allegedly spread over 9,000 social media posts on energy policies and climate change, stirring up environmental activist groups. They also allegedly exploited the US' long-running debate on gun control. After the February 2018 school shooting in Florida, the foreign agents were said to have amplified conspiracy theories claiming that the shooting had never happened and was instead a secret government operation.[62] Dr Mathew Mathews (Senior Research Fellow, Institute of Policy Studies, LKY SPP, National University of Singapore) observed that this made the debate on gun control in the US even more toxic than before.

113. Economic fault lines usually lie between the rich and the poor. For example, the Ukraine Crisis Media Centre found that pensioners and groups in poor economic conditions were vulnerable to foreign disinformation campaigns in Ukraine.

114. Identity-based fault lines are particularly potent. According to Dr Cherian George (Professor of Media Studies, School of Communication, Hong Kong Baptist University), who has studied hate propaganda for several years, "tribal identities [can be activated by simple falsehoods] in a way that is difficult to fight."[63] Disinformation is used in hate propaganda to keep us-versus-them attitudes simmering, and to make one group feel threatened or victimised by another group. False stories can be used to whip up indignation and outrage, instigating people to take action.

115. For example, sectarian and racist narratives used by online hoax campaigns are threatening social stability in Indonesia, according to RSIS researcher, Ms Yang. The hoaxes exploit long-running anti-Christian, anti-Chinese and anti-communist fault lines. As mentioned above, such falsehoods led to massive protests against 'Ahok' in Indonesia.

116. *Incitement of public unrest and violence.* Around the world, falsehoods that rupture societal fault lines have also often led to public unrest, and endangered lives. For example, in North Sumatra, angry mobs sought to burn down Chinese temples and Buddhist monasteries after a Chinese lady complained about noise

[60] Carol Soon and Shawn Goh, Appendix III: Written Representations, Paper No. 62, page B359, para 8.
[61] Andrew Kramer, "To Battle Fake News, Ukrainian Show Features Nothing but Lies", *New York Times* (26 February 2017).
[62] Sheena Frenkel and Daisuke Wakabayashi, "After Florida School Shooting, Russian 'Bot' Army Pounced", *New York Times* (19 February 2018).
[63] Cherian George, Appendix IV: Minutes of Evidence, page C696, para 5830.

from a nearby mosque.[64] According to Mr Septiaji Eko Nugroho from Indonesia, founder of fact-checking organisation MAFINDO, the violence was instigated by disinformation spread via chat applications.

117. India has seen a proliferation of online falsehoods that have inflamed communal unrest between Hindus and Muslims. For example, after the circulation of an offensive cartoon of Prophet Muhammad on Facebook, ongoing violence was given additional fuel by a photograph claiming to be of a Muslim man trying to disrobe a Hindu woman, when it was in fact a scene from a film.[65] In another incident, violence between Hindus and Muslims was worsened by an online video clip of two young men being killed by a mob in a clash that had purportedly arisen amidst the ongoing unrest, when it was in fact recorded in Pakistan several years before.[66] The violence eventually spread to neighbouring villages, leaving several dozens dead and over 40,000 displaced.[67]

118. In the US, there was the well-known Pizza-gate conspiracy theory, which fuelled anger amongst right-wing citizens who believed that political figures connected with the Democratic Party were allegedly running a paedophilia ring in a particular pizza restaurant. This led eventually to an angry American firing shots into the pizza restaurant. At a college in Minnesota, a fake note containing a racist threat against a black student led to campus-wide protests.[68] In the UK, as explained above, anti-Muslim conspiracy theories contributed to the radicalisation of Darren Osborne, who drove his van into a crowd outside a mosque.

119. *Instigation of public disorder and instability.* Falsehoods have caused public alarm, and in some cases threatened financial stability in the process. A group from Nanyang Polytechnic cited the example of a false claim in China that salt would ward off potential radiation poisoning from Japan's nuclear emergency. This triggered panic buying and led to a ten-fold increase in the price of salt. Falsehoods could also lead to a bank run, as pointed out by a group of SMU law students. This would both cause public alarm and impact financial stability.

120. Falsehoods that affect financial markets could lead to de-stabilising effects for the country. One example is the false tweet that the White House had been bombed, as mentioned at [104] above. The falsehood led to a massive fall in the stock market, which was fortunately quickly reversed.[69] Had the tweet not been quickly de-bunked, there could have been serious damage to investors and the financial system.

---

[64] Apriadi Gunawan, "Vihara, pagodas burned down, plundered in N. Sumatra", *The Jakarta Post* (30 July 2016).
[65] Sam Jawed, "The Vicious Cycle of Fake Images in Basirhat Riots", *AltNews* (7 July 2017); "What is Behind the Religious Violence in India's West Bengal?", *BBC* (11 July 2017).
[66] "Muzaffarnagar: Tales of Death and Despair in India's Riot-hit Town", *BBC* (25 September 2013).
[67] Pamposh Raina, "A Village in Muzaffarnagar Recounts Rape and Murder", *New York Times* (30 September 2013.
[68] "St Olaf: Racist note that prompted protests was fake", *Kare* (11 May 2017).
[69] Heidi Moore and Dan Roberts, "AP Twitter hack causes panic on Wall Street and sends Dow plunging", *The Guardian* (23 April 2013).

c. Harm to democratic institutions, free speech

121. One of the biggest threats that online falsehoods pose to society is their harm to the cornerstones of a well-functioning and democratic society. These would include citizen engagement in public discourse, trust in public institutions, and the right of citizens to have a representative government.

122. The damaging impact of online falsehoods on democratic institutions was a concern highlighted by several representors, including local law students and constitutional law academics such as Dr Thio Li-Ann (Professor of Law, Faculty of Law, National University of Singapore) and Associate Professor Eugene Tan (Associate Professor of Law, School of Law, Singapore Management University). Law students from the National University of Singapore (NUS) also quoted the observation of the UK House of Lords that "the working of a democratic society depends on the members of that society … being informed and not misinformed."[70] Dr Ullrich Ecker from Australia (Associate Professor, School of Psychological Science, University of Western Australia) termed it a "truism that a functioning democracy relies on a well-informed public."[71] Evidence on how online falsehoods harm various fundamental aspects of democracy is set out below.

123. *Damaging society's shared public space and impeding informed participation in public discourse.* Citizens should be able to engage in public debate and thereby participate in shaping their society. This is vital to a democracy. Dr Thio highlighted that the ability of citizens to engage in political discussions was important. Through this process, "citizens gain an understanding of public issues and are better equipped to participate in the workings of a democratic society."[72] She also explained the concept of a shared public space, as one "where plural viewpoints are exchanged, interrogated, debated, with all sides better understanding the complexities of a public issue and the range of positions taken on such questions."[73]

124. Crucially, the accuracy and diversity of information that citizens receive are pivotal to public discourse at two levels:

    a. First, as Dr Thio explained, information enables citizens to understand public affairs and issues of public interest.

    b. Second, by enabling people to understand the viewpoints of others, to achieve necessary compromise and accommodation. Accepting that

[70] Er Shengtian Rachel and Joel Jaryn Yap Shen, Appendix III: Written Representations, Paper No. 51, page B233.
[71] Stephan Lewandowsky et al, "Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era", *Journal of Applied Research in Memory and Cognition 6* (2017) 353, p 354.
[72] Thio Li-Ann, Appendix III: Written Representations, Paper No. 55, page B291, para 4.2.
[73] Thio Li-Ann, Appendix III: Written Representations, Paper No. 55, page B302-301, para 10.6.3(d).

society is plural, including in philosophies and world views, she emphasised that being exposed to a diversity of views was necessary "for understanding accurately where another citizen is coming from… and for facilitating compromise and overlapping consensus where possible." It enabled people to have "a common framework for social experience and a sense of a shared common good" in order to have "the ability to compromise and arrive at reasonable accommodations."[74] Dr Thio emphasised the importance of a commitment to pluralism, for a harmonious society.

125. Other representors made similar points. For example, lawyer Mr Darius Lee observed that "[f]or a healthy democracy to function, it must be fuelled by a healthy supply of accurate information from diverse sources."[75]

126. Online falsehoods can damage society's shared public space and public discourse in the following ways. *First*, they can make it difficult for people to understand each other, and inhibit diverse views from being shared.

a. According to Mr Nimmo and Dr Mathews, falsehoods can appeal to emotions and cause people to react with anger, and make people's emotions on an issue stronger than before. When people are angry, it could make it more difficult to have a rational debate. Falsehoods can therefore make public discourse ugly and lacking in civility.

b. Falsehoods can crowd out other voices, thereby preventing people from being exposed to a diversity of views, and discouraging pluralism in democratic debate. Political data scientists from Germany, Dr Simon Hegelich (Professor, Bavarian School of Public Policy, Technical University of Munich) and Mr Morteza Shahrezaye (Researcher, Bavarian School of Public Policy, Technical University of Munich) made the point that the flooding of social media platforms with negative comments seemed to have deterred those who were more sympathetic to the plight of refugees. The negative comments circulating online in Germany included anti-refugee falsehoods, such as the falsehood that a Syrian refugee who had taken a selfie with German Chancellor Angela Merkel was an ISIS terrorist.[76]

127. *Second*, falsehoods can erode trust in authoritative sources of information. This prevents the formation of a shared foundation of facts necessary for public debate.

a. There was the view that "the most salient danger" associated with "fake news" was that it "devalues and delegitimizes voices of expertise, authoritative institutions, and the concept of objective data – all of which

[74] Thio Li-Ann, Appendix III: Written Representations, Paper No. 55, page B302, para 10.6.3(a).
[75] Darius Lee, Appendix III: Written Representations, Paper No. 32, page B105, para 4.
[76] Stephanie Ott, "How a selfie with Merkel changed Syrian refugee's life", *Al Jazeera* (21 February 2017).

undermines society's ability to engage in rational discourse based upon shared facts." This was a finding of participants at a workshop by Yale's Information Society Project and the Floyd Abrams Institute for Freedom of Expression, which representors drew attention to.[77]

b. The Czech experience may offer evidence of such distrust in facts. A survey referred to by Mr Janda showed that 53% of Czechs believed that there was both pro-Russian and anti-Russian propaganda in the Czech public space and they could not trust anything.[78]

c. Falsehoods can lead to sections of the population relying on different realities in debates online. Dr Thio expressed concern that without a sense of solidarity and common framework of experience, "'tribes' championing single-issue agendas" would emerge.

d. Law academic Associate Professor Eugene Tan emphasised the need to ensure that public discourse in Singapore did not become a "post-truth" one, where the line between fact and fiction was dangerously blurred. He quoted the observation that "people are entitled to their own opinions but not their own facts."[79]

128. Several representors also spoke of deliberate efforts to undermine trust in the mainstream media. Mr Nimmo said that, in the UK experience, the distrust of mainstream media has been actively fostered by "alternative" news outlets from various political extremes, who have a shared interest in weakening the political centre and the credibility of established outlets. In Singapore, representatives from Singapore Press Holdings (SPH) and Mediacorp spoke of a "constant drip feed online" of attacks on the credibility of the mainstream media.[80]

129. *Third*, online falsehoods can cause citizens to disengage from public discourse altogether. Psychologist Dr Ecker cautioned that being exposed to large amounts of misinformation has been shown to have the psychological effect of making people stop believing in facts altogether, and decreasing their engagement in public discourse. This potential impact was also highlighted with concern by several other representors, such as Dr Thio, Mr Fang, and Dr Wardle.

---

[77] Zhulkarnain Abdul Rahim, Appendix III: Written Representations, Paper No. 80, page B477-478, para 10; Sui Yi Siong et al., Appendix III: Written Representations, Paper No. 130, page B1138, para 24, citing Baron, Sandra & Crootof, Rebecca, "Fighting Fake News: Workshop Report", *Yale Law School (The Information Society Project) and the Floyd Abrams Institute for Freedom of Expression* (7 March 2017).

[78] Kremlin Watch Program of the European Values Think-Tank, "Kremlin Hostile Disinformation Operations. Situational report on Czech Republic and Central European context", *European Values Think-Tank* (18 October 2016), p 12, citing poll by Czech Television (August 2016).

[79] Eugene Tan, Appendix III: Written Representations, Paper No. 150, page B1315-1316, para 27.

[80] Warren Fernandez, Appendix IV: Minutes of Evidence, pages C492, para 4251; Walter Fernandez, Appendix IV: Minutes of Evidence, page C510, para 4380.

130. *Obstructing public institutions in policy-making and the delivery of public services.* Society depends on public institutions to carry out their governance functions and make policies in the public interest. The public also relies on public institutions as a key source of information. Falsehoods can obstruct governance functions by obfuscating public debate, as well as by eroding trust in public institutions. The erosion of domestic trust in public institutions diminishes the ability of public institutions to defend their reputations, respond effectively to threats and crises, and to govern. It also weakens the role of public institutions as a source of information to foster a common foundation of facts for public debate.

131. Several representors expressed concern about the impact of online falsehoods on trust in public institutions.

a. By undermining trust in public institutions, governance in a country can be weakened. Hence, hostile actors often seek to weaken public institutions by undermining public trust in them, as explained by experts Dr Janis Berzins from Latvia (Director, Center for Security and Strategic Studies, The National Defense Academy of Latvia) and Mr Janda. For example, falsehoods spread in the Czech Republic that sought to discredit the US and NATO have apparently found success. According to Mr Janda, 50.2% of Czechs believed that the US was responsible for the influx of Syrian refugees, and 38% believed that the Ukrainian crisis was caused by the US and NATO. This shows that major sections of the population can be influenced by falsehoods, over time.

b. The influence of falsehoods on public trust has been shown by psychological research. According to psychologist Dr Ecker, conspiracy claims have been found to adversely affect trust in public services and institutions, even those unconnected to the claims.[81] Further, the mere exposure to falsehoods has been found to make people less likely to accept official information.[82]

c. Dr Damien Cheong (Research Fellow, National Security Studies Programme, RSIS), identified public institutions in Singapore as a potential target of disinformation operations. He said that incidents targeting trust in the police had occurred.

132. Several representors expressed concern about how the impact of online falsehoods on public understanding of issues of public interest in turn affected policy-making.

[81] Lewandowsky et al, "Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era", *Journal of Applied Research in Memory and Cognition* 6 (2017) 353, p 355.
[82] Lewandowsky et al, "Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era", *Journal of Applied Research in Memory and Cognition* 6 (2017) 353, p 355.

a. Information forms the basis for the political and societal decisions made by individuals, social groups and communities.[83] Dr Thio observed that for elected representatives, information was important to effective public debate and informed policy-making.

b. By undermining deliberative political debate, online falsehoods "destroy the feedback loop between the government and the governed", as pointed out by undergraduates from the NUS Law Faculty.[84]

c. Some examples show this impact in practice:

    i. In the UK's 2012 Leveson Inquiry on the culture, practices and ethics of the press, Lord Leveson found that the cumulative impact of inaccurate news, whether online or offline, about political issues could have serious consequences for policy-making. In particular, the Leveson Inquiry report highlighted how false stories published about Europe by some parts of the press made it difficult for the political leaders of that period to adopt particular policies or achieve certain political ends in relation to the EU.[85]

    ii. Falsehoods may erode overseas support for countries, cutting them off from important aid and economic cooperation. For example, according to Kremlin Watch's Mr Janda, falsehoods portraying the Ukrainian government as fascist have been spread in the Czech Republic. According to Mr Janda's research, a quarter to a third of Czechs believed that the Ukrainian government is fascist. This was said by Mr Janda to have impeded the Czech government's ability to render humanitarian aid to Ukraine.

    iii. Similarly, falsehoods may have contributed to Ukraine's inability to enter into a trade agreement with the EU. The issue had been the subject of a referendum in the Netherlands, and a significant number of Dutch people had voted against the trade agreement with Ukraine. Ukrainian foreign ministry officials suggested that this was due to Dutch voters' beliefs that the Ukrainian government was corrupt and that Ukraine had shot down MH17, an event that killed 193 Dutch citizens.[86]

---

[83] Lewandowsky et al., "Misinformation and its Correction: Continued Influence and Successful Debiasing", *Psychological Science in the Public Interest* 13(3) (2012) 106, p 107.

[84] Er Shengtian Rachel and Joel Jaryn Yap Shen, Appendix III: Written Representations, Paper No. 51, pages B232-233.

[85] "An Inquiry into the Culture, Practices and Ethics of the Press – Report", *The Leveson Inquiry* (November 2012), Volume II, [9.55].

[86] "Putin's Asymmetric Assault on Democracy in Russia and Europe: Implications for US National Security", *US Congress* (10 January 2018), pp 113-115.

133. *Undermining citizens' right to a representative government and representative politics.* In a democracy, citizens should be able to exercise their right to vote in an informed manner. This, as pointed out by Dr Thio, constitutes effective participation in the political process. Elections and other national voting processes such as national referendums are important exercises of popular sovereignty and self-determination, as explained by Dr Gillian Koh (Deputy Director, Institute of Policy Studies, LKY SPP, National University of Singapore). Informed voting is important for a genuinely representative government that has legitimacy in the eyes of the people. Without this legitimacy, instability and weakened governance would ensue.

134. Several representors highlighted how online falsehoods undermine these fundamental means of self-determination and political participation. For example:

a. Dr Thio emphasised that when one's vote was based on misinformation about an electoral candidate, one's "positive liberty" to effectively participate in the political process in an informed manner was thwarted by the confusion caused by the falsehood.

b. NUS law undergraduates highlighted that deliberate online falsehoods have undermined representative government, as voters were unable to make informed choices between competing candidates and policies.

135. The Committee notes several examples of the prevalence of online falsehoods in elections and other fundamental voting processes.

a. Falsehoods have been used to try to cast doubt on the legitimacy of the outcome of an election. They have sought to do so in two ways. First, they may make outright claims of vote rigging and lack of due process. For example, Mr Nimmo has written about how false claims that Scotland's 2014 independence referendum was rigged, or did not meet international standards, led to a petition for a re-vote that gathered over 100,000 signatures (a portion of which appeared to have been generated by bots).[87]

b. Second, falsehoods may inundate the elections in large volumes, raising questions as to whether voters were induced by falsehoods to vote in a particular way. The doubt created over whether voters were equipped with what they needed to make good decisions may impact the legitimacy of the outcome.[88] For example, in the 2016 US Presidential Election, research by the Oxford Internet Institute found that as a whole, more misinformation and polarising and conspiratorial content was being

[87] Ben Nimmo, "#ElectionWatch: Scottish Vote, Pro-Kremlin Trolls", *Digital Forensic Research Lab* (13 December 2017).
[88] Philip Howard and Bence Kollanyi, "Social media companies must respond to the sinister reality behind fake news", *The Guardian* (1 October 2017).

shared than professionally produced news in the country.[89] Average levels of misinformation were disproportionately higher in swing states than in uncontested states.[90] In Michigan, a swing state, the amount of professionally researched political news and information shared was smaller than the amount of "junk news" shared.

c. Mexico has reportedly seen a "sea of misinformation" on multiple online platforms, in advance of its Presidential Election in July 2018. Rival candidates have been the target of these false stories. President Trump's national security adviser has warned of foreign meddling in this election that was intended to create trouble along the US-Mexico border. One New York Times report observed that "whatever the impact on polls, the spread of lies stains public debate…"[91]

136. The evidence on whether falsehoods affect the way people vote is unclear, at this point.[92] A study by researchers from Ohio State University found that misinformation during the 2016 US Election had a very strong correlation to the voting behaviour of a particular subset of voters, namely, supporters of Obama in the 2012 US Election.[93] If targeted Facebook Ads can "significantly shift" voter intent, as Facebook claims,[94] it seems likely that falsehoods can in some situations change votes. Another study found that messages encouraging people to vote might influence decisions on whether to vote at all.[95]

137. A few representors expressed scepticism about whether falsehoods can actually influence people's voting behaviour. Two studies, which respectively concerned the 2016 US Presidential Election and Brexit, were cited for the proposition that falsehoods do not influence people's voting behaviour, or at the most, did so only at the margins.[96] However, neither study clearly concludes that falsehoods do not influence voting behaviour.[97] Even so, the Committee takes the point that not

[89] Philip Howard et al, "Social Media, News and Political Information during the US Election: Was Polarizing Content Concentrated in Swing States?" *Computational Propaganda Research Project, Data Memo 2017.8* (28 September 2017), p 3.

[90] Philip Howard et al, "Social Media, News and Political Information during the US Election: Was Polarizing Content Concentrated in Swing States?" *Computational Propaganda Research Project, Data Memo 2017.8* (28 September 2017), p 4.

[91] Ioan Grillo, "Fake News Crosses the Rio Grande", *New York Times* (3 May 2018).

[92] See also Briony Swire et al, "Processing political misinformation: comprehending the Trump phenomenon", *Royal Society Open Science* (March 2017); Hunt Allcott and Matthew Gentzkow, "Social Media and Fake News in the 2016 US Election", *Journal of Economic Perspectives*, Vol. 31, No. 2 (Spring 2017).

[93] Richard Gunther et al, "Fake News Did Have a Significant Impact on the Vote in the 2016 Election", *Ohio State University*.

[94] Adam Pasick, "Facebook says it can sway elections after all – for a price", *Quartz* (1 March 2017).

[95] Robert Bond et.al, "A 61-million-person experiment in social influence and political mobilisation", *Nature* (13 September 2012).

[96] Andrew Guess et al, "Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign" (9 January 2018); Vidya Narayanan et al, "Russian involvement and junk news during Brexit", *Computational Propaganda Data Memo 2017.10* (19 December 2017).

[97] One study sought to assess the extent of the "filter bubble" effect online by comparing the proportion of fake news consumed by individuals with their exposure to other sources of information. What the study does not

everyone can or will be taken in by falsehoods, and that levels of discernment would naturally vary across a population.

138. Nonetheless, the Committee's view is that the fundamental question is whether such online falsehoods should be allowed in the public space, if it was deliberately intended to mislead, and particularly if such falsehoods have serious consequences similar to those that have manifested in many places, leading to violence and bloodshed, loss of lives and the polarisation of societies. Some of the evidence in this regard has been set out at [106]-[107] and [116]-[118] above.

d. Harm to individuals

139. Falsehoods have harmed individuals at a personal level in different ways. At a fundamental level, falsehoods can confuse the decisions people make, and affect how people interact with the world around them. Falsehoods have also harmed people by making them the target of harassment and insults, causing them anxiety and leading them to make decisions that are bad for their health and well-being.

140. *Interference in individual decision-making.* Falsehoods can affect people in fundamental and everyday ways. They tend to influence decisions people make, such as how people participate in the political process. They can make people feel more concerned or threatened than warranted. It is reportedly becoming increasingly difficult even for experienced and well-informed news consumers to reliably distinguish accurate information from false information. In that regard, a 2017 US survey cited by representatives from TrendMicro showed that even when respondents felt they could tell fake from real news, many experienced considerable confusion.

141. *Provocation of harassment and insults.* Several representors recounted the distress they were subjected to as the result of falsehoods. According to one account, falsehoods posted in an online forum about the representor, a woman, led to her being sexually harassed.

142. Individuals or groups, from politicians and celebrities to ordinary people, may suffer public humiliation as a result of falsehoods put out by website operators seeking financial gain, as observed by the representatives from TrendMicro. This reflected the personal experience of representor Mr Prakash Hetamsaria, whose

suggest is that exposure to other sources of information meant that the individuals were not influenced by the falsehoods. The study in fact found "fairly widespread exposure to fake news websites" among Americans, and that fact-checking largely failed to effectively reach consumers of fake news (Andrew Guess et al, "Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign" (9 January 2018)). The other study sought to assess the extent of junk news on Twitter in the lead-up to Brexit, and the extent to which related Twitter conversations had links to information from a foreign country. It found that there was little evidence of links to this foreign country. This is not evidence of the influence of falsehoods on people. In fact, the study expressed concern about "the large number of accounts both human and automated, that shared polarizing and provocative content over the social media platform in days leading up to the referendum." (Vidya Narayanan et al, "Russian involvement and junk news during Brexit", *Computational Propaganda Data Memo 2017.10* (19 December 2017).

photograph was posted on the *All Singapore Stuff* website and who was falsely identified as a new citizen disappointed with Singapore and considering giving up his citizenship. The article was shared over 44,000 times. Mr Hetamsaria and his family, including his young daughter, were impacted by the xenophobic comments that followed. The falsehood hence also inflamed xenophobic and anti-immigrant sentiments in Singapore.

143. *Cause anxiety*. Falsehoods can also have the effect of causing anxiety in people. In July 2018, the databases of SingHealth – the largest group of healthcare institutions in Singapore – were hacked. Personal particulars such as the names, NRIC numbers, addresses, gender and race of 1.5 million patients were stolen.[98] However, in the aftermath of the cyber-attack, some patients received SMSes falsely claiming that, in addition to the particulars above, their phone numbers, financial details and medical records had also been accessed.[99] SingHealth eventually clarified in a Facebook post that the SMSes were untrue.

144. *Harming of health*. Falsehoods can threaten the health of individuals. Mr Nugroho of Mafindo highlighted how quack procedures promoted online have led to deaths in Indonesia, and patients have declined to continue with medical treatment because of what they read on the Internet. A group from Nanyang Polytechnic observed that falsehoods affected public health tools such as vaccines, by drowning out expert voices.

145. The American Journal of Public Health study referred to at [38] above which described how foreign troll accounts had spread false and unverified content about vaccines also shows how falsehoods can harm public health. One of the authors of the study noted that the foreign troll accounts, by "playing both sides" in the vaccination debate, "erode public trust in vaccination" and "[expose] us all to the risk of infectious diseases".[100] This should be seen in light of the views of experts who have attributed the recent surge in cases of measles in Europe to the drop in the number of people being vaccinated.[101] Despite the abundance of scientific evidence in favour of immunisation,[102] falsehoods relating to the benefits of vaccination may lead to people resisting getting vaccinated and harming public health as a result.

146. *Causing of financial harm*. Falsehoods that affect financial markets may have a wide-scale impact on a country's financial stability and on businesses, as well as a deep impact on individual investors who suffer financial losses, as pointed out by one student representor.

---

[98] "SingHealth cyber attack: How it unfolded", *The Straits Times* (20 July 2018).
[99] "SingHealth warns of fake SMS claiming access to phone numbers, financial details", *ChannelNewsAsia* (20 July 2018).
[100] "Russia trolls 'spreading vaccination misinformation' to create discord", *BBC* (24 August 2018).
[101] "Russia trolls 'spreading vaccination misinformation' to create discord", *BBC* (24 August 2018).
[102] "Russia trolls 'spreading vaccination misinformation' to create discord", *BBC* (24 August 2018).

e. Harm to businesses

147. Representors such as the Singapore Corporate Counsel Association and NTUC FairPrice emphasised how corporations too are not spared from the negative consequences of deliberate online falsehoods. Falsehoods may harm the reputation of businesses, erode customers' confidence, goodwill and trust, and cause financial loss, potentially transferring costs to consumers. They may also go beyond private concerns and trigger concerns over public health and safety.

148. *Triggering of alarm over food product safety.* False claims about the safety of food products appear common. Examples given by representors include falsehoods in Singapore, China, Malaysia and the US that food products were made of plastic, contained harmful lead, or contained parasites. Such claims caused needless public alarm.

149. *Straining of ties with customers.* Clever hoaxes can create tensions between businesses and their customers. NTUC FairPrice recounted how a doctored image of a Pasar pork product with a Halal sticker label repeatedly surfaced online, suggesting that NTUC FairPrice was religiously insensitive. In another case, perpetrators impersonated NTUC FairPrice and claimed that NTUC FairPrice was holding a survey with vouchers to be won. People fell for the ruse, and subsequently sought to claim the vouchers from NTUC FairPrice.

150. *Smearing of business reputation.* Business competitors may use falsehoods to target their competitors, as explained by the representative from TrendMicro. One method used is the posting of falsehoods undermining competitors in the comments sections of review websites.

151. *Causing of financial loss.* Several representors noted that businesses wasted manpower and resources when dealing with falsehoods. NTUC FairPrice agreed that this could possibly lead to costs to consumers, and hoped to avoid this outcome.

(4) Difficulties in Combatting Online Falsehoods

152. Online falsehoods are difficult to combat. The evidence before the Committee pointed to three key difficulties: (a) human cognitive tendencies, (b) the weakness of truth compared with falsehoods, and (c) the further and faster reach of falsehoods than the truth. These difficulties apply offline as well, but they tend to be greater in the online world. The evidence received by the Committee on the difficulties in combatting online falsehoods is set out more comprehensively in **Annex D**.

a. Human cognitive tendencies

153. Individuals are their own first line of defense against falsehoods. However, people often use mental shortcuts when processing information.[103] Conditions on social media encourage people to rely more on these mental shortcuts, making it easier to fall prey to falsehoods online than offline.

154. *Mental shortcuts.* There is an innate tendency towards confirmation bias, which leads people to believe information consistent with their preferences and world views. Dr Carol Soon described such world view-consistent information as having the same effect as falling in love, having sex, or eating chocolate. One of the earlier well-known experiments demonstrating this effect was conducted at Stanford University in 1976. When presented with two contradictory sets of data on the deterrent effect of capital punishment, students who supported capital punishment found the pro-deterrence data more credible than the anti-deterrence data. The reverse was true for students who were against capital punishment.[104] Confirmation bias has been demonstrated by other studies since then.[105]

155. Falsehoods tend to be believed when seen repeatedly. This is also known as the illusory truth effect. This can occur even when people are knowledgeable about the topic the falsehood relates to. The more often people see the falsehood, the stronger its effect, especially if they see it from different sources.

156. People tend to believe falsehoods in order to conform to the expectations of those they are close to (conformity cascades). People also tend to believe falsehoods because many others do so (informational cascades).

157. Finally, the ability to detect falsehoods is commonly overestimated. This is known as optimism bias. The above tendencies may therefore exert greater influence than anticipated.

158. *Heuristic tendencies are greater online.* Conditions on social media platforms encourage these tendencies. It is common for people to be repeatedly exposed to the same articles shared by others within their social networks. Due to conformity and informational cascades, many regard social media posts as trustworthy sources of information, despite the absence of traditional editorial verification.

159. An individual's defences against falsehoods may be weaker online than offline. A large volume of information is usually received online, especially on social media.

---

[103] Stephan Lewandowsky et al, "Misinformation and Its Correction: Continued Influence and Successful Debiasing", *Psychological Science in the Public Interest* 13(3) (2012) 106, pp 107-108; Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 18.

[104] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 20.

[105] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 21.

The common use of re-posting and re-tweeting on social media has made people accustomed to sharing information online without knowing its original source.

160. These conditions increase reliance on mental shortcuts. Today, the number of "likes", shares and re-tweets have frequently become compelling indicators of credibility. The identity of the original source tends to become less salient to Internet users. In that regard, research has shown that the identity of the closest source, which on social media is often friends, family, or acquaintances, exerts the most influence on how people assess the credibility of information.

161. When online, people tend to engage more in "skimming" rather than the "deep processing" required for critical thinking. People are less able to accurately evaluate the credibility and accuracy of online information.[106] It is generally easier for people to accept what they read online as true, than to take the effort to verify and reject it as false.

162. The highly educated are not immune to human cognitive flaws. Dr Carol Soon's research showed that even people with higher education levels could be susceptible to online falsehoods, in some cases even more so than others (see [167] below). A research study by Stanford University found that highly-educated people often misjudged the credibility of websites based on how the website looked.[107] Mr Nugroho shared from his own experience how online falsehoods in Indonesia also impacted educated people, who may be well-versed in particular topics, but not others. Some representors had a differing view that Singaporeans were less vulnerable to falsehoods compared to people in other countries because of our higher levels of education; however, the basis of these views was not explained.

b. Weakness of truth compared with falsehoods

163. Truth is generally weaker than falsehoods due to human psychology, and conditions online.

164. <u>*Human psychology.*</u> There are several psychological reasons why falsehoods may have a stronger effect on us than the truth. *First*, the influence of falsehoods is by its nature difficult to reverse, as shown by substantial psychological research. Exposure to misinformation can have long-term effects, while corrections may be short-lived.[108] Even when people believe a correction, they may forget what is true

[106] Ullrich Ecker, "Why rebuttals may not work: the psychology of misinformation" *44(2) Media Asia (2017) 79*, p 2.
[107] Sam Wineburg and Sarah McGrew, "Lateral Reading: Reading less and learning more when evaluating digital information", *Stanford History Education Group Working Paper No. 2017-A1* (September 2017).
[108] "Combating fake news: An agenda for research and action", a conference held at the Harvard Shorenstein Centre on Media, Politics and Public Policy (17-18 February 2017); Emily Thorson, "Belief echoes: The persistent effects of corrected misinformation", *Political Communications* 13(3) (2016) 810.

and "re-believe" the falsehood.[109] Falsehoods tend to trigger more emotions, especially negative emotions,[110] making them generally harder to correct, as such falsehoods leave strong impressions.

165. *Second*, people engage in motivated reasoning, which means finding justifications for their existing wrong conclusions, despite conflicting facts. People tend to reject corrections when they are inconsistent with their beliefs. For example, Mr Nugroho from fact-checking body Mafindo found it very difficult to persuade radical communities with ideological agendas in Indonesia. They would strongly defend their positions, even though these positions were factually false.

166. *Third*, in some cases, corrections can backfire, by increasing people's belief in the falsehood. For example, one study found that conservatives presented with false information that Iraq possessed weapons of mass destruction became even more likely to believe this claim after reading a news article correcting the falsehood.[111]

167. Such biases are not found only in those with extreme views. Dr Carol Soon acknowledged that these could apply to all sectors of the population, including the middle ground. Her research had also found evidence that there may be a correlation between higher levels of education and resistance to corrections that conflict with existing beliefs,[112] showing that the educated may not be less susceptible to prejudices and biases. Similarly, a group from Nanyang Polytechnic cited the observation in a research report published by the Tow Center for Digital Journalism that all people have emotional resistance to being wrong.

168. *Biases are worsened by conditions online.* Conditions on social media can encourage motivated reasoning. It has been found that interacting within an online cluster of like-minded people amplifies polarisation and heightens intolerance to different views. One study found that within such online clusters, few users would engage with posts that de-bunked falsehoods. Those who did instead reacted negatively to the de-bunk.[113] This finding was supported by research by political data scientist Dr Hegelich, who observed that within online clusters, responses to different views usually involved a "frantic curtailment" of those views, and escalation of ideological conflict.[114]

---

[109] Briony Swire et al, "Processing political misinformation: comprehending the Trump phenomenon", *Royal Society Open Science* (March 2017); "Combating fake news: An agenda for research and action", a conference held at the Harvard Shorenstein Centre on Media, Politics and Public Policy, 17-18 February 2017; Thorson, "Belief echoes: The persistent effects of corrected misinformation," *Political Communications* 13(3) (2016) 810.
[110] Soroush Vosoughi et al, "The spread of true and false news online", *Science* 359, 1146-1151 (2018), p 1.
[111] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 32.
[112] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 54.
[113] Fabiana Zollo et. al, "Debunking in a World of Tribes", *PLoS ONE* 12(7) (24 July 2017).
[114] Simon Hegelich and Morteza Shahrezaye, "Disruptions to political opinion", *Konrad Adenauer Stiftung* (July 2017), p 6.

169. The real influence that online "echo chambers" can have on people's beliefs and biases was supported by a study by researchers from the University of Warwick, who investigated anti-refugee attacks in Germany over a two-year period. The study cited an incident involving a young individual living in an otherwise pro-refugee German town, who broke into a refugee group house and tried to set it on fire, after "isolating himself in an online world of fear and anger", and in particular, "his Facebook echo chamber". The study also found that there would be about 50% more attacks on refugees in towns where per-person Facebook use was one standard deviation above the national average.[115] According to the researchers, their findings suggested that social media has played a role in propagating violent hate-crimes and motivating real-life violence.[116]

170. One representor cited a Google-funded study to suggest that people in Singapore, being in a highly wired society, were less susceptible to the influence of falsehoods. The Committee could not agree with this suggestion, as it considered that the study focused on how search engines influenced public opinion and political view-points.[117] It found that Internet users were generally exposed to and sought out diverse sources of information online.[118] The study did not address whether these users were influenced by the falsehoods they were exposed to, or deal with the psychological research on people's responses to falsehoods. In fact, the study found that "the number of people who report not fact-checking is sizeable and is a potential reason why fake news has been able to spread."[119] The Committee thus opined that this study needs to be appreciated alongside the psychological research-based studies or observations submitted as evidence above.

c. Further and faster reach of falsehoods

171. The speed at which falsehoods travel online was a key concern of several representors. Corrections usually lag behind falsehoods, for reasons that are often difficult to overcome. This hinders our ability to mitigate and remedy the damage done.

172. *Corrections lag behind the falsehood.* A 2018 study by the Massachusetts Institute of Technology (MIT) found that falsehoods spread farther, faster and deeper than the truth in all categories of information, and especially for politically-false news.[120] In particular, it found that falsehoods are 70% more likely to be re-

[115] Karsten Muller and Carlo Schwarz, "Fanning the Flames of Hate: Social Media and Hate Crime", *University of Warwick Working Paper Series* (May 2018), p 21.
[116] Karsten Muller and Carlo Schwarz, "Fanning the Flames of Hate: Social Media and Hate Crime", *University of Warwick Working Paper Series* (May 2018), p 42.
[117] William Dutton et al., "Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain, and the United States" *Quello Center Working Paper No. 5-1-17* (1 May 2017), p 5.
[118] William Dutton et al., "Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain, and the United States" *Quello Center Working Paper No. 5-1-17* (1 May 2017), p 5.
[119] William Dutton et al., "Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain, and the United States" *Quello Center Working Paper No. 5-1-17* (1 May 2017), p 130.
[120] Soroush Vosoughi et al, "The spread of true and false news online", *Science* 359, 1146-1151 (2018), p 1.

tweeted than the truth.[121] Another 2018 study by the University of Buffalo found that this phenomenon can be observed as well during disasters (such as during Hurricane Sandy and the Boston Marathon bombing). The researchers examined more than 20,000 tweets sent during such disasters, and found that between 86 to 91 percent of users would retweet or like false tweets, while less than 20 percent of the same users would offer any clarification even after the false tweet had been debunked on Twitter and traditional news media outlets.[122] Earlier research by a well-known expert on fake news compared the online spread of several false stories with that of corrections by major news outlets. The same conclusion was reached – the truth cannot outrace the false.[123]

173. This conclusion was supported by representors such as Dr Kalina Bontcheva from the United Kingdom (Professor, Department of Computer Science, University of Sheffield), who is developing technology to automatically detect online falsehoods. Locally, Mr Hetamsaria's experience revealed how the falsehood about him was shared over 44,000 times on Facebook, while his Facebook clarification was shared only a handful of times.

174. *Reasons for lag are difficult to overcome.* There are two reasons why the truth often lags behind the false. *First*, falsehoods generally enjoy an inherent time advantage. This is in some cases worsened by the difficulty of identifying a falsehood. For example, it often takes an expert hours to conclude if a photograph is fake or authentic.[124] Yet, as computer scientist Dr Farid points out, "on the Internet, two hours is an eternity [and] things go viral in a matter of minutes or hours".[125] During the 2017 Catalan independence referendum, Mr Nimmo recounted how a fake photograph of police pushing back against demonstrators under a Catalan flag was uploaded by a Twitter user. Within an hour and a half, a Spanish fact-checking organisation had managed to tweet the truth – that the image was a fake, with the flag included in the photograph using Photoshop. However, the tweet containing the truth was retweeted over 3,700 times, while the fake was retweeted over 12,600 times.[126] Established news organisations are also experiencing the pressure of having to respond quickly to falsehoods, when verifying and cross-checking information online is in fact a heavily resource-intensive one.

---

[121] Soroush Vosoughi et al, "The spread of true and false news online", *Science* 359, 1146-1151 (2018), p 4.
[122] Wang Bairong & Zhuang Jun, "Rumor response, debunking response, and decision makings of misinformed Twitter users during disasters", *Natural Hazards* (2018) 93:1145; see also Cory Nealon, "During disasters, Twitter users likely to spread falsehoods", *University at Buffalo* (16 May 2018), available at <https://www.buffalo.edu/ubnow/stories/2018/05/twitter-disasters.html>.
[123] Craig Silverman, "Lies, Damn Lies, and Viral Content", *Tow Center for Digital Journalism, Columbia Journalism School* (2015), pp 129-131.
[124] Hilke Schellmann, "The dangerous new technology that will make us question our basic idea of reality", *Quartz* (5 December 2017), available at <https://qz.com/1145657/the-dangerous-new-technology-that-will-make-us-question-our-basic-idea-of-reality/> (Interview with Hany Farid)
[125] Hany Farid, Appendix IV: Minutes of Evidence, page C625, para 5313.
[126] Ben Nimmo, "#Election Watch: Fake Photos in Catalonia?", *Digital Forensic Research Lab* (23 October 2017).

175. *Second*, people are less likely to share corrections due to psychological factors. Due to confirmation bias, information that is consistent with beliefs and world views is often shared and sought more than information that is inconsistent with these beliefs.[127] Due to negativity bias, negative information (which falsehoods usually are) is usually shared more than positive information (which corrections usually are). This may explain why, in the US during the 2016 US Presidential Election, the most popular fake election news stories garnered more engagement on social media than the most popular true election news stories.

176. *Ability to mitigate and remedy the damage is hindered.* The slower speed and reach of corrections have three implications. *First*, falsehoods often cause damage long before corrections can be put in motion.[128] *Second*, corrections cannot reach people fast enough to stop them from unwittingly spreading the falsehood.

177. *Third*, corrections are less likely to reach those exposed to the falsehood. A 2018 study by US and UK academics of selective exposure to misinformation during the 2016 US Election found that *none* of the respondents who read fake news articles saw the de-bunks for those falsehoods.[129] Similarly, a study of the 2017 French Election showed that there was almost no overlap between the audience for the rumour that then-Presidential candidate Emmanuel Macron was funded by Saudi Arabia, and the audience of its de-bunk.[130]

d. Social transformations caused by the digital revolution

178. Underpinning the difficulties in combatting online falsehoods are social transformations caused by the digital revolution. How the Internet and social media tend to increase human cognitive biases has been explained above. In addition, the digital revolution has led to online "echo chambers" on social media, the disruption of the news ecosystem, and fundamental changes to the nature of political discourse. This has in turn created fertile conditions for online falsehoods to gain traction.

179. *Online "echo chambers"*. Online "echo chambers" are said to facilitate the spread of deliberate online falsehoods and accentuate the difficulties in combatting them. "Echo chambers" refer generally to online clusters where individuals discuss

[127] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), pp 21, 45, citing Del Vicario, "The spreading of misinformation online", *Proceedings of the National Academy of Sciences* 113(3) (2016) 554-559; Jieun Shin and Kjerstin Thorson, "Partisan Selective Sharing: The Biased Diffusion of Fact-checking Messages on Social Media: Sharing Fact-Checking Messages on Social Media", *Journal of Communication* (February 2017); Jisun An et al, "Partisan Sharing: Facebook evidence and societal consequences", *COSN* (October 2014).

[128] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), pp 59, 67.

[129] Andrew Guess et al., "Selective Exposure to Misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign" (9 January 2018).

[130] Claire Wardle and Hossein Derakhshan, "Information Disorder: Toward an interdisciplinary framework for research and policy making," *Council of Europe report* (27 September 2017), 67, citing Saper Vedere, a Belgian start-up: <http://www.saper-vedere.eu/>.

similar views with like-minded people.[131] As pointed out by political data scientists Dr Hegelich and Mr Shahrezaye, this phenomenon is not surprising since social media platforms are designed for private exchanges between friends. The attraction of the like-minded to each other is hence welcomed on social media. Dr Hegelich and Mr Shahrezaye cautioned against "simple explanations like filter-bubbles or echo-chambers", as reality on social media was more complex.[132] While they disagree that there are "filter bubbles", as users are generally exposed to different ideological viewpoints, they also found that there is "polarisation caused by the uneven distribution of information on social media."[133] Other studies have similarly found evidence of a filtering effect in online "echo chambers".[134]

180. Online "echo chambers" appear to make it more difficult to combat the spread of online falsehoods. They tend to facilitate the spread of deliberate online falsehoods consistent with the beliefs of those in the "echo chamber", and hinder the effectiveness and spread of corrections. Studies cited in research by Dr Soon and Mr Goh found that people in online "echo chambers" tend to selectively filter and share information aligned with their ideological views. Such "echo chambers" are a primary driver of online misinformation.[135] Dr Hegelich and Mr Shahrezaye found that when confronted with ideologically different information, users in online clusters would attack the information instead.[136]

181. *Disruptions to the news ecosystem.* Fundamental disruptions to the news ecosystem have facilitated the spread of online falsehoods in three ways. *First*, the barriers for non-professional sources of news to enter the news ecosystem, regardless of their quality, have been lowered. This development has come with its pros and cons. Social media platforms have become the go-to sources for information globally. Individuals, who are not subject to checking mechanisms and editorial oversight, are able to gain popularity through promotion on social media. Some take advantage of the anonymity of the Internet to be reckless or

[131] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 44.
[132] Simon Hegelich and Morteza Shahrezaye, Appendix III: Written Representations, Paper No. 74, page B443.
[133] Simon Hegelich and Morteza Shahrezaye, Appendix III: Written Representations, Paper No. 74, page B443.
[134] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), pp 44-45; citing E Pariser, *The filter bubble: What the Internet is hiding from you*, Penguin UK (2011); Sunstein, *On rumors: How falsehoods spread, why we believe them, and what can be done*, Princeton University Press (2014); Del Vicario, "The spreading of misinformation online", *Proceedings of the National Academy of Sciences* 113(3) 554-559; Halberstam and Knight, "Homophily, group size and the diffusion of political information in social networks: Evidence from Twitter", *Journal of Public Economics* (2016) 143.
[135] Carol Soon and Shawn Goh, "What Lies Beneath the Truth", *Institute of Policy Studies* (30 June 2017), pp 44-45; citing Pariser, *The filter bubble: What the Internet is hiding from you*, Penguin UK (2011); Sunstein, *On rumors: How falsehoods spread, why we believe them, and what can be done*, Princeton University Press (2014); Del Vicario, "The spreading of misinformation online", *Proceedings of National Academy of Sciences* 113(3) 554-559; Halberstam and Knight, "Homophily, group size and the diffusion of political information in social networks: Evidence from Twitter", *Journal of Public Economics* (2016) 143.
[136] Simon Hegelich and Morteza Shahrezaye, "Disruptions to political opinion", *Konrad Adenauer Stiftung* (July 2017), p 6.

negligent with the "news" they report or share using their individual social media accounts. In that regard, a representative from SPH explained that "many players [had] come into the space and as a result… the sphere [was opened up] to a lot more content that could be false."[137]

182. *Second*, the unprecedented connectivity of the Internet has encouraged a public expectation for news to be issued in real time, although this may be at the expense of robust news verification. This makes it easier for rumours and conflicting accounts to spread and confuse. The representative from NGO Monitor gave the example of a terrorist attack. He explained that previously, the deadline for journalists to file their reports could be the next morning. Today, however, anyone with a camera and a social media account could put out an image and a narrative of the attack immediately, with limited verification, if any. The publication could then go viral. The authorities, who "do not have the luxury of making instant claims", would be at a disadvantage, as the publication would be circulating online while they investigated the facts.[138]

183. *Third*, the business model of newspapers has been disrupted. The advertising revenue they depend on is being channelled instead to social media platforms, which provide news aggregation and digital advertising. According to representors from the journalism sector, this has put a strain on the ability of newspapers to sustain themselves, and to carry out their missions of educating and informing the public.

184. *Transformation of political discourse.* Social media has transformed political communication, by making it emotional and convenient, rather than reasoned and considered. This was explained by political data scientists Dr Hegelich and Mr Shahrezaye, who observed that we are seeing an exceptional "digital revolution of the public and private sphere":[139]

a. Digital technology has blurred the line between the private and public spheres.

i. The amount of public information online has exploded. What is considered part of the public sphere is growing exponentially.

ii. Social media was supposed to be a channel of private communication. Instead, it has allowed many people to address the public directly.

iii. Decisions on what should be made public are increasingly being made by social media algorithms.

[137] Warren Fernandez, Appendix IV: Minutes of Evidence, page C497, para 4289.
[138] Gerald Steinberg, Appendix IV: Minutes of Evidence, page C556, para 4748.
[139] Simon Hegelich and Morteza Shahrezaye, Appendix III: Written Representations, Paper No. 74, page B44.

iv. The institutions that used to guard the line between the public and private spheres, such as the media, are losing influence.

v. What was previously confined to the private sphere is now entering the public sphere, even though what was not wrong in the private sphere may be wrong in the public sphere.

b. Further, social media was not designed for political communication. The use of social media to share political content has led to "an enormous misfit in design."

i. Social media was designed for convenient communication among private persons. Such private communication is guided by private affinity and emotions.

ii. Political discourse should not be a matter of convenience. Democratic political discourse should be the result of debates, which are often arduous, in order to find the right compromise among legitimate interests.

185. Social media is therefore said to have negatively impacted the quality of information that enters public discourse. At the same time, Dr Hegelich and Mr Shahrezaye were of the view that the importance of social media for political communication will only continue to grow.

(B) Disinformation Operations: Attacks on National Sovereignty and Security

186. Studying the use of online falsehoods by State actors in disinformation operations is important to fully understand the potential national security risks posed by the phenomenon of deliberate online falsehoods, and to shed light on the more sophisticated techniques and tools that could be used against Singapore. The Committee invited and received extensive evidence on this aspect of the problem.

187. As mentioned above at [26], the Committee is not in a position to draw any conclusions on whether any country is indeed responsible for the alleged actions or intentions attributed to them by others. It is also not within the Committee's remit to assess whether these alleged actions were conducted for geopolitical or other reasons. Statements set out below and in the annexes concerning the actions of any such country should be regarded as statements made by representors. These statements do not reflect the Committee's views.

188. Part I(B) addresses the following issues:

a. The use of disinformation operations as a military doctrine or tool;
b. Disinformation operations allegedly conducted by Russia; and
c. Disinformation operations allegedly conducted by an Asian country.

(1) The use of disinformation operations as a military doctrine or tool

a. The use and goal of disinformation operations as "non-kinetic" warfare

189. Evidence was given that online falsehoods systematically spread by foreign States – *i.e.* disinformation operations – have the potential of harming the national sovereignty and security of another State. Disinformation operations involve the spread of false or deliberately misleading information in another State to provoke or to push an agenda that is politically favourable to the aggressor State. In Czech expert Mr Janda's view, governments should treat such operations as a "national security threat".[140] This was echoed by national security experts from RSIS,[141] who have jointly argued in a paper that online falsehoods seeking to undermine the foundations of society should be recognised as a "national security issue".

190. *Established part of military arsenal*. Various representors were of the view that disinformation operations are now an established part of the military arsenal of some nation States, and have become just as important as conventional "kinetic" warfare. It is now clear that States *do* engage in disinformation operations, just as

[140] Jakub Janda, Appendix IV: Minutes of Evidence, page C168, para 1436.
[141] Muhammad Faizal bin Abdul Rahman et al, "Countering Fake News: A Survey of Recent Global Initiatives", *S. Rajaratnam School of International Studies, Policy Report* (March 2018), p 4.

they engage in warfare or diplomacy. According to national security expert Dr Shashi, these "non-kinetic" means of warfare may be deemed more important today in geopolitical conflicts, especially against a militarily superior enemy. While the use of disinformation is not new, Ukraine Media Crisis Centre (UCMC) recognised that it has been adapted very effectively for the age of Internet and social media. The information space is allegedly now the main battlefield in this conception of warfare.[142] In the view of Dr Michael Raska (Assistant Professor, RSIS), the use of social media today in conflicts is important, as it allows tailored information to be delivered to influence events in real time.

191. Various national security experts also highlighted the importance of understanding how disinformation operations fit into the use of "non-kinetic" warfare by aggressor States. It is alleged that aggressor States which engage in disinformation operations tend to conduct themselves as follows.

192. *First*, they adopt an "unrestricted approach" to warfare. Dr Gulizar Haciyakupoglu (Research Fellow, RSIS) observed that aggressor States do not separate between wartime and peacetime, and nothing (*i.e.* no military or non-military tool) is considered "off the table".[143] Various representors recognised that disinformation operations are persistent and permanent – *i.e.* one can *expect* deliberate online falsehoods to be spread by aggressor States even in the absence of an open conflict. In fact, during what would appear to be peacetime, it is said that disinformation operations can work on "slow burn" issues which can be equally if not more pernicious, being in a guise which makes them hard to identify. By the time a target State detects the presence of these operations, severe damage may already have been done.

193. *Second*, they are quick to leverage on new types of advanced military technologies, and engage in various forms of information and cyber warfare. According to Dr Shashi, aggressor States would have a comprehensive suite of tools to achieve their desired outcomes, with disinformation operations being just one of them. Some of these other tools include mobilising different segments of the population in a society, infiltrating local non-government organisations, and bribing or paying off politicians. Besides disinformation operations, "non-kinetic" warfare can come in the form of cyber-attacks like large-scale hacking operations (or Advanced Persistent Threats), malware attacks and Distributed Denial of Service attacks. These cyber-attacks may allow the aggressor States to collect information on the target State's citizens, to guide their online actions and increase their impact accordingly.

194. <u>*Fundamental goal*</u>. According to some representors, the fundamental goal of disinformation operations launched by some aggressor States is to undermine the social resilience of the target State, and in the process weaken the target State on

[142] John R. Haines, "Russia's Use of Disinformation in the Ukrainian Conflict", *Foreign Policy Research Institute* (17 February 2015), p 3.
[143] Gulizar Haciyakupoglu, Appendix III: Written Representations, Paper No. 56, page B311, para 1(ii).

various fronts – its values, culture, political system or institutions. By reducing the morale of the armed forces and civilian population of a target country, for example, the aggressor State reduces the need to deploy hard military power to achieve the same aims. According to Dr Berzins, a national security expert from Latvia, once a critical mass of civilians in the target country believe their country is a failed State that does not care for the interests and needs of its population, they may be tempted to believe that the loss of current statehood will bring better living conditions. This would usher the State into a zone of significant security vulnerability, achieved solely by "non-kinetic" means alone.

b. The attractiveness of disinformation operations to aggressor States

195. According to various representors, disinformation operations are attractive to aggressor States for the following reasons.

196. *First*, the costs and manpower needed for disinformation operations are disproportionately low. In Dr Raska's view, disinformation operations can have a disproportionately high impact on target societies, with the costs of disinformation operations being relatively low when compared to the amount a State would have to spend on conventional military options to achieve the same aims. This is because quite often, all it takes for such operations to be carried out successfully are: (a) a small number of people who are sympathetic to the aggressor State's cause; and (b) many computers and Internet connections, appropriate VPN masking, and fake phone numbers to create fake accounts.

197. *Second,* disinformation operations are highly effective, and reduce the need to deploy kinetic means of warfare. Dr Raska observed that disinformation operations employed strategically can create a pressuring situation in which the target State makes a decision leading to its own defeat. This would reduce the necessity for the aggressor State to deploy kinetic means of warfare at all, due to the "chaos within the [target State] or the visible collapse of the national will to resist the enemy".[144]

198. *Third,* disinformation operations carry a low risk of detection and allow aggressor States to disclaim responsibility. This is especially so if the agents carrying out the disinformation are not paid by the State, and appear simply as ordinary citizens of that State looking to advance the interests of their motherland. Disinformation operations can be so insidious, subtle and obscure that a target State may not know that it is under attack. An aggressor State may also mix falsehoods with real stories to gradually change opinions, making it difficult for the messages spread to be identified as being part of a disinformation operation. Such disinformation operations can be, in the words of Dr Alan Chong (Associate Professor, Centre for Multilateralism Studies, RSIS), "*theoretically...undetectable*".[145] Also, unlike

[144] Alan Chong, Appendix III: Written Representations, Paper No. No 91, page B906, para 17.
[145] Alan Chong, Appendix III: Written Representations, Paper No. No 91, page B906, para 17.

conventional military operations, it will not be easy to quickly identify the real aggressors, even if one realises that one is a victim of a disinformation operation.

199. Various national security experts were thus of the view that many States today (whether large or small) do engage in disinformation operations. Representors agreed that it is reasonable to assume that any large State would have already started developing its own capabilities to engage in disinformation operations, or would in fact already possess the capabilities to do so. Representors also acknowledged that disinformation operations can be attractive even to smaller States, or States without strong technological or conventional capabilities, as a form of asymmetric warfare against larger or more well-resourced States. As Mr Nimmo explained, even if a small group (or for that matter, State) does not have the capability to create its own botnet to engage in disinformation operations, it can always "*rent a botnet for the occasion*".[146] This is not just a theoretical possibility; there is at present a proliferation of commercial entities, who would offer their expertise on cyber and disinformation operations for sale.

(2) Disinformation operations allegedly conducted by Russia

200. The Committee received substantial evidence on disinformation operations allegedly conducted by Russia, and the consequent impact of these operations, on various States. The following paragraphs provide a summary of the evidence received, while the detailed evidence is set out in **Annex E**.

201. Before reviewing the evidence on operations allegedly conducted by Russia, it bears noting that the researchers and experts who appeared before the Committee updated the Committee that various States (both big and small, and including some Asian States as well) have been involved in disinformation operations. On a related note, the Committee received a confidential briefing by a security agency in Singapore, and was similarly informed at the briefing that Singapore has been the subject of foreign disinformation operations by *various* States. It was also recently reported that "Russia is no longer the only country to use social media's openness against the unwitting populations of Western democracies", and that "other autocracies are now following suit".[147] This section simply details the substantial evidence received in relation to alleged Russian disinformation operations, and does not exhaustively cover all the case studies of all relevant States.

202. According to various representors, disinformation operations are part of Russia's complex toolkit of instruments which are allegedly used to undermine the sovereignty and security of target States, especially in Eastern Europe. It has been claimed that these information operations are used on a perpetual basis by Russia,

[146] Ben Nimmo, Appendix IV: Minutes of Evidence, page C210, para 1779.
[147] Casey Michel, "It turns out Russia isn't the only country turning Facebook and Twitter against us", *The Washington Post* (23 August 2018).

to manipulate an adversary's perceptions, shape its decision-making process and strategic choices, while minimising the scale of kinetic force needed.

203. Representors also described the multiple platforms and actors Russia has allegedly employed in the conduct of its disinformation operations. These include the following:

a. State-sponsored media outlets (*e.g. Russia Today* and *Sputnik*): These media outlets allegedly broadcast disinformation in target States, using "emotionally engaging" propaganda and entertaining means to spread the Russian narrative, and exploit the ideals of freedom of information and expression to create mistrust amongst local populations towards their own mainstream media or established news networks.

b. Social media: Social media has allegedly been used as a cheap distribution channel or gateway for Russian-linked media outlets to disseminate falsehoods, through the use of "quirky" articles with catchy titles, and sensational or emotional content.

c. Bots and Trolls: Bots and trolls have also allegedly been used to artificially amplify pro-Kremlin messages on the Internet, attack Russia's opponents and drown out constructive debates online.

d. Local actors: According to some representors, local actors in some target States have been key to the success of the spread of Russian-based disinformation, through their circulation of content disseminated by Russian media outlets within their respective spheres of influence, whether knowingly or not. This allows the content disseminated by Russia to appear to originate from locals.

204. Evidence was also submitted to the Committee on the impact such disinformation operations have allegedly caused to countries like Ukraine, the Czech Republic, the US and France. As the detailed evidence can be found in **Annex E**, the Committee would merely highlight a few of the claimed experiences and impacts at paragraphs [205]-[206] below, which reveal the highly powerful and dangerous nature of State-sponsored disinformation operations.

205. <u>*Ukraine*</u>: Russian disinformation operations in Ukraine are said to have achieved considerable success, with Ukraine being a neighbouring State with a huge proportion of Russian-speaking people who identify as being ethnically Russian. The disinformation tactics allegedly used by Russia in Ukraine include targeting groups vulnerable to Russian influence, to support overarching and emotive narratives. Russian disinformation operations have allegedly fuelled existing tensions between different communities, discredited Ukraine's standing in other EU countries, and even resulted in the loss of territorial sovereignty and lives in Ukraine. It was claimed, for example, that Russian-linked fighters who took up

armed conflict in Crimea and other parts of Eastern Ukraine had been motivated to fight because of the false Russian television coverage of Ukrainian "atrocities" against Russian-speaking citizens.

206. *The US*: Russian disinformation operations were allegedly launched in the US to undermine public faith in the US democratic process (*i.e.* the 2016 US Presidential Election), denigrate and harm Hillary Clinton's electability, and sow discord and discontent in US society generally. One of the key strategies attributed to a Russian-linked group (known as the "Internet Research Agency" (IRA)) was to infiltrate US communities on social media by ingratiating themselves with genuine members of the community, and then using the approval of those members to take a stance on political or divisive issues. The actions of the IRA have allegedly inflamed social divides, undermined the US democratic processes, and even incited public protest. For example, IRA initiated the creation of opposing Facebook groups, which allegedly triggered an actual standoff on the streets between supporters and opponents of an Islamic centre in Texas. Commentators have pointed out that the US was vulnerable to Russian disinformation operations because they were complacent and unprepared, and that despite all that has happened, the US is still struggling to find a coherent and effective response, due to its domestic politics and legal constraints in imposing effective countermeasures.

(3) Disinformation operations allegedly conducted by an Asian country

207. Evidence was also given as to how disinformation operations have been conducted in Asia, allegedly by an Asian country.

(C) Singapore's Context

208. Representors gave evidence on the nature of the phenomenon of deliberate online falsehoods specifically in relation to Singapore's context. Three key observations emerged: (i) foreign disinformation has likely occurred and can be expected to occur in Singapore; (ii) our societal conditions are fertile ground for insidious "slow drip" falsehoods which can cause long-term damage in Singapore; and (iii) our region's tensions and circumstances are a source of vulnerability.

(1) Foreign disinformation in Singapore

209. The evidence showed that disinformation operations have been conducted by various States. There are also a range of State and non-State actors who are engaging in disinformation operations. This evidence contradicted the views of a few representors, who asserted that the threat of foreign disinformation was posed only by Russia and that Singapore was not at risk because Singapore was not a threat to Russia. These representors did not explain the reasons for their assertions, and did not claim to have particular expertise on hostile disinformation operations. To the contrary, as mentioned earlier, researchers and experts, who had studied the field of hostile disinformation operations, had testified that the threat of disinformation is posed not only by one country, and that Singapore has been and can expect to be subject to foreign disinformation operations. The Committee was also informed at the confidential briefing which it received from a security agency that Singapore has indeed been the subject of foreign, State-sponsored disinformation operations. The evidence received by the Committee is set out below.

210. *Foreign State-linked disinformation has occurred in Singapore*. Dr Gulizar Haciyakupoglu from RSIS described some of the indicators of information warfare being carried out against Singapore. This included a State using news articles and social media to influence the minds of segments of the Singapore population, and to legitimise the State's actions in the international sphere.

211. As mentioned at [193] above, disinformation operations and cyber-attacks form a comprehensive set of tools which aggressor State and non-State actors rely on to wage a form of "non-kinetic" warfare against target States. In this regard, the Committee notes that there have also been a number of cyber-attacks against Singapore, including attacks on sensitive ministries and critical institutions, in the recent past.

a. In its annual report, the Cyber Security Agency of Singapore (CSA) reported that in 2017, Government agencies experienced a range of cyber threats, including system intrusions and spoofed websites.[148] According to American technology company F5 Networks and its data partner

[148] "Singapore Cyber Landscape 2017", *Cyber Security Agency of Singapore*, pp 10-11.

Loryka, Singapore was also the top cyber-attack target around the world during the Trump-Kim summit, with the country experiencing close to 40,000 attacks during the 12 June 2018 meeting. F5 Networks reported that during this period, 88% of attacks on Singapore were launched from a particular foreign State. Additionally, 97% of all attacks attributed to this foreign State during the same period were specifically targeted at Singapore. It remains, however, unclear whether the attacks were sponsored by that foreign State.[149]

b. It was revealed in July 2018 that the databases of SingHealth, the largest group of healthcare institutions in Singapore, had been hacked.[150] The personal particulars (such as names, NRIC numbers, addresses, gender and race) of 1.5 million patients, including that of several ministers, were stolen.[151] Initial investigations by the CSA and the Integrated Health Information System confirmed that this was a "deliberate, targeted and well-planned cyberattack" and "not the work of casual hackers or criminal gangs".[152] Further detailed analysis by the CSA determined the attack to be the work of an Advanced Persistent Threat group, which refers to a class of sophisticated cyber attackers, typically State-linked, who conduct extended, carefully-planned cyber campaigns, to steal information or disrupt operations.[153] As mentioned earlier, in the aftermath of the cyberattack, SMSes falsely claiming that patients' phone numbers, financial details and medical records had been accessed began circulating, as malicious actors sought to exploit the vulnerable situation with false information.

212. In a similar vein, Dr Shashi cautioned that it would be a mistake to assume that foreign State-led disinformation was not already happening here. National security experts Dr Shashi and Dr Cheong stressed that disinformation campaigns were usually a "long game", where infiltration and influence were covert and cumulative.

213. *Foreign disinformation can be expected to occur in Singapore*. Various representors highlighted several reasons why Singapore should expect to be a target of disinformation operations.

[149] Cynthia Choo, "Singapore top cyber attack target during Trump-Kim talks: Report", *Today* (18 June 2018).
[150] On 24 July 2018, a Committee of Inquiry was convened to, amongst other things, inquire into the events and contributing factors leading to the cybersecurity attack on SingHealth. The Committee of Inquiry is expected to submit a report on its findings and recommendations by 31 December 2018. See: Ministry of Communications and Information Press Release, "Appointment of a Committee of Inquiry into the SingHealth cybersecurity attack on or around 27 June 2018" (24 July 2018). As at the date of this Report, the Committee of Inquiry had yet to submit its report.
[151] "SingHealth cyber attack: How it unfolded", *The Straits Times* (20 July 2018).
[152] Ministry of Health Press Release, "SingHealth's IT System Target of Cyberattack" (20 July 2018), para 3.
[153] Ministry of Communications and Information, "Statement by Mr S Iswaran, Minister-in-Charge of Cybersecurity, on the cyber-attack on SingHealth's IT system, during Parliamentary Sitting on 6 August 2018", para 6.

214. *First,* information warfare against Singapore is a more attractive strategy than conventional warfare. National security expert Dr Raska submitted that considering Singapore's conventional military strength, foreign States who cannot challenge Singapore through conventional warfare will engage in subtle information campaigns that target the friction points in Singapore society, weakening Singapore and undermining Singapore's will to defend itself. This form of asymmetric warfare may offset a foreign State's military inferiority and achieve political aims similar to conventional warfare. Dr Shashi also observed that the key in future conflicts would be to employ asymmetric attacks on all elements of national power as a means to deter, intimidate or defeat a militarily-superior enemy. He pointed out how militaries are increasingly viewing non-kinetic actions as being just as important (or even more important) than conventional warfare.

215. *Second,* the means and tools for disinformation campaigns are allegedly readily available in our region, and can easily be turned against Singapore. For example, some national security experts pointed out that cyber armies which have been deployed to aid sectarian or political agendas exist in several of our neighbouring countries, which can easily be repurposed and deployed against Singapore.

216. Evidence was given on how "data-driven political consultants", who use sophisticated targeting techniques to influence public opinion, appear to have been engaged by political parties and politicians in Malaysia. According to Dr Shashi, this will allow techniques to be tried for a local environment, and will build up the expertise and capabilities of such techniques in the region. Dr Shashi agreed that this has the potential to be leaked to other entities and individuals in the region, who may have their own reasons to attack Singapore.

217. *Third,* our digital connectedness allows foreign actors easy reach to wide segments of Singapore's population. Evidence of how the Internet has made it easier for foreign agents to infiltrate and impersonate locals has been set out above. Singapore has been described as a "hyper-connected" community where people rely heavily on technological platforms to communicate, obtain and share information. A Nielsen survey indicated that more Singaporeans access their news over the Internet and social media, compared to those who access their news through hardcopy newspapers.[154] A report in the Business Times published on 24 January 2017 stated that 70% of Singaporeans are active social media users on mobile devices, more than double the global average of 34%, and that more than three in four Singaporeans use social media.[155]

218. According to Mr Nugroho, the wide use of English in Singapore could also lower the barrier posed by language differences, since English is widely spoken around the world. He contrasted Singapore with Indonesia, where people generally speak

[154] Ang Peng Hwa, Appendix III: Written Representations, Paper No. 143, page 1259, para 1.3.
[155] Singapore Press Club and Singapore Corporate Counsel Association, Appendix III: Written Representations, Paper No. 155, page B1364, para 2.4.

Bahasa Indonesia, where foreign disinformation operators would be forced to learn the language if they wanted to launch disinformation campaigns there.

219. Crises (such as terrorist attacks) are flashpoints that could be exploited by disinformation agents. Dr Mathews warned that during crises, suspicion and anxiety are heightened, and any misinformation spread online will almost certainly have an effect on Singaporeans' minds, affecting trust among different communities. This creates potential for foreign nefarious elements to use deliberate online falsehoods to de-stabilise Singapore in moments of crisis.

(2) Real risks of "slow drip" falsehoods causing long-term damage to society

220. Singapore's diverse society provides fertile ground for insidious "slow drip" falsehoods to cause longer-term damage to society that may not always be visible, until too late. The dangers of "slow drip" falsehoods were elaborated on by various expert representors (see [105] to [108] above). The Committee also received first-hand and expert evidence on how such falsehoods have manifested in Singapore.

221. *"Slow drip" falsehoods.* Expert representors repeatedly warned of the insidious "slow drip" effect of falsehoods on our society. In the words of Dr Mathews, a researcher who has examined issues related to social cohesion for over a decade, it is "in our everyday lives where deliberate online falsehoods could harm our social cohesion."[156] He explained that "low-level" falsehoods could raise tensions little by little; emotions may not be high initially, but falsehoods could make them stronger.

222. Dr Mathews explained that such falsehoods may include "reports that intentionally feature misinformation about particular ethnic, religious or immigrant groups and their loyalty to Singapore, their potential to commit anti-social acts or crimes, their lack of contribution to society, their overuse of state resources, or highlight and speculate about aspects of their culture which may not be well understood but deemed as at odds with majority culture."[157]

223. One example was the falsehood spread by news website The Real Singapore about how a complaint by a Filipino family resulted in commotion between Hindu participants and the police during a Thaipusam procession in 2015. The story quickly gained traction among netizens, who did not question its veracity. It led to xenophobic comments online. In Dr Mathews' view, the story would have shaped the opinions of some Singaporeans towards immigrants, Hindus, and an important religious festival. He warned that such falsehoods could cumulatively have a corrosive effect on social cohesion over time.

[156] Mathew Mathews, Appendix III: Written Representations, Paper No. 100, page B969.
[157] Mathew Mathews, Appendix III: Written Representations, Paper No. 100, page B969.

224. The experience of Mr Hetamsaria, described earlier above, is also apt. The falsehood that Mr Hetamsaria was a new citizen disappointed with Singapore stirred up xenophobic and anti-immigrant sentiment in Singapore.

225. Mr Hazrul Jamari provided evidence from the perspective of a member of the Singapore Malay community. He described how the spreading of falsehoods of a communal or religious nature via WhatsApp and Facebook was common. In his view, such falsehoods tread on very sensitive territory. Simple falsehoods about Halal stores selling pork could create a sense of panic within the community. ISIS propaganda online had stirred tensions between local Sunnis and Shias.

226. A few representors did not think that the problem was serious because Singapore has yet to experience any "significant harm". They noted local incidents such as the false story of the 2015 Thaipusam procession, but did not consider these types of incidents significant because, amongst other reasons, no visible impact arose from these incidents. However, as Dr Mathews and some of the other representors mentioned above have testified, while some online falsehoods have a visible impact, others can have a hidden and insidious impact over time, which is no less significant.

227. *Singapore's social conditions.* Singapore's diversity means that extra care must be taken to protect our social cohesion. Dr Shashi observed how Singapore can be a "sandbox for subversion", given our small size and highly-wired population. He explained that as Singapore was polyglot, multiracial, and data rich, an aggressor could "peel off" a particular group and set it against other groups or public institutions.[158]

228. Representatives from religious organisations shared how divides and fault lines were very real in Singapore. The representative from the Singapore Buddhist Federation spoke of how people motivated by religious zeal or bigotry have spread falsehoods about the Buddhist faith or Buddhism. Representatives from the Roman Catholic Archdiocese of Singapore (RCC) and the National Council of Churches in Singapore (NCCS) observed that tensions could exist among people of the same faith as well, including along ethnic lines.

229. The representative from NCCS expressed that these fault lines could not be eradicated, and that Singapore's high level of cohesion did not render us immune from the eruption of conflict. A representor who was previously from the Inter Religious Organisation cautioned that the effect of online falsehoods of a communal nature could be exacerbated among mixed and densely-populated housing estates.

230. Dr Mathews gave evidence from his experience in examining issues relating to social cohesion and exploring Singapore's fault lines for over a decade. His

[158] Shashi Jayakumar, Appendix III: Written Representations, Paper No. 59, page B334.

research showed that Singapore is still not a race-blind society, and our differences still matter.

a. A 2016 study conducted by Channel NewsAsia (CNA) and the Institute of Policy Studies in 2016 on 2000 Singapore citizens[159] found that stereotypes and prejudices are held by a sizeable proportion of the Singapore population. Almost half of the respondents agreed that people from some races compared to others would be more violent, not friendly, and more likely to get into trouble. About half of the respondents also reported that most Singaporeans were mildly racist. More than half of the respondents perceived new migrants as racist.

b. Dr Mathews' past research has shown that some Singaporeans perceive that discrimination and prejudice still exist, especially when it comes to getting jobs and top positions.

c. In another study conducted with more than 2000 Singapore citizens and permanent residents in April and May 2017, the results showed that considerable numbers of Singaporeans find it hard to trust people of other races. Around 60% of Chinese respondents thought that less than half of Malays or Indians would return their wallet.

231. Any source of difference may be exploited, not just racial and religious differences. Divides along class lines were flagged by lawyer Mr Zhulkarnain Abdul Rahim. Ideological differences may also be susceptible to the effect of falsehoods. Some of the largest Facebook advocacy groups in Singapore may be sites where falsehoods are gaining traction, according to preliminary empirical research by a group of Singaporean representors. In these online groups, they found news from unreliable sources and hate propaganda to be prevalent. Also pertinent is evidence of how international NGOs may use misleading accounts of the facts to interfere in Singapore's politics.

232. In August 2018, the RCC had to issue an official notice to de-bunk WhatsApp messages which suggested that the Archbishop was unhappy with the Government. The message circulated on WhatsApp took an extract of Archbishop William Goh's "Scripture Reflections" out of context, to achieve this purpose.[160] This shows that deliberate online falsehoods can be used not only to drive a wedge between one community against another in Singapore, but between a specific community and the Government as well.

[159] Mathew Mathews, "Key Findings from the Channel NewsAsia – Institute of Policy Studies Survey on Race Relations, *Institute of Policy Studies* (2016), p 3.

(3) Vulnerability due to regional circumstances

233. The Committee received evidence on how our regional context can contribute to Singapore's vulnerability to harmful online falsehoods.

234. *First*, the sources and drivers for information warfare that could affect Singapore are deeply embedded in our regional security conflicts, with the conflicts in the region being reflected online as well.

235. *Second*, societal fault lines run across national borders. Dr Liew Kai Khiun (Assistant Professor, Wee Kim Wee School of Communication and Information, Nanyang Technological University) cited an example relating to the crisis faced by Muslims in the Rakhine State, Myanmar. When local media CNA and *The Straits Times* reported on the crisis, comments were posted on their social media pages refuting their reports. These denials appeared to come from Myanmar-based user accounts, and were accompanied by comments with Islamophobic overtones, triggering backlash from accounts that appeared to belong to Singaporean Muslim users.

236. The spill-over of tensions from the region into Singapore is a cause for concern. According to Dr George's study of hate propaganda, hate groups in the region and around the world "are far more formidable than anything we have needed to deal with."[161] Similarly, Singapore could be impacted by the religious and racial policies of our neighbouring countries. Dr George cautioned that it would be reckless to assume otherwise.

237. *Third*, as mentioned above, the resources and tools used for information warfare are available and accessible elsewhere in the region, and can be effectively used by actors familiar with the local and regional context against Singapore.

(4) Other Matters

238. Before moving on, the Committee should make reference to one representor, Dr Thum Ping Tjin. Dr Thum made some claims regarding falsehoods in Singapore, the details of which are set out in the Addendum. The Committee has given no weight to Dr Thum's views. Based on his conduct in relation to the Committee, the Committee does not find Dr Thum to be a credible representor. *First*, he misrepresented his academic credentials in his evidence, to suggest that he held more distinguished roles at Oxford University (*e.g.* a "visiting professorship") than the unpaid positions he held, and visiting scholar arrangements he obtained in return for paying a fee. His claim that his repeated misrepresentations were unintentional (*e.g.* a "typographical oversight") is not believable. He has clearly lied. *Second*, when questioned about his key allegation that Operation Coldstore had no national security basis, he admitted that he had not read or had chosen not

[161] Cherian George, Appendix IV: Minutes of Evidence, page C704, para 5876.

to give any weight to accounts by senior cadres of the Communist Party of Malaya that he acknowledged contradicted this thesis; he also admitted he had not in his publications explained why he chose to disregard them. *Third*, he failed to follow up with documents to substantiate his claim that he had indirectly dealt with contradictory evidence in his publications, a claim significant to his credibility as a historian, despite submitting an additional follow-up representation. The facts of what transpired, including those that led the Committee to draw its conclusions, are set out in the Addendum.

(D) Conclusions on the Nature of Deliberate Online Falsehoods

239. The Committee concludes that deliberate online falsehoods are a real and serious problem for the world, and Singapore. It is a problem that is more potent than before, due to technological advances and social changes in the Internet era. The advent of the Internet, and social media in particular, came with much promise. At the same time, they have also come with problems. There is overwhelming evidence of one of the problems, namely, deliberate online falsehoods. Singapore is not immune, and must take action to combat it.

240. There is clear evidence that deliberate online falsehoods have caused serious harm in many countries. They can influence the emotions, beliefs and actions of many people, as shown by psychological research set out above. In summary:

a. Falsehoods influence memory, reasoning and decision-making. Falsehoods can also have a strong influence on beliefs, when they appeal to emotions, particularly our negative emotions.

b. Falsehoods affect emotions. They can make people feel more concerned or threatened by something than the evidence warrants. They can arouse fear, and make people react with anger. Simple falsehoods can activate "tribal identities" in a way that is difficult to fight.

c. Falsehoods affect trust and sow doubt. Falsehoods erode trust in public institutions, and official information. They can decrease people's desire to engage in politics, and prevent people from believing in valid information. They can have this effect, and thus sow doubt, even if people do not believe in the falsehood itself.

d. Exposure over time to falsehoods that promote or attack a particular point of view can gradually change people's views.

e. The above applies to many of diverse backgrounds, including the well-educated and literate.

241. The examples in other countries of serious harm caused by online falsehoods are quite clear. These have been set out in the sections above and **Annexes A, B, C and D**. In essence:

a. Online falsehoods can lead to violence and the loss of lives. By provoking hate, online anti-Muslim falsehoods led a British man to drive a van into a mosque in the UK, leaving one dead and others injured. By arousing anger and reinforcing confirmation bias, the conspiracy theory in the US about a paedophilia ring allegedly run by political figures connected to the Democratic Party went viral in online chat rooms, and led to demonstrations and an armed man firing into a pizza restaurant.

b. Online falsehoods can also have longer-term and more fundamental effects on society and public discourse:

i. Falsehoods impede rational and reasoned political debate, by stoking negative emotions, such as anger and hate. They encourage polarisation and the entrenching of misinformed ideological beliefs. They seed negativity and anger among the middle ground, and polarise society by leveraging on or exploiting existing cleavages in society, making it difficult for rational discourse to take place. They can ultimately influence elections as well, as some studies indicate.

ii. Falsehoods also affect policy-making. Surveys in the Czech Republic and the Netherlands showed how falsehoods about the Ukrainian government could influence public opinion against Ukraine. This was said to have impeded their government policies on cooperation with Ukraine. Representors from Europe shared how the pollution of discourse by falsehoods has made it difficult for policy-making in Europe in relation to issues such as migration.

iii. The problem has worsened in the Internet era. Social media has blurred the long-standing divide between public discourse and private speech, and encouraged political discourse online to be convenient and emotional, when it should instead be reasoned and considered. This has boosted the ability of falsehoods to proliferate and influence. Online "echo chambers" allow falsehoods to gather strength. Social media and online targeted advertising tools can be manipulated to tailor and directly send political falsehoods to people with certain ideological views, exploiting their confirmation bias.

242. Online falsehoods are therefore antithetical to the philosophical argument concerning the "marketplace of ideas." Allowing the free flow of speech in the public marketplace is intended to result in the truth prevailing. This has been elaborated on further below at [426]-[429]. In summary:

a. For the "marketplace of ideas" to properly function, certain conditions are necessary.

b. One condition is that people will rationally debate and engage with different ideas. However, online falsehoods impede this, by arousing anger, fear and other negative emotions, and by exploiting our cognitive biases.

c. Another condition is that everyone has equal access to the range of ideas available in the "marketplace". However, online falsehoods crowd out reliable facts and news, especially when the falsehoods appeal to emotions. As shown by psychological research, they can make people stop believing in accurate information.

d. Online falsehoods can therefore interfere in the proper functioning of the "marketplace of ideas". They have a negative impact on democratic contestation, by resulting in people being misinformed, rather than informed.

243. Among the most egregious of consequences that online falsehoods can have is their threat to the national sovereignty and security of a country. Evidence was led on how disinformation campaigns by foreign actors can and have led to greater friction, distrust and anger in the target society, political leaders being influenced, elections being undermined, public protests taking place, and even the loss of territorial sovereignty.

244. The digital age has made the spread of falsehoods a more potent problem than ever before, as shown by the growing scale of the problem, which societies have struggled to cope with. In summary:

a. The Internet has allowed people with agendas to prolifically spread falsehoods online, including through the deliberate use of false amplifiers like botnets. Falsehoods are circulating on the Internet in large quantities and copies. They exploit the "illusory truth" effect, which is the human cognitive tendency to believe a falsehood the more one sees it. As the Center for European Policy Analysis put it, the age of information is turning into the age of disinformation.[162]

b. The overload of information online also allows online falsehoods to take advantage of cognitive limitations on people's ability to discern truth from falsity.

c. The digital revolution has lowered the barriers to entry for producers of falsehoods. The range of possible perpetrators of impactful online falsehoods now extends beyond well-resourced State actors, to include ordinary people.

d. Technological advances have made disinformation tactics easier and more effective. Foreign disinformation experts spoke of how the Internet has brought foreign propagandists a growing audience. Unlike before, foreign actors are able to use online tactics to easily infiltrate local

[162] Edward Lucas and Peter Pomeranzev, "Winning the Information War", *Center for European Policy Analysis*, (August 2016) p 2.

communities through impersonation, and directly send their own messages into the target country on a large scale.

e. The low cost and high impact of online disinformation campaigns have made the adversaries of a country, whether other States or non-State actors, more formidable.

f. These developments are taking place within a social landscape that has changed rapidly due to the Internet. Today, Internet users are accustomed to receiving information from purported news websites set up by anonymous people or untrained citizen journalists. They are also accustomed to posting such content on social media even if it is false or likely to be taken out of context. The influence of the traditional media, which traditionally safeguarded the accuracy of information in the public sphere, has been diluted. News breaks in real-time, making traditional news verification more challenging. Social media has made it easy for the like-minded to connect with each other regardless of physical location, and feed off each other's ideological biases. Information on social media is now, to some degree, filtered to us based on our personal preferences. Social media is also transforming political communication, as explained at [184] above.

245. 14 representors submitted that the problem was not so serious as to warrant new measures by Singapore. They held the view that the problems posed by deliberate online falsehoods today are no greater than or different from the age-old problems posed by falsehoods before the Internet era or, as mentioned above, that the phenomenon of deliberate online falsehoods poses no credible threat to Singapore. The Committee's views on these submissions are as follows:

a. None of these representors considered the range of serious consequences that online falsehoods could have, and have had. During the hearing, after being shown the evidence of expert representors, a number of the representors who had raised questions about the nature of the problem agreed with one or more of the following:

i. The reach and speed of the Internet and social media have escalated the potential impact of disinformation.
ii. One purpose of disinformation is to slowly undermine trust in an institution.
iii. Disinformation poisons public debate and is a threat to democracy.
iv. There is no guarantee that even effective counter-campaigns can defeat the high volume flow of malicious communications.
v. Falsehoods can be used to incite unrest and, when that happens, some form of response is necessary.

b. They also did not provide evidence to show that the phenomenon of deliberate online falsehoods is no different from that before the Internet age. Instead, they generally relied on the argument that the Green Paper had not shown why the phenomenon warranted new measures. However, the purpose of the Green Paper was to serve as a reference point for the Committee's further work, and invite viewpoints and evidence on various points, rather than set out extensive evidence. Since then, the evidence presented to the Committee has made clear that the problem is real and requires specific measures.

c. None of them gave contrary evidence on the serious influence that falsehoods tend to have on people's emotions, beliefs and actions.

d. On the issue of whether falsehoods can influence people's voting behaviour:

    i. As explained above at [137], studies cited on the specific issue of impact of falsehoods on voting behaviour did not make findings on this issue. Studies on the matter are, at this point, not conclusive.

    ii. Voting behaviour is only a narrow issue. Falsehoods have a range of other effects on people, which were not contested. Beyond voting behaviour, there are many other ways in which falsehoods impact people and societies, in highly damaging ways. That was not contradicted.

    iii. Even in the narrow context of voting behaviour, none of the representors disputed that serious attempts have been made to influence votes, and these attempts will continue.

    iv. There are likely to be continued attempts to influence votes, including the use of insidious "drip feed" falsehoods. There is no reason to believe that these attempts will not succeed, especially in view of the evidence of how falsehoods influence people.

e. The evidence received by the Committee (see [209]-[219] above) clearly illustrates that Singapore is not immune to online falsehoods. Singapore has been targeted by foreign actors in the past, and can be expected to be targeted in the future.

f. None of the 14 representors gave concrete evidence why Singapore would be immune to online falsehoods.

    i. There was no evidence to show that Singapore is less vulnerable because of our context and social conditions.

ii. The argument that Singapore is immune runs counter to the evidence from studies on how falsehoods can influence all people, regardless of background. Notably, in an earlier survey, around two in three Singaporeans said they could not always recognise news as fake,[163] and one in four of them admitted to sharing news they later found was fake.[164] Further, the problem of online falsehoods has manifested in countries with diverse contexts, including countries like the US, UK, Germany and France, which have high levels of education and literacy.

iii. The argument that the high level of trust in public institutions and the lack of polarised politics in Singapore make Singapore invulnerable, is difficult to accept. Falsehoods may have the very objective of eroding trust in public institutions and polarising society. The evidence shows that creative and sophisticated disinformation tactics are already being employed, with success. As Mr Janda from the Czech Republic, and Head of the Kremlin Watch Program, cautioned, even as fault lines in society are being bridged, disinformation agents will keep adjusting their strategies to capitalise on other vulnerabilities, whatever they may be.

iv. The view that the problem is confined to online disinformation campaigns by certain foreign countries, including but not limited to Russia, and Singapore is hence not affected, is also not tenable.

    1. Foreign disinformation experts were clear that there were State and non-State actors who were conducting such campaigns, using a range of digital tools and techniques, and others would learn as well.

    2. Evidence also indicated that Singapore has been a target of foreign actors, who have used online falsehoods to influence Singaporeans. The Committee is satisfied that Singapore has been the target of foreign actors with capabilities similar to what Russia is alleged to have, *i.e.* the ability and propensity in engaging in "non-kinetic" warfare. Besides disinformation operations, it is also publicly known that Singapore has increasingly been a target of cyber-attacks, including against government agencies.

    3. Capabilities such as cyber armies, troll farms and data-driven targeting exist in our region and in other countries. They can easily be turned against Singapore.

[163] Media Release – Findings of Poll on Attitudes towards Fake News, *REACH* (26 March 2018), p 4.
[164] Kelly Ng, "Laws tackling fake news to be introduced next year: Shanmugam", *Today* (19 June 2017).

g. In the face of the overwhelming evidence received by the Committee – our geopolitical situation, the incidents which have occurred in Singapore and the experience of other countries as recounted by experts – it is clear that the problem is serious enough to warrant a robust response and new measures. Similar to how Singapore has tackled other challenges such as cybersecurity and terrorism, precautionary steps should be taken to tackle deliberate online falsehoods.

246. Evidence was led on what some of Singapore's vulnerabilities were. This included surveys showing that racial divides exist, and differences do matter, as they always will to an extent in any country. Online disinformation campaigns often seek to exploit such divides.

247. Representors gave evidence of online falsehoods targeting trust in public institutions and social divides in Singapore. In the case of website *The Real Singapore's* false story about a Filipino family's complaint about noise from a Thaipusam procession, it was argued that the prosecution of the website's co-founders was an adequate response. However, that took place over a year after the story was published. It did not stop the story from spreading and gaining xenophobic responses online, which amplified what were otherwise views from the margins. Over time, such incidents can change public attitudes towards social harmony and tolerance of other communities. The "drip feed" effect of falsehoods should not be under-estimated.

248. Importantly, Singapore's approach has been to act early and ensure that we are prepared. On racial and religious issues for example, Singapore's multi-pronged approach involves building trust with leaders of religious communities, who in turn have the trust of their communities, implementing strong social policies, and having in place legislation such as the Maintenance of Religious Harmony Act, to nip in the bud inflammatory conduct. A similar approach should be taken in relation to falsehoods which are spread online, to prevent the erosion of the trust which has been painstakingly built up between the different communities in Singapore. This is especially important in light of the evidence received by the Committee on how deliberate online falsehoods have sowed discord and incited public protests on the streets between different communities in societies.

249. The experience of the US offers an apt lesson. The US was reportedly slow to appreciate the threat of foreign State information warfare, despite early warning signs, due to a "misguided belief" in the resilience of American society and its democratic institutions.[165]

250. Deliberate online falsehoods are therefore a problem that Singapore should take action against. The key question is what should be done.

[165] Adam Entous et al, "Kremlin trolls burned across the Internet as Washington debated options", *The Washington Post* (25 December 2017).

(II) RESPONDING TO THE PHENOMENON

251. The purpose of Part I of the Report is to help readers *understand the phenomenon* of deliberate online falsehoods, in terms of the real and serious problems this phenomenon has caused, and will cause, to the world generally, and to Singapore as well. Part I summarises the evidence received by the Committee on how different actors spread online falsehoods for different objectives; how digital technologies today enable them to do so easily; how such falsehoods can threaten our national security, public institutions, and also individuals and businesses; and why online falsehoods are so difficult to combat today.

252. Part II of the Report sets out the various countermeasures proposed by representors on how we should *respond to the phenomenon*, and the Committee's respective observations and recommendations. A common view taken by many representors, which the Committee agrees with, is that there is no one silver bullet to combat this complex problem, and a multi-pronged approach is necessary. This multi-pronged approach would involve both near-term and long-term efforts, and the involvement of multiple stakeholders – including media organisations, technology companies, community leaders and volunteers, and the Government.

(A) Desired Outcomes

253. It is critical to first understand what our collective, desired outcomes are. These desired outcomes would then help to shape, and guide, our understanding and evaluation of the countermeasures recommended, or those which are eventually implemented. A comprehensive understanding of the harms posed by deliberate online falsehoods, and careful consideration of the countermeasures proposed by representors, have allowed the Committee to crystallise what these desired outcomes are:

a. *A population that is well-informed and digitally literate*. The outcome envisioned here is a population which has easy access to accurate facts, but also equipped to assess information critically in the digital age. A well-informed and digitally literate population is one where citizens are empowered to make good and accurate decisions, on both the individual and societal levels, benefitting not just themselves, but those around them as well. This requires trusted and credible sources or mechanisms to be in place, which people can reliably rely on for the true facts; and for every individual to also be sufficiently discerning and sophisticated to tell truth from falsity online.

b. *A society that is cohesive and resilient.* The outcome envisioned here is that of a well-functioning society where there is a high level of trust and respect between different communities, and between the people and public institutions as well. It is a society where harmony and tolerance prevail, not just in good times, but even in challenging moments. Such a

society is one which will stand resilient in the face of efforts seeking to sow discord between different groups. It is also a society where public institutions are able to govern and serve the public effectively, because of the high trust reposed in them by the people.

c. *An information ecosystem which values and protects the truth above all*. The outcome envisioned here is an overall confidence amongst people that information transacted online can be trusted, and healthy, fruitful discussions can flourish. It entails building and maintaining effective mechanisms, to regularly ensure our information ecosystem is not polluted by deliberate online falsehoods. This means having in place measures that can effectively and authoritatively take to task both the malicious agents and the falsehoods they spread. The ultimate goal of these measures is *not* censorship, but the exact opposite – to ensure our freedom of speech can be meaningfully exercised, in a properly-functioning "marketplace of ideas" that is not drowned out by fake actors or false content.

d. *A nation with our sovereignty and freedom safeguarded*. The outcome envisioned here is that our sovereign right to debate issues rationally and passionately, and make vital decisions for and by ourselves, will not be interfered with by malicious foreign agents seeking to advance their own political agendas. It means ensuring that our democratic processes and public institutions are protected from the negative influence of State-led disinformation operations.

(B) Proposed Countermeasures

254. The Committee received evidence on a suite of possible countermeasures proposed by representors, covering different dimensions of the problem, and involving a range of different stakeholders. They comprise measures that will allow swift action to be taken in the short-term, and those that will take effect over the longer-term. Many representors were of the view that there was no one silver bullet in combatting deliberate online falsehoods. Each component of the suite of countermeasures is necessary, and plays an important role, to achieve the desired outcomes outlined above. As Dr Soon and Mr Goh put it, "a combination of [near-term and long-term] measures will mitigate the shortcomings of each, thus providing a holistic solution to the problem at hand".[166]

255. In the Committee's view, the suite of countermeasures proposed by representors can be categorised under one of the following components:
    a. Nurture an informed public;
    b. Reinforce social cohesion and trust;
    c. Promote fact-checking;

[166] Carol Soon and Shawn Goh, Appendix III: Written Representations, Paper No. 62, page B366, para 22.

d. Disrupt online falsehoods; and
e. Deal with threats to national security and sovereignty.

256. The report will discuss each component in turn below. In each section, the Committee will first outline the rationale and context for the countermeasures within each component; second, summarise the views and recommendations put forth by representors in relation to the countermeasures proposed; and third, set out the Committee's observations and recommendations on the proposed countermeasures.

(1) Nurture an Informed Public

257. Nurturing an informed public is recognised as a critical, long-term safeguard against the threats posed by deliberate online falsehoods. Ultimately, deliberate online falsehoods are only as effective as the negative impact or response they engender from their recipients. As a representor described, an informed citizenry that is able to effectively discriminate between what is factual or not, and knows which are the reliable sources of accurate and unbiased reporting they can trust, is "the first line of defence against disinformation and misinformation".[167] In general, representors proposed two key approaches in nurturing an informed public: public education and quality journalism.

a. Public education

(i) Rationale and context

258. Public education is key to strengthening the resilience of our citizenry against deliberate online falsehoods. In this context, public education refers to the inculcation of media and digital literacy, and also critical thinking skills. The goal is to build up the immunity of our citizenry against deliberate online falsehoods by equipping them with the knowledge and skills to discern truth from falsehood, effectively interrogate information sources and understand how and why online falsehoods are spread in the digital age. Specifically on news literacy, the 2018 Reuters Institute Digital News Report highlighted that improving people's knowledge about how the news is made – who makes it, how it is selected and how it is financed – would enable people to better separate fact from fiction.[168] Public education also seeks to teach people to be responsible social media users.

(ii) Representors' Views and Recommendations

259. Many thoughtful views and recommendations in relation to public education were put forth by representors for the Committee's consideration. They relate to the following areas:

[167] Chong Ja Ian, Appendix III: Written Representations, Paper No. No 53, B248, para 3.
[168] Nic Newman et al, "Reuters Institute Digital News Report 2018", *Reuters Institute for the Study of Journalism, University of Oxford*, p 33.

a. Content of public education;
b. Incorporating critical thinking skills in schools;
c. Target audience;
d. Modalities of public education; and
e. Other measures supporting public education initiatives.

260. To assist the Committee in understanding the current state of play, various representors shared what different entities in Singapore are already doing or are planning to do to promote media and digital literacy.

a. *National Library Board (NLB):* NLB has been promoting the importance of information searching and discernment since 2013 through its S.U.R.E. campaign, which comprises four key concepts: Source, Understand, Research and Evaluate. NLB uses different avenues — from conducting learning journeys and workshops to providing open-source resource guides — to promote information literacy at a national level. It has also partnered with the Ministry of Education to develop and embed into the school curriculum appropriate educational resources. NLB's new strategy for S.U.R.E. (known as "SURE 2.0") will be geared towards being more "citizen centric",[169] *i.e.* by focusing on the contextual application of information literacy skills, to enable citizens to make informed decisions in their daily lives. SURE 2.0 will have three main thrusts targeted at different segments of the population – School (students), Work (working adults) and Life (general population). These three thrusts will be supported by marketing and digital engagement efforts.

b. *Media Literacy Council (MLC)*: MLC has sought to develop public awareness and education programmes relating to media literacy and cyber wellness. Through its annual Better Internet Campaign, it has sought to raise public awareness and educate the public on detecting and debunking falsehoods. Moving forward, it has made plans to: (a) organise and facilitate partnerships between practitioners and academics to identify and articulate key concerns for research and public education; (b) develop content and resources on critical thinking and fact-checking, to be disseminated through online and offline platforms, *e.g.* online videos, social media posts, website articles, printed materials and a fake news game; (c) support ground-up projects by young people to raise awareness of online falsehoods and its consequences; and (d) work with technology companies and other institutions to promote media literacy and equip the public with fact-checking skills.

[169] National Library Board, Appendix III: Written Representations, Paper No. 40, page B161.

c. *Ministry of Education (MOE)*: Through MOE's Cyber Wellness programme, students at both primary and secondary school levels have been taught information literacy skills.

d. *Media organisations*: Media organisations have also sought to promote media literacy and educate their readers and the public at large:

    i. The mainstream media have given the issue of fake news significant coverage to increase public awareness of the issue and the fight against online falsehoods. SPH has (a) published numerous reports and commentaries about the fight against online falsehoods; (b) organised public talks to highlight ways in which readers can spot fake news; and (c) proposed opening up SPH's news archive to the public to raise media literacy by empowering citizens to have easier access to information.

    ii. Mothership has participated in public forums and community efforts to address the issue of fake news.

e. *Technology companies*: Several technology companies have launched various initiatives, public and private campaigns, workshops and dialogues to promote media literacy. Some specific examples were mentioned by Facebook, Google and Twitter in their written representations:

    i. Facebook: In September 2017, Facebook partnered with the MLC to distribute 130,000 posters (in English, Mandarin, Bahasa Melayu and Tamil) to local neighbourhoods around Singapore, to raise awareness of how to spot false news. In the same month, Facebook also launched a public service announcement on "How to Spot False News" on the Singapore Facebook page, which reached over tens of thousands of people in Singapore.

    ii. Google: Google has invested in media literacy initiatives by partnering MLC to help citizens, regardless of age, develop critical thinking, and to promote an astute and responsible participatory culture online. For example, Google is an active supporter of the MLC's *Better Internet x Youths Call for Proposals (CFP)* – where it provided co-funding support and advice to community projects and initiatives focused on tackling misinformation.

    iii. Twitter: For the Better Internet Campaign 2018, Twitter partnered the MLC for the language translation of the Digital Intelligence (DQ) Parent Handbook into the vernacular languages which benefitted Singaporean parents of different races. It hosted and conducted workshops for parents together with others like the NLB

and MENDAKI. This outreach to parents complements Twitter's ongoing work to build up discernment in Singaporean youth, through the National Youth Council, the National Council of Social Services, as well as educational institutions such as the Institute of Technical Education.

261. <u>*Content*</u>. A wide range of views were shared on what the content of public education should entail, to effectively immunise our population against the effects of deliberate online falsehoods:

a. *Digital and informational awareness:* A central component of digital literacy is to ensure people are educated on how falsehoods are spread online today, and this was emphasised by a number of representors. Mr Nimmo proposed that the public be taught how to identify a bot or troll, and the tricks and techniques of those who spread falsehoods. Dr Ecker also shared that explaining disinformation strategies and exposing misleading argumentation to people can reduce their susceptibility to future misinformation. Dr Soon and Mr Goh pointed out the importance of making people more aware of how the online space works – in particular, how the social media environment and our heuristic tendencies (*i.e.* mental shortcuts) hinder our assessment of the veracity of online information. Professor Lim Sun Sun (Professor of Media & Communication, Head of Humanities, Arts and Social Sciences, Singapore University of Technology and Design) also reminded the Committee that media literacy education today must be sophisticated enough to keep pace with the transformations that we have seen in the media landscape.

b. *The political economy of falsehoods:* Professor Lim Sun Sun highlighted the importance of educating people on why falsehoods are spread, and not only the characteristics that falsehoods have. In her view, understanding the political and economic motivations behind the production and spread of falsehoods will help make individuals more discerning and circumspect when they consume online information.

c. *Civic responsibilities and political rights:* Dr Shobha Avadhani, (Instructor at the Centre for English Language Communication, National University of Singapore) suggested educating young people on civic participation through the media; for example, to engage on social issues and meaningfully participate in debates. In Dr George's view, it is also important to educate the public on core political commitments and rights, such that the rights of different groups and communities in our society will always be respected.

d. *Current affairs:* Dr Chong was of the view that members of the public should be exposed to an open-ended series of 'current affairs' talks to

facilitate their general awareness and knowledge of world events, in order to guard against disinformation campaigns.

e. *The immorality of falsehoods:* Mr Anthony Chia was of the view that our citizenry needs to be educated on why the creation and spread of falsehoods is morally wrong.

262. *Incorporating critical thinking skills in school*. A significant number of representors also expressed the importance of incorporating critical thinking skills into our school syllabus.

263. Mr Gaurav Keerthi, a former competitive debater, proposed that our secondary school syllabus be revamped to emphasise critical reading and debating. Mr Adrian Kwek (Senior Lecturer, Singapore University of Social Sciences) proposed that critical thinking skills should be taught in all subjects, and incorporated into exams and continuous education. The goal, in Mr Aloysius Kwok's view, is to enable students to independently and critically evaluate news spread by people, to determine if they are credible news. It was also pointed out that while critical thinking is important, it should be combined with character education, so that the skills that are picked up by students will be used for positive ends.

264. *Target audience*. Representors proposed that public education be extended to all segments of our population, with the content and mode of delivery specifically tailored to the relevant audience.

a. *Students and youth*: As highlighted above, various representors had submitted on the importance of critical thinking and digital literacy being a fundamental part of school curriculum. Ms Kirsten Han proposed that such education should be a key part of the schooling experience for Singaporeans, right from a young age. Outside of the school environment, MLC has made and implemented plans to reach out to youths via various projects and initiatives, to raise awareness of fake news and its consequences, and provide guidance on how to assess the validity of online content.

b. *Working and non-working adults*: Mr Benjamin Chen proposed that digital literacy modules should be incorporated as part of training in companies for working adults, while free seminars should be provided for non-working adults.

c. *Elderly*: Many representors recognised that the elderly can fall prey easily to deliberate online falsehoods, if they are not sufficiently equipped to detect and take the necessary precautions against such scams. In response to this problem, the Info-communications Media Development Authority included a fact-checking module in the Silver Infocomm programme to

promote literacy among the elderly. NLB has also been organising talks for the elderly, contextualising principles of informational literacy for them based on the types of online falsehoods they tend to experience, such as health and financial scams.

d. *People familiar with different languages*: Dr Nekmat observed that discussions about the problem of deliberate online falsehoods tended to be in the English language, and there was a need to take into account that people may be more comfortable in a different language, such as their mother tongue.

265. *Modalities*. There are various ways in which public education can be conducted. In this regard, the Committee heard evidence from representors on the different modalities of reaching out to people, to effectively educate them on the importance of media and digital literacy. These various modalities are as set out below.

266. *First, adopt engaging methods of outreach.* Some representors highlighted the importance of using engaging and innovative methods to reach the target audience. Simply putting out information or content on media and digital literacy and hoping that people would read it may no longer be viable today.

a. *Use of a digital playbook*: Representors from the Roses of Peace (ROP) shared their plan to use a "digital playbook", to lay out the strategies that ordinary citizens can adopt online to identify and push back on fake news. ROP plans to work with its partners to develop scalable online programmes to share the playbook that can reach a wider group of netizens.

b. *Incorporating educational messages into interactive games*: Mr Keerthi shared about his experience creating www.confirm.sg as a gamified tool online to help users sift fact from fiction. He found it successful as it managed to reach a sizeable amount of audience in a more engaging manner. Similarly, teenager Mr Zubin Jain also expressed that to connect with young people, it is important for the medium to be interesting and engaging – one example being to disguise the educational message in a form of a video game.

c. *Effective use of case studies*: Mr Alan Soon proposed that the school curriculum should include a more hands-on and pragmatic approach in teaching media and digital literacy, such as the use of case studies.

267. *Second, tap on grassroots networks and committed volunteers*. Representors also shared how grassroots networks and committed volunteers can be relied on to promote media and digital literacy on the ground.

a. *ROP Ambassadors Programme*: ROP shared that it had launched its inaugural ROP Ambassador Programme, where 30 selected "Peace Ambassadors" will receive training on digital media engagement and facilitation skills, which they will be called upon to deploy in helping identify and tackle dissemination of online falsehoods. ROP believes that with the right support from the Government and partner organisations, these 30 Peace Ambassadors can help develop suitable media content in their spheres of influence to push back against the spread of online falsehoods.

b. *ROP Advocates Network*: ROP will also be launching the ROP Advocates Network, by working with different constituencies to train citizens to be peace advocates and help counter online falsehoods that affect racial and religious harmony. Once trained, these advocates will work in the heartlands to interact and educate the elderly and/or those who are not as IT-savvy. They will also act as influencers and engagers online to create a safe space for citizens to have questions relating to various religions answered.

c. *Mafindo's literacy education programme*: Mr Nugroho shared that in Indonesia, fact-checking organisation Mafindo has sent out volunteers in 15 cities to work with schools, mosques and churches, and deliver messages on using social media responsibly, and how to detect and avoid falsehoods. The mission of these volunteers is to create an "anti-hoax mindset".

268. *Third, work with NGOs*. It was also proposed that Singapore should consider cooperating with experienced NGOs such as the International Research & Exchanges Board (IREX) to roll out similar, nationwide media literacy campaigns in Singapore.

269. *Other measures supporting public education initiatives*. Representors also made recommendations on how the Government or other entities can implement other measures that would support public education initiatives.

a. *Strengthen the work of MLC*. It was suggested that the Government should strengthen the MLC, by giving it more resources to conduct more effective public education campaigns to help people better distinguish deliberate falsehoods from credible news.

b. *Deepen collaboration and research*. MLC proposed that academics and practitioners should collaborate more to identify and articulate key concerns in relation to the topic of fake news for the purposes of research and public education. The results or findings of this collaborative effort and research can then be shared with the public for their edification.

c. *Monitor the ability of the population to discern and counter fake news*. Mr Andrew Fung proposed that a university or research institution develop a "Fake News Maturity Index" to measure an individual and the population's ability to discern and counter fake news. This could provide a basis for Singapore to measure how well-prepared its population is against the threat of deliberate online falsehoods, as compared to other countries.

(iii) Observations and Recommendations

270. The Committee fully agrees that public education on media and digital literacy and critical thinking have an essential role to play in strengthening our individual defences against deliberate online falsehoods. This is a necessary endeavour, and is one long-term solution against deliberate online falsehoods. Existing initiatives are a strong base to build on. In this regard, the Committee commends the existing efforts by various ground-up groups, the media, technology companies and public agencies. It is essential, however, that the existing efforts by these various bodies be specifically reviewed for their effectiveness in tackling the new phenomenon of deliberate online falsehoods, including whether they are able to reach all concerned segments of society. This should include a review of the curricula developed by schools and tertiary institutions.

271. In terms of the possible content of public education, the Committee agrees with various representors that the content has to be broad-based. This broad-based education must aim to equip people with the skills to assess the veracity and credibility of information and sources. Findings from the 2018 Reuters Institute Digital News Report, which surveyed respondents from 18 developed and English-speaking countries, suggest that people may be lacking in such skills. In response to a question on how news is selected on social media platforms, it was found that 40% of respondents did not know how news was selected for them: 12% thought that the decisions were made by journalists working for news organisations, 11% thought that it was by journalists employed by Facebook, and 9% thought that it was a random process. Only 29% of respondents correctly identified that the news shown to them was as a result of computer algorithms.[170]

272. In view of the sophisticated techniques used by malicious actors, the Committee stresses that public education has to effectively explain the motivations and agendas of disinformation agents and the strategies they employ, or in Professor Lim Sun Sun's words, educate people on the "political economy of falsehoods". The curricula used should be regularly updated in view of the evolving digital media landscape and insidious techniques used by malicious actors, which are increasingly difficult to detect. The Committee is also of the view that it is important to educate people on their responsibilities in producing and sharing

[170] Nic Newman et al, "Reuters Institute Digital News Report 2018", *Reuters Institute for the Study of Journalism, University of Oxford*, p 34.

content online – for example, how they should promptly pass on corrective information received to those in their spheres of influence.

273. The Committee accepts the recommendation put forth by representors on the importance of moral and civic education – in particular, education on civic participation, engagement in public discourse, and respect for different communities in society. Such education will not only help to immunise individuals against deliberate online falsehoods; it will also have the benefit of strengthening our social cohesion, which deliberate online falsehoods often seek to tear apart. There is value in considering how such education – which takes place predominantly in schools today – can also be implemented outside of the formal education system.

274. The importance of incorporating critical thinking skills in schools was also highlighted by various representors. The Committee is of the view that the Government should consider the proposal, as suggested by one representor, to have "more hands-on and pragmatic approaches" to inculcate critical thinking skills in schools.

275. The Committee agrees that public education needs to reach all segments of society, from children, to adults, to the elderly, and to people comfortable in different languages. Initiatives such as the Silver Infocomm Curriculum, which focus on the specific needs of particular segments of society such as the elderly population, should be encouraged and promoted. The Committee stresses the importance of ensuring there are sufficient outreach programmes not just to the young and educated, but to those who are less educated and less Internet-savvy as well.

276. To have an effective impact and reach, it is crucial for public education efforts to be mounted on the appropriate medium. The Committee commends the different, innovative modalities which ground-up initiatives such as the ROP have adopted to improve public education efforts. The Committee urges agencies and entities involved in public education to consider the various modalities proposed by representors. This includes the suggestion that educational messages should be incorporated into interactive games to reach out effectively to the youth. There is also room for greater partnership between government agencies and committed individuals, volunteers and NGOs who are dedicated to public education efforts in the community, but may be constrained by lack of resources in their work.

Recommendation 1. To ensure that public education efforts have the necessary scope and scale, the Government should consider putting in place a national framework to coordinate and guide public education initiatives. This framework should have the following elements:

a. An expanded, broad-based curriculum for schools that would include:
    i. a component specifically on the motivations and agendas of disinformation agents and their techniques and strategies;
    ii. moral and civic education, to foster active and constructive public discourse and responsible online behaviour; and
    iii. imparting critical thinking skills creatively.

    This curriculum should be regularly updated with the latest research and knowledge about the problem of online falsehoods.

b. A framework of desired skills and outcomes to:
    i. guide public education efforts in building information and media literacy among Singaporeans. This framework should similarly be informed by research on the problem of online falsehoods; and
    ii. coordinate ministry actions, including overarching outreach, to ensure coverage of all segments of society.

Recommendation 2. The Government should consider encouraging and providing the necessary support for innovative and ground-up campaigns or initiatives for public education, to widen effective outreach beyond Government-led initiatives.

b. Support quality journalism

(i) Rationale and context

277. Besides public education, supporting quality journalism is also an essential tool in nurturing an informed public. Quality journalism aims to publish information in a manner that is accurate, informative, purposeful, and helps readers make sense of what is going on in the world.

278. Quality journalism serves two important functions. *First*, it helps prevent otherwise credible news sources from becoming (intentionally or not) agents in amplifying deliberate online falsehoods. This is especially so in an age where, as Dr Bontcheva observed, blatant lies often make the rounds, are re-posted and shared thousands of times, and sometimes even published by mainstream media thereafter.

279. *Second*, quality journalism also provides an option for those who might otherwise turn to websites peddling false news or other questionable online platforms for

news. The importance of maintaining a trusted mainstream media was emphasised by some representors.

280. The challenges faced by the news industry today make quality journalism more imperative, yet more difficult to achieve.

a. *First*, the competitive environment and profit-driven nature of the news industry does not lend itself naturally to good journalism. Dr Bontcheva pointed out that today's highly competitive online media landscape has resulted in poorer quality journalism and partisan reporting by media outlets, with misinformation, bias and factual inaccuracies routinely creeping in. This has, in Dr Wardle's view, led to a proliferation of clickbait headlines and sensational, image-dependent, and emotionally-driven coverage. The line between news and entertainment is also blurred, one example being the coverage of the 2016 US Presidential Elections.[171] Non-sensationalised news in turn can become disregarded as a result, simply because it is less entertaining.

b. *Second*, trust in the media is now at an all-time low globally. Based on the 2018 Edelman Trust Barometer Global Report, the media is now the "*least trusted institution*" compared with businesses, governments and NGOs. Out of 28 countries surveyed, the media in 22 countries are considered "*distrusted*" by their respective populations, having fallen below the 50% mark; including countries like the US and the UK.[172] In their written representations, representatives of TrendMicro also highlighted a survey which shows that in the US, 58% of respondents believed that the mainstream media spread online falsehoods.

281. The importance of ensuring accuracy in journalism was emphasised by local journalists, representors from traditional media organisations as well as online news platforms who testified before the Committee. Former journalist Mr Fang acknowledged that ensuring accuracy in reporting is "first and foremost the role and should be the core value of the media and of journalism".[173] The senior editors of SPH and CNA also affirmed that they have "zero tolerance for falsehoods, regardless of whether it has a low impact or high impact".[174] Representatives from

[171] For example, people appeared to be more drawn to jokes made about the supposed incompetence of the political candidates, rather than their actual policy positions.

[172] According to the Edelman Report, in Singapore, trust in journalism (mainstream and online-only media) is at 66%, which is significantly higher than the situation in the UK (53%), the US (53%) and Australia (52%). See "2018 Edelman Trust Barometer: Global Report", *Edelman Trust Barometer Annual Global Study* (2018), p 20.

[173] Nicholas Fang, Appendix IV: Minutes of Evidence, page C1175, para 10682.

[174] Walter Fernandez, Appendix IV: Minutes of Evidence, page C505, para 4344; Warren Fernandez, Appendix IV: Minutes of Evidence, page C505, para 4346. In their written representations they also set out their commitment to quality journalism. SPH recognised the critical role of its news platforms as "honest brokers", helping readers to stay informed and to distinguish between credible news and misleading or false reports. This entails newsmakers and advertisers doing their part to verify or disclaim rumours in a timely manner, and avoid speculative, misleading information. CNA pledged to continue to be an accurate, credible and trusted source of news and information. CNA believes that its capabilities and newsrooms must continue to be strengthened, to

Mothership described how they have a group of editors who closely check their articles for factual accuracy before publication.

(ii) Representors' views and recommendations

282. The following views and recommendations were submitted by representors on how quality journalism can be better supported in Singapore:

a. Encourage more fact-checking and investigative journalism;
b. Improve the standards of online citizen journalists;
c. Set similar standards for online news platforms; and
d. Remove financial pressures within the news industry.

283. *Encourage more fact-checking and investigative journalism*. Various representors recognised and highlighted the importance of training journalists in fact-checking and investigative journalism.

284. Dr George proposed that journalists should get advanced training in verifying content, to know how to spot sophisticated falsehoods like doctored images. He was of the view that funding should be provided for universities and other organisations to run free workshops to train journalists. The need to provide adequate support and funding mechanisms to strengthen media institutions was also echoed in the written representations of The Independent Singapore.

285. The importance of training journalists is recognised by technology companies like Google as well, which through its Google News Lab Training Network has trained and continues to train many journalists around the world. Dr Wardle also highlighted the importance of providing additional investment and training opportunities for employees of non-partisan media.

286. According to other representors, news platforms should take their own initiative to encourage more fact-checking. Mr Benjamin Goh suggested that news platforms can work together with fact-checkers to allow users to be quickly and effectively notified of a false claim to which they have been exposed. Dr Wardle pointed out that different news platforms should consider working collaboratively to ensure that manipulation tactics will be flagged and shared between themselves, and to prevent duplication of efforts amongst different newsrooms.

287. *Improving the standards of online citizen journalists*. Given the rise of alternative news platforms today, Mr Andrew Loh, co-founder of The Online Citizen, proposed that the Government fund the training of online journalists as well. Besides monetary support, Mr Loh suggested that the authorities meet more frequently with online journalists, to have dialogues and better understand each other's perspectives.

address the problem of deliberate online falsehoods through maintaining trust in the mainstream media via high-quality journalism, fact-checking and in-depth reporting.

288. *Set similar standards for online news platforms*. Mr Zhulkarnain and Mr Goh proposed that online news platforms should be held to similar standards of journalism as the mainstream media. One way of doing so would be to encourage or require the alternative media to subscribe to a code of conduct.

289. *Remove financial pressures within the news industry*. Mr Teymoor Nabili and Mr Fang were of the view that quality journalism can be enhanced if news organisations could be freed from financial pressures, so as to focus their attention on accountable journalism.

a. Mr Nabili shared that one possibility was for news organisations to be allowed to enjoy tax breaks and crowdfund their operations based on "a new legal space to allow for a hybrid model of news funding".[175]

b. Mr Fang proposed to separate the news functions of news organisations from the rest of the business, and be held under a not-for-profit umbrella where the sole mandate is to deliver excellence in journalism. He proposed that the funding of local news organisations could be modelled after the British Broadcasting Corporation (BBC) in Britain, which is funded principally by an annual television licence fee charged to all British households, companies and organisations using any type of equipment to receive or record live television broadcasts.

(iii) Observations and Recommendations

290. The Committee is of the view that quality journalism is an important public good which needs to be continually supported and nurtured. Quality journalism is a pillar of a society's information ecosystem. It ensures effective communication between the Government and the people, and between different segments of society. It also helps the public understand the world around them. The Committee agrees with the representors who said that having trusted sources of reliable information put forth by quality journalists is critical because it militates against a culture of doubt and confusion that can be brought about by the spread of deliberate online falsehoods.

291. The 2018 Edelman Trust Barometer also found that trust in journalism as a source of general news and information is on the increase; compared to the trust in various online and social media platforms. The Committee notes the "*zero tolerance for falsehoods*" approach of journalists and online news platforms who appeared before the Committee.

292. It is important that journalists are trained to engage in accurate reporting to ensure quality journalism. News organisations, institutes of higher learning and

[175] Teymoor Nabili, Appendix III: Written Representations, Paper No. 31, page B101.

technology companies have a role to play in this regard. The Committee believes that it will be good to have courses and workshops where journalists of all stripes can obtain further training, as recommended by various representors.

293. The Committee also agrees that there is room for greater dialogue between the Government and news platforms committed to quality journalism, including those which solely operate online. This will help both sides to better understand each other's perspective, and build up a relationship of trust that is committed to the pursuit of truth in the information ecosystem in Singapore.

294. The Committee accepts the view of representors who have argued that both the mainstream and alternative media should be held to the same journalistic standards of intellectual integrity and factual accuracy. They should be free to air views in a responsible way, which encourages considered discussion and critical thinking. Quality journalism can be maintained if (a) journalists (whether mainstream or online) maintain professional standards and are factual and accurate in their reporting; and (b) readers are sufficiently discerning to shun news platforms which are consistently unwilling or unable to abide by basic journalistic standards.

295. The Committee recognises the commercial challenges which news organisations face today. It is beyond the Committee's remit to deal with this issue. The Committee suggests this matter be further considered by the Government and the media organisations. In this era of digital disruption, news organisations will have to actively reinvent themselves to meet the demands of their readers – both in terms of the speed and interface which news are expected to be presented today, and also the quality and accuracy of the content itself.

Recommendation 3. News organisations, technology companies and institutes of higher learning should consider ways to ramp up the training of journalists of all backgrounds, especially in techniques for ensuring accuracy in a new and rapidly evolving digital news environment.

Recommendation 4. Journalists should also proactively find ways to update their skills in digital fact-checking, and arm themselves with knowledge of how online falsehoods and disinformation campaigns work.

Recommendation 5. Both the mainstream media and the alternative news platforms should hold themselves to the same professional standards of journalism, ensuring there is fairness, accuracy and integrity in reporting.

Recommendation 6. The Government should consider how it can support the objectives in Recommendations 3 to 5.

(2) Reinforce Social Cohesion and Trust

296. Trust holds a country and society together despite attempts to divide. At the same time, it is this trust that disinformation agents seek to erode. The lack of trust in society is a vulnerability often exploited by online falsehoods, resulting in division and polarisation.

297. Many representors, including experts, religious representatives, and concerned laypersons, therefore emphasised the importance of shoring up trust and cohesion in society. The importance of this endeavour was underscored by Dr Shashi, who quoted the observation that "[i]t is easy to manufacture a lie, and relatively cheap to distribute it widely. To demolish that lie takes intensive effort, and meanwhile the nature of the Internet ensures that it lives, breeds and reinforces other lies." He described efforts to reinforce resilience and a national consensus as "painstaking work that will require constant tending".[176]

298. Representors spoke of trust in two main areas: (i) among people and communities, and (ii) in the Government.

a. Strengthen trust among people and communities

(i) Rationale and context

299. Undermining social cohesion is one serious impact that deliberate online falsehoods have had, as described above at [111]-[118]. Strengthening trust in society is an important means of bridging the fault lines that could be exploited by perpetrators of deliberate online falsehoods. Experts and representors who spoke as members of their ethnic and religious communities underscored this. For example, a representative from NCCS explained that nurturing a culture of trust would enable people to be more discerning and sceptical in the face of divisive disinformation.

300. Prejudices and vulnerabilities in trust may occur along multiple fault lines, such as between and within racial and religious groups, between home-grown and new citizens, citizens and immigrants, between people of different socio-economic classes, and between groups with different ideological world views. The evidence showing some of the vulnerabilities in Singapore's social cohesion has been set out above at [220]-[237].

301. While Singapore is doing well in maintaining a harmonious society, there will always be vulnerabilities. As acknowledged by the representative from the RCC, efforts are needed to keep society's fissures and fractures as narrow and minimal as possible.

[176] Shashi Jayakumar, Appendix III: Written Representations, Paper No. 52, page B338.

(ii) Representors' views and recommendations

302. The topic of social cohesion was addressed in the evidence of a significant number of representors, including religious organisations, experts on the topic, and those who spoke as members of their respective communities.

303. The following specific measures were proposed:

a. *Convening people and community leaders*. Representors proposed convening people and community leaders to discuss their different perspectives on an issue, and to address the influence of divisive falsehoods. Dr Shashi emphasised the need for in-person interaction when doing so, drawing an analogy with efforts to de-radicalise people with extremist views. He also explained that this should go beyond the current Inter-Racial and Religious Confidence Circles, in order to reach people outside of established networks.

b. *Platforms to clarify and respond to falsehoods*. Representors saw a need for more "safe spaces" for people to discuss sensitive issues related to race and religion, and for divisive issues and falsehoods to be responded to by serious and reasoned opinions put forward by ordinary citizens. Different types of platforms were proposed for doing so, including dedicated websites and credible organisations.

In that regard, Dr Mathews shared how 64% of respondents in a 2016 study by CNA and the Institute of Policy Studies agreed that it was very hard to discuss issues related to race without someone getting offended. Around 25% had questions about other races, but did not ask them out of concern about possible ramifications. The issues they were concerned with largely related to the religious beliefs and practices of racial groups.

c. *Grassroots outreach*. Representors proposed proactively conducting outreach among communities to counter divisive falsehoods. This included collaborating with local cultural and religious leaders, in order to better contextualise falsehoods within their communities' contexts. According to Mr Nugroho, in his experience, such direct outreach had helped to positively change people's views.

d. *Advocacy against hate groups*. "Hatewatch" NGOs in other countries were referred to as a possible idea that could be adopted in Singapore. These NGOs could track hate networks, name-and-shame key players, and educate the media.

e. *Monitoring and research*. Calls were made for the gathering of data and evidence to be able to respond early and adequately. One recommendation

was real-time monitoring of online messages to help alert society to emerging problems relating to hate. Another recommendation was to research how online falsehoods in different languages can sow discord in Singapore. In that regard, Dr Nekmat highlighted how falsehoods in a particular language may be more relatable to a particular community, and may open up local communities to influence from neighbouring countries. He also flagged the existence of online ethnic community networks on closed messaging platforms, which could be vulnerable to deliberate online falsehoods.

(iii) Observations and Recommendations

304. The Committee appreciates the candid views shared by representors about Singapore's diversity and vulnerabilities. Indeed, a necessary first step toward maintaining social cohesion is to accept that there are and always will be primeval differences in society; navigating the compromise needed to live in peace and harmony is the result of deliberate efforts and not chance.

305. The social harmony seen in Singapore today was achieved by taking an activist approach towards fostering multi-culturalism and multi-racialism. The representatives from religious organisations in Singapore agreed that the harmony enjoyed by Singapore today was due to several factors working together: recognition by different communities of their responsibility to the nation, partnership and trust between the State and religious groups, social policies to promote integration, and laws. While these have worked well, efforts to address underlying tensions and grievances have to be persistent and sustained, particularly due to the insidious effect of online falsehoods, as Associate Professor Alton Chua (Wee Kim Wee School of Communication and Information, Nanyang Technological University) explained.

306. There are existing platforms that undertake some of the recommendations raised, such as convening people and community leaders, and responding quickly to divisive issues. For example, Inter-Racial and Religious Confidence Circles were established to serve as bridges between religious, ethnic and community groups at the local level, to deepen people's understanding of the various faiths, beliefs and practices, and to respond quickly to racial and religious tensions.

307. There have also been ground-up initiatives to address issues that may divide communities. These comprise both sustained efforts as well as *ad hoc* dialogues and events.[177] Singapore has a supportive environment for ground-up efforts. Examples of support sources include the National Volunteer & Philanthropy Centre, and the Ministry of Culture, Community and Youth's (MCCY's) drive to promote active citizenry, including by funding relevant community projects.

[177] *E.g.*, "More than Just", an online community and series of small group workshops funded by MCCY's Our Singapore Fund; "A Good Day", an event that includes a discussion on what it means to be in the majority or minority, funded by the National Volunteer & Philanthropy Centre.

Ground-up initiatives are important to complement State-led platforms. The Committee commends those involved in ground-up initiatives, and encourages the continued growth of community-driven efforts.

308. Singapore's efforts to maintain social harmony will need to evolve to address new problems. The evidence showed that the following specific areas relating to deliberate online falsehoods could be strengthened:

a. One area was on equipping individuals to raise and discuss sensitive issues relating to the differences amongst themselves. Empirical research showed that people faced perceived barriers in doing so, such as a fear of offending others. This also affects the ability to address divisive falsehoods. Addressing these barriers may involve both providing the "safe spaces" to do so, which would involve having skilful moderators and facilitators, and familiarising people with how to put forward their views and queries reasonably and with sensitivity.

b. Another area was on staying abreast of how divisive falsehoods, xenophobia and hate manifest in Singapore, and identifying new and evolving vulnerabilities, in order to formulate appropriate responses early.

309. The task of maintaining social cohesion is never complete, as several representors have emphasised. It has been and should continue to be a priority for Singapore.

Recommendation 7. Organisations and initiatives for the promotion of social cohesion, both old and new, should consider providing clarifications and information on distortions and falsehoods affecting social cohesion. In doing so, they should consider adopting the following principles recommended by representors, where relevant:

a. Employ people-to-people interaction and communication.
b. Create "safe spaces" for exchanging views and perspectives on sensitive issues.
c. Serve as voices of influence in society, to cultivate a strong core of people who are less susceptible to deliberate online falsehoods.
d. Mediate honest discussion among differing groups.
e. Reach into and across "echo chambers".

Recommendation 8. The Government should consider supporting or conducting research to understand society's vulnerabilities.

b. Maintain trust in public institutions

(i) Rationale and context

310. Strong trust in public institutions makes it harder for deliberate online falsehoods to take effect. Dr Jayakumar stated that once there is underlying trust between a government and its people, the people would be less disposed to believe disinformation. Trust is necessary to enable public institutions to effectively intervene during crises, according to Dr Liew.

311. Conversely, mistrust in public institutions facilitates the uptake of falsehoods. According to Dr Berzins, a national security expert from Latvia, any gap between the authorities and society is a key vulnerability that can be used as leverage by adversaries. He cautioned that when people lose faith in public institutions, the chances of success for disinformation operations increase significantly. In that regard, public institutions are a central source of information for society. If people lose trust in public institutions, they may turn instead to less reliable alternative sources of information. Other representors echoed the similar view that deliberate online falsehoods thrived on the lack of public trust.

(ii) Representors' views and recommendations

312. Various representations touched on how to reinforce trust in public institutions, in order to ensure society stays resilient against deliberate online falsehoods. In summary:

a. Two key recommendations were made, including by experts, on how public institutions could maintain public trust when responding to or taking measures against online falsehoods. This recognised that responses to online falsehoods could themselves be exploited by disinformation agents to further erode trust.

i. The first recommendation was pre-emptive, and called on public institutions to explain to the people in advance the nature of the disinformation threat, as well as the proposed approach to responding to the disinformation threat. Expert representors such as Dr Jayakumar and Dr Gulizar Haciyakupoglu recommended using offline interactions and non-governmental initiatives in doing so.

ii. The second was reactive, and involved quickly exposing the falsehood and the techniques behind the falsehood, and putting out the facts and providing explanations in a timely manner. Representors tended to see the authorities as playing the primary role in doing so, especially when the disinformation was against public institutions. Other representors thought that non-

governmental initiatives would help lend credibility to State efforts. It was explained that responding quickly with information would mitigate speculation and conspiracy theories, which could otherwise worsen the situation.

b. Broader recommendations about governance generally were also made. These reiterated the importance of well-established principles of governance, namely, communication, transparency, participation and accountability. Several suggestions were made, including the following:

   i. explain the rationale for public policy decisions;
   ii. be candid about failures and problems faced;
   iii. undertake continuous and transparent communication with the public;
   iv. involve the public in policy and decision-making processes;
   v. demonstrate willingness to be held accountable by the public; and
   vi. foster civil society and an active citizenry.

c. Some specific prescriptions were also made for improving transparency and accountability of Government, generally. There was a suggestion to enact a Freedom of Information Act, to enable the public to request for and obtain information from public institutions. Related recommendations were made to establish an ombudsman, to assess what classified data could be disclosed, to regularly de-classify archival material, and to investigate complaints against public institutions.

(iii) Observations and Recommendations

313. Trust-building by public institutions is an important consideration when it comes to how public institutions respond to deliberate online falsehoods specifically. The Committee agrees with the pre-emptive and responsive measures recommended on this issue.

314. The Committee also notes the observations made by various representors about communication, transparency, participation and accountability. These are important facets of trust-building between Government and society generally. The rationale for the recommendations was that they will engender good governance and greater trust between Government and society, and that should in turn help to deal with deliberate online falsehoods. The recommendations were generally broad (*e.g.* Government should explain the rationale for public policy decisions). They raised issues that pervade many facets of the Government's work. Assessing these recommendations would require investigation of the following: (1) the extent to which the suggestions are already being implemented/practised, across the range of governmental actions (*e.g.* the extent to which Government explains the rationale for public policy decisions, and undertakes continuous and transparent communications); (2) the necessary reasons why there may have to be

some qualifications to such suggestions in specific areas, on grounds of security or other considerations; (3) how these suggestions align with other policy considerations of the Government, and how they should be considered together; and (4) what the gaps (if any) are, and how the recommendations (in the context of points 1 to 3 above) will, in concrete terms, help in dealing with deliberate online falsehoods. These issues, in direct relation to deliberate online falsehoods, were not dealt with in the representations.

315. The Committee makes the same observation, as regards the specific recommendations on Freedom of Information Act and ombudsman. Nevertheless, the Committee recognises that there can be different, and legitimate, points of view on whether these recommendations are good, in general, for a country. The Committee's task is however to consider measures to deal with deliberate online falsehoods. As there are countries which have such legislation and institutions, the Committee suggests that the Government studies the experience of these countries, and whether having a Freedom of Information Act and an ombudsman will help in dealing with deliberate online falsehoods.

316. The Committee recognises that many of the recommendations, though lacking specificity, were motivated by valid and important considerations, including transparency, public participation and an active citizenry. The Committee has made some recommendations for consideration by the Government (see Recommendation 10 below). The Committee has also sent a summary of these recommendations to MCCY due to its oversight of citizen engagement efforts, for the Government's notice and consideration. MCCY's response can be found in **Annex G**.

Recommendation 9. Public institutions should, wherever possible, provide information to the public in response to online falsehoods in a timely manner. They should also seek to pre-empt vulnerabilities and put out information in advance, where appropriate, to inoculate the public. They should ensure that they communicate with the public in clear and comprehensible terms.

Recommendation 10. Existing efforts should be reviewed, to consider whether they are adequate to achieve the following:

a. *Transparency.* Swiftly communicating information in response to online falsehoods, the reasons for any Government action against online falsehoods, and the reasons for decisions to not disclose information to the public.

b. *Participation and communication.* Engaging the public on Government strategies against online disinformation operations.

c. *Accountability*. Assuring the public of the integrity of the information the Government puts forward concerning public institutions.

(3) Promote Fact-checking

a. Rationale and context

317. Promoting fact-checking initiatives which can promptly debunk falsehoods is recognised by many representors as a possibly useful countermeasure in combatting deliberate online falsehoods. This is because of the two functions which fact-checking initiatives serve.

318. The first function is to counter falsehoods by informing the public of corrections and facts. In this regard, the Committee received substantial evidence on the strengths and limitations of fact-checking. Some representors pointed out that corrections and falsehoods often do not overlap, such that those exposed to a particular falsehood may not come across the necessary corrective information. Research also shows that corrective information may not change the beliefs held by some, especially if the corrective information conflicts with a person's pre-existing beliefs.[178] Given the speed and volume in which falsehoods are spread online, Dr Ecker warned that fact-checking efforts will remain an "uphill battle".[179]

319. Despite these limitations of fact-checking, Dr Ecker did find that corrections can be effective, depending on the manner in which these corrections are framed.[180] The ability of corrections to reduce the influence of falsehoods was also supported by studies referred to by Dr Soon and Dr Bontcheva.[181] The Committee also notes the view expressed by Mr Janda, Head of the Kremlin Watch Program, that while fact-checking is largely ineffective against entrenched ideological misconceptions, it plays a useful role in defending the mainstream from the extremes. It is also important to bear in mind that while fact-checking is not always effective on every person, leaving falsehoods uncorrected is not a viable

[178] Brendan Nyhan and Jason Reifler, "When corrections fail: The persistence of political misperceptions", *Political Behavior* (2010) 32(2), 303-330; Stephan Lewandowsky et al, "Misinformation and its correction: Continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106.
[179] Ullrich Ecker, Appendix III: Written Representations, Paper No. 44, page B183, para 17.
[180] Stephan Lewandowsky et. al, "Misinformation and its correction: Continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, pp 115-116.
[181] Fridkin et al, "Liar, liar, pants on fire: How fact-checking influences citizens' reactions to negative advertising", *Political Communication* (2015) 32(1), 127-151; Min, "Intertwining of campaign news and advertising: The content and electoral effects of newspaper ad watches", *Journalism and Mass Communication Quarterly* (2002) 79(4), 927-944; Ullrich Ecker et al, "Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction", *Psychonomic Bulletin & Review* (2011) 18(3), 570; "*New studies on political fact-checking: Growing, influential; but less popular among GOP readers*", *American Press Institute* (2015).

alternative. This is because not responding to the falsehood at all creates more space for it to take hold in the collective consciousness, and the falsehood will become harder to dislodge the longer it goes unchallenged.[182]

320. The second function fact-checking initiatives serve is to encourage people to pursue accuracy and veracity of information, as observed by Dr Bontcheva. Ultimately, the corrections and facts put out by fact-checking initiatives help to create a culture or mindset in society that emphasises the importance of truth, reinforcing the importance of being accurate and properly informed before coming to any decision.

b. Representors' views and recommendations

321. The Committee received many views and recommendations from representors on the different ways in which fact-checking initiatives can be established and operate. The diversity of these recommendations show the varied roles and purposes that fact-checking initiatives can have. At [322]–[344] below, the Committee sets out the evidence received in relation to fact-checking initiatives, as follows:

a. Possible types of fact-checking initiatives;
b. Degree of Government involvement in fact-checking initiatives;
c. Scope, tools and responsibilities of fact-checking initiatives; and
d. Related measures that can aid or encourage fact-checking.

(i) Possible types of fact-checking initiatives

322. There are broadly speaking four main types of fact-checking initiatives which were shared or proposed by representors.

323. *First, a fact-checking initiative can be run by journalists from media organisations.* This is the case for many existing fact-checking bodies, which are part of the work of news organisations, *e.g.* BBC's Reality Check, Le Monde's Decodex, and BuzzFeed's Fact Checker. Dedicated fact-checking organisations such as PolitiFact, Snopes, and StopFake, are also staffed by journalists.

324. *Second, some representors proposed establishing a fact-checking coalition, made up of media players, industry practitioners and other interested parties like technology companies and non-government entities.* SPH shared that there are already such coalitions in other countries, such as the CrossCheck project in France. Besides being able to verify information with entities from different industries, Mr Warren Fernandez (Editor-in-Chief, SPH) explained that the advantage of having such a coalition is that different media organisations will be

[182] Craig Silverman, "Lies, Damn Lies, and Viral Content", *Tow Center for Digital Journalism, Columbia Journalism School* (2015), p 152.

able to work collaboratively to verify claims, and reduce the "competitive instinct" between themselves to be the first to publish unverified claims.[183]

325. *Third, some representors were of the view that a fact-checking initiative should be volunteer-driven and/or community-based*. According to Mr Wilson Na, fact-checking initiatives should be run by community partners and grassroots volunteers, given that it is the community which stands to "lose the most" from the propagation of fake news.[184]

326. There are various examples of such fact-checking initiatives. In Indonesia, Mafindo runs a fact-checking initiative that relies on mainly citizen volunteers on a crowdsourced platform. A fact-checking body known as the "Baltic elves" was also set up in Lithuania by citizens who have banded together voluntarily through social media to debunk falsehoods. Locally, Professor Lim Sun Sun highlighted a fact-checking platform being developed by students at the Singapore University of Technology and Design (SUTD) as an example of a community-driven fact-checking effort.

327. *Finally, a fact-checking initiative can be established and maintained by the Government.* Mr Rajesh Sreenivasan suggested that the Government create and actively maintain, with the help of major online content platforms in the private sector, a local fact-checking/myth-busting database that members of the Singapore public can refer to as a trusted first port-of-call should they wish to verify the truth or veracity of any Singapore-related news circulating online.

328. Other fact-checking initiatives were brought to the Committee's attention, as follows:

    a. *"Factually" in Singapore*: In Singapore, the Ministry of Communications and Information had set up *Factually*, which seeks to dispel and clarify false information that has gained sufficient public attention. This dedicated fact-checking website had helped, amongst other things, to refute false rumours being spread about Singaporeans' CPF savings, an issue close to many Singaporeans' hearts. Such "direct responses" are said to be far better in capturing people's attention than mere explanations of Government policy.

    b. *"StopFake" in Ukraine*: Mr Deynychenko shared that his fact-checking organisation "StopFake" had, over four years, collected thousands of examples of Russia's purposeful dissemination of fakes and manipulations. Disclosing such information to the public, according to Mr Deynychenko, contributed to the decrease in Ukraine of public trust in sources of foreign disinformation.

[183] Warren Fernandez, Appendix IV: Minutes of Evidence, page C520, para 4448.
[184] Wilson Na, Appendix III: Written Representations, Paper No. 30, page B96.

c. *"Maltido Bulo" in Spain*: According to Mr Nimmo, Spanish fact-checking group Maltido Bulo helped to expose a number of fake news during and after the Spanish referendum in 2017.

(ii) Degree of Government involvement in fact-checking initiatives

329. There was a diversity of opinions from representors on the issue of how involved the Government should be in fact-checking initiatives.

330. *Strictly independent from the Government*. Some representors expressed that fact-checking initiatives should be strictly independent from the Government. To these representors, fact-checking initiatives draw their influence from their credibility, rather than from authority conferred on them by the State. The reasons offered as to why fact-checking initiatives should be strictly independent from the Government are as follows:

a. *First, to increase the credibility of the fact-checking initiative*. Some representors expressed concern that any form of Government intervention or influence would lead the fact-checking initiative to be *perceived* as spreading propaganda rather than unbiased facts. According to Mr Deynychenko, the fear of losing credibility is the reason why StopFake does not depend on Government support at all.

b. *Second, to allow the fact-checking initiative to be able to fact-check a wide range of political issues*. Mr Shaun Lim argued that a fact-checking initiative which is not independent from the Government, such as *Factually*, may not be able to objectively conduct fact-checks when issues of politics or governance are involved – and these are precisely the areas today which may require fact-checking.

c. *Third, to prevent the Government's own reputation from being harmed*. Mr Benjamin Ong shared the concern that if the fact-checking initiative is not completely independent from the Government, the Government may be seen as arrogating to itself a purported monopoly on truth, which can backfire by leading to a perception that the Government is acting in a self-interested manner.

331. *Independent, but financially supported by the Government*. A slightly different position is that while a fact-checking initiative should be fully independent from the Government in terms of its everyday functions, the Government can nonetheless provide funding to the fact-finding initiative to support its work, so long as this does not affect the latter's independence.

332. *Government involvement in limited circumstances*. Some representors stated that while a fact-checking initiative should be run independently, Government

involvement may be necessary when State-backed information is needed. For example, when issues pertaining to national security are involved, both SPH and CNA agreed that the Government would have to be part of the process of fact-checking. Both Mr Janda and Mr Nimmo also said that while civil society should play the main role in investigating and refuting falsehoods, the government in question may need to be involved if foreign powers are trying to interfere in local elections, or stage a large-scale attack on the information environment in its country.

333. *Government as one representative amongst different stakeholders in a network of actors*. Another modality proposed was for the Government to be represented amongst various actors in a fact-checking network. This is because having other actors, such as non-profit entities, collaborate with State agencies in a network of fact-checking platforms will allow for better cross-verification of information.

334. *Established and maintained by the Government*. As mentioned earlier, a few representors were comfortable with a fact-checking initiative being established and maintained by the Government. Whether it is linked to the Government or not, representors are in agreement that a fact-checking initiative *ought to* focus on presenting true and accurate facts to the public.

(iii) Scope, tools and roles of fact-checking initiatives

335. Representors also shared on the different scope, tools and roles of fact-checking initiatives.

336. *Scope of fact-checking*. The scope of fact-checking initiatives can differ widely, depending on the intention for which they were set up.

a. *Fact-checking user-submitted information*: Some fact-checking initiatives, such as the one proposed by SPH, would operate by only verifying and debunking information which is submitted to them by users.

b. *Fact-checking for specific events*: Other fact-checking initiatives provide detection to debunking services for *specific events* – such as during elections or in political debates.

c. *Fact-checking specific subject matters*: A proposal was also made for fact-checking initiatives to focus on specific subject matters. For example, there could be a fact-checking body which focuses on matters related to public institutions, to help verify whether any alleged document or press release truly came from Government agencies.

337. *Use of technological tools to detect falsehoods*. Various representors proposed that fact-checking initiatives should employ advanced technological tools to help verify facts speedily and accurately.

338. Mr Zhulkarnain proposed the use of blockchain technology to assess content and verify documents. He also proposed that online document authentication be used. This could come in the form of applications that allow users to input images or facts to be authenticated.

339. As computer scientists, Dr Bontcheva and Dr Farid shared with the Committee the technological tools they have worked on, which can help to automatically detect false information online. Significantly, both of them also cautioned that many of these automated tools are not yet accurate enough to operate on a large scale, and that human reviewers are still required in the process of detecting falsehoods.

340. *Recommend or decide on appropriate enforcement action*. SPH and CNA proposed that the fact-checking body they have respectively recommended should have the power to either recommend or decide on appropriate enforcement actions. This would come in handy when the fact-checking body encounters incidents involving malicious falsification of information or dissemination of such false information.

341. *Engage in cross-border fact-checking*. Mr Nugroho and Ms Yang emphasised the importance of taking a regional approach towards fact-checking. Mr Nugroho shared about Mafindo's involvement in the Asia Pacific Fact Checker Network, which handles cross-border debunking of falsehoods. In Mr Nugroho's view, this is important given how falsehoods which begin in one country can spread easily to others. Ms Yang also proposed that there should be a regional fact-checking task force set up in the Association of Southeast Asian Nations (ASEAN) to support the work of fact-checking bodies like Mafindo, similar to how the EU East StratCom Taskforce was established in 2015 to counter foreign disinformation operations.

(iv) Related measures that can aid and encourage fact-checking

342. Various representors shared how providing access to more information can aid and encourage fact-checking, as it allows more entities and individuals to evaluate information for themselves.

343. *First*, some representors called for non-sensitive Government information to be published on a more regular basis. Such information can then be used as reference for fact-checking when falsehoods are being spread online, in order to debunk the falsehood. Given the use of "paywalls" by some news websites, Mr Timothy Tan also proposed that a website be created to store and archive published news, so that people will have easy access to them.

344. *Second*, some representors took the view that content producers should be encouraged or compelled to disclose where they had sourced their information.

Senior editors of SPH and CNA shared the importance of media organisations being honest by marking sponsored content upfront, so that the audience will know the origins of the content and who has had a hand in crafting it. Mr Zhulkarnain expressed that social media and online news websites should be transparent on its funding and/or political affiliations, to provide readers with the necessary information to discern the agenda or slant behind their news reporting. On a related note, some representors also proposed that online publishers be encouraged or required to post citations of original sources for the claims they publish online, in order to foster accountability and allow other readers to fact-check these claims.

c. Observations and Recommendations

345. The Committee shares the view expressed by many representors that fact-checking is a tool that can be deployed in tackling deliberate online falsehoods. At the same time, the Committee also notes the concerns raised by several representors on the limitations of fact-checking initiatives. In particular, fact-checking has been shown to have limited effect on those with ideologically-entrenched views. De-bunks and corrections may also fail to reach those who have been exposed to the falsehood. This explains why Dr Soon and Mr Goh – whilst supportive of fact-checking efforts – have argued that fact-checking is "*not a panacea and has its limitations*".[185]

346. This is not to say that fact-checking initiatives should not be encouraged and undertaken. The Committee recognises that having trusted fact-checking initiatives can help remedy the influence of falsehoods on people, and prevent particular falsehoods from spreading further by warning the wider community in advance. Such initiatives can also play a broader role in promoting a culture of accuracy and veracity in society. Furthermore, the process of de-bunking falsehoods may also expose, to a significant segment of the public, the nature and use of deliberate online falsehoods, thus serving as an important tool of public education concurrently.

347. The Committee supports the proposal put forth by some representors for a fact-checking coalition, comprising different media organisations and partners from other industries (like technology companies) in Singapore, to be established. The Committee is of the view that such a coalition could pull together valuable resources from otherwise competing media organisations, and tap on the expertise of partners from different industries to fact-check the falsehoods quickly and accurately. The involvement of different media organisations can help increase the coalition's credibility and its success rate in debunking falsehoods.

348. The Committee also encourages the setting up of other ground-up fact-checking initiatives. One laudable example is the fact-checking platform being developed

[185] Carol Soon and Shawn Goh, Appendix III: Written Representations, Paper No. 62, page B370, para 29.

by students at the SUTD, as highlighted by Professor Lim Sun Sun. As Dr Soon and Mr Goh put it, these independent fact-checking initiatives should not be viewed "*as a threat and should [be accorded] with the independence to develop their own structure and processes*".[186] Ultimately, an ecosystem of credible fact-checking initiatives committed to the common pursuit of accuracy and veracity would only benefit society.

349. The Committee notes the divergence of opinions expressed by representors on the role which the Government should play in fact-checking, and that most of the fact-checking initiatives presented to the Committee are industry-led or ground-up initiatives which are independent from public institutions. The Committee is of the view that the role, if any, that public institutions can play in supporting fact-checking initiatives in general, or a fact-checking coalition specifically, needs to be further considered, taking into account, amongst other things, the following:

    a. *First*, based on the 2018 Edelman Trust Barometer, trust in public institutions by the general population in Singapore is high at 58%. This is significantly higher than other countries (*e.g.* US (43%), France (40%), and UK (39%)) where many fully-independent fact-checking initiatives operate.

    b. Second, there may be resource constraints on a fact-checking coalition set up solely by participating media organisations in Singapore, given the relatively small size of our news media industry. It is noteworthy that the CrossCheck Project in France which SPH referred to had involved a total of *thirty-seven* newsrooms and technology partners at its inception, with its fact-checking efforts focused primarily on the lead-up to the French Presidential election in May 2017 (*i.e.* over a few months). A fact-checking coalition that is meant to operate credibly in the long haul will require a substantial amount of resources.

350. Ultimately, whether a fact-checking coalition will be trusted and relied upon by people depends on its credibility and its effectiveness. A fact-checking coalition that ends up being a partisan, propaganda mouthpiece of the Government of the day will very quickly lose its credibility, be of no utility to people, and, as one representor pointed out, end up damaging the Government's own reputation in the process. A fact-checking coalition (or any fact-checking initiative for that matter) must have sufficient independence and competence, where the fact-checking initiative is ultimately committed to presenting the truth to the public.

351. In this regard, the Government can consider whether it should or is able to provide support to credible fact-checking initiatives as appropriate. The UK Digital, Culture, Media and Sport Committee Interim Report published on 29 July 2018 ("UK Committee Interim Report"), for example, has suggested that the UK

[186] Carol Soon and Shawn Goh, Appendix III: Written Representations, Paper No. 62, page B370, para 28.

Government initiate a working group of experts to create a credible annotation of standards describing the level of verification of different websites.[187]

352. While cross-border fact-checking is a laudable proposal, the Committee is of the view that this proposal should be considered at a later stage, after an eco-system of fact-checking has been successfully entrenched in Singapore.

353. The Committee also shares the view that having easy access to credible and accurate information in general can aid and encourage fact-checking, especially by individuals who wish to evaluate information for themselves. The Committee is also of the view that content producers online should consider the proposals put forth by representors on the importance of marking sponsored content upfront, and disclosing their sources of information comprehensively. This will aid in creating a culture of accuracy in society, to deter the creation of online falsehoods in Singapore.

Recommendation 11. There is a role for trusted fact-checking initiatives in combatting deliberate online falsehoods. Different media organisations and partners from other industries should consider establishing a fact-checking coalition in Singapore to debunk falsehoods swiftly and credibly, or providing relevant support to such credible fact-checking initiatives as appropriate. There are differing views on the role, if any, that the Government can play in supporting fact-checking initiatives. Thus, this aspect needs to be further considered.

(4) Disrupt Online Falsehoods

354. The playing field for the "contestation of ideas" is not at all equal when it comes to online falsehoods; this was so even before the digital age. The truth tends to be inherently weaker than falsehoods in influence, due to human tendencies such as memory and ideological pre-dispositions. These tendencies are worsened online, where people receive information in large quantities, and "echo chambers" encourage polarisation and intolerance. Shielded by anonymity online, bad actors have readily employed digital tools and techniques to amplify falsehoods, capitalised on "echo chambers", and crowded out the facts. Corrections have been unable to out-race the speed and reach of online falsehoods. While nurturing an informed public through education, quality journalism and fact-checking are important, they are, as explained at [358] below, insufficient to deal with these realities.

355. The following sections deal with the following issues, namely, (i) countering and deterring the spread and influence of online falsehoods and (ii) the nature of online platforms.

[187] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 18.

a. Counter and deter the spread of online falsehoods

(i) Rationale and context

356. A significant number of representors, both experts and laypersons, called for measures to swiftly stem the spread of online falsehoods. They were of the view that such measures are vital because of the speed at which online falsehoods can cause irreparable damage. Some of them further explained why longer-term or indirect measures such as fact-checking and public education were not enough to deal with online falsehoods.

357. *Importance of stemming the spread of online falsehoods*. Proponents of measures to stop the spread of online falsehoods made the following supporting points:

a. *First*, allowing people to be exposed to online falsehoods can lead to serious consequences. Drawing on their experience dealing with disinformation campaigns, Mr Janda and Mr Deynychenko emphasised that the gravity of the threat posed by online disinformation required swift interventions to stop the spread of disinformation from the outset. Dr Mathews emphasised the danger of allowing the circulation of falsehoods of a racial or religiously sensitive nature in the wake of an incident that could undermine social trust. Associate Professor Chua warned that passivity in the face of damaging online falsehoods would facilitate the "illusory truth effect", *i.e.* the tendency to believe what one repeatedly sees.

b. *Second*, the effects of online falsehoods can be considerably more aggravated than content spread over traditional media. This was highlighted by a group comprising a lawyer and SMU law students, who noted that the European Court of Human Rights had similarly found that content online would have greater effects than content on traditional media, due to "the ease, scope and speed of the dissemination of information on the Internet, and the persistence of the information once disclosed."[188]

358. *Indirect measures are inadequate.* The evidence was clear that long-term or indirect measures, as set out in Parts (II)(B)(1), (2) and (3) above, while important, are inadequate to prevent and remedy the damage that online falsehoods can cause. The reasons for this are as follows:

a. With regard to public education:

[188] Sui Yi Siong et. al, Appendix III: Written Representations, Paper No. No. 130, paras 19, 20 and 30.

i. Public education will not be able to completely overcome our cognitive biases and heuristic tendencies, and immunise us against falsehoods. As explained above at Part I(A)(4), deliberate online falsehoods are difficult to overcome because of our fallibility in assessing information, and our psychological tendencies and prejudices.

ii. In a similar vein, many of the reasons why falsehoods influence us may not be issues that media literacy can ever address. Various expert representors agreed that politically, racially or culturally divisive falsehoods are difficult for public education to overcome. Adopting ideological positions, such as believing in the Pizza-gate conspiracy, may also not be issues of media literacy to begin with, as pointed out in a report by Data & Society, a research institute based in the US.[189]

iii. It is also doubtful whether media literacy in itself could be effective against the deliberate techniques used in disinformation campaigns.[190] When a particular community is being targeted with hate speech, it would be unwise and futile, in Dr George's words, to respond by "*distribut[ing] media literacy leaflets*".[191] As Professor Lim Sun Sun aptly put it, media literacy itself cannot "*go far enough or fast enough*" to deal with the challenges of deliberate online falsehoods,[192] especially when we consider the speed and reach in which deliberate online falsehoods spread.

iv. Further, education is a long-term measure which takes time to take effect. This was a limitation recognised and highlighted by representors from a range of backgrounds, including Dr Nekmat, who is a strong proponent of media literacy education. SPH's editor-in-chief Mr Warren Fernandez suggested that the timeframe one is looking at for public education to have its desired effect could be "many generations".[193]

v. Studies show that media literacy efforts may not always be effectively designed and implemented;[194] this is an ongoing challenge both globally and in Singapore.

---

[189] Monica Bulger and Patrick Davison, "The Promises, Challenges, and Futures of Media Literacy", *Data & Society* (February 2018), p 6.
[190] Monica Bulger and Patrick Davison, "The Promises, Challenges, and Futures of Media Literacy", *Data & Society* (February 2018), p 6.
[191] Cherian George, Appendix IV: Minutes of Evidence, page C700, paras 5850 – 5857.
[192] Lim Sun Sun, Transcript (29 Mar), Appendix IV: Minutes of Evidence, page C1177, para 10698.
[193] Walter Fernandez, Appendix IV: Minutes of Evidence, page C496, para 4280.
[194] Mariska Kleemans and Gonnie Eggink, "Understanding news: the impact of media literacy education on teenagers' news literacy", *Journalism Education 5(1)* (June 2016), p. 74

b. With regard to quality journalism:

    i. *First*, as with fact-checking, quality information generally cannot outrace falsehoods, or overcome cognitive biases. It was said that "[g]ood news sells, but bad news sells better, and faster".[195] Good journalism may fall on deaf ears. Based on the 2018 Edelman Trust Barometer Global Report, over 60% of respondents agreed that the average person does not know how to tell good journalism from rumours or falsehoods.[196] This suggests that public education and quality journalism have to work in tandem. It also underscores the power of our cognitive biases.

    ii. *Second*, quality information may not reach all segments of the population. Without sufficient reach, it cannot be as effective. According to the Nielsen Media Index Report 2017, only 55.9% of adults read print and online newspapers in Singapore today.[197] In that regard, traditional media may need time to overcome the marginalisation of its role due to the digital revolution. Representatives of Mothership shared how technology has significantly diminished the role of traditional media and journalists, as it is now cheap and easy for anyone to broadcast information online.

c. With regard to social cohesion and trust, vulnerabilities in society will never be completely eradicated. As Czech expert Mr Janda warned, building social resilience is necessary, but is not enough. Disinformation agents will keep adjusting their strategies to capitalise on vulnerabilities, whatever they may be.

d. With regard to fact-checking:

    i. Fact-checking faces the fundamental limitations identified above at Part I(A)(4), namely, (i) human cognitive tendencies, (ii) weakness of truth compared with falsehoods, and (iii) the further and faster reach of falsehoods. In particular, the inability of fact-checks to be delivered proactively and directly to the public and those exposed to the falsehood, is a key limitation.

    ii. In addition, several expert and lay representors pointed out that fact-checkers cannot cover a significant proportion of the online

[195] Carol Soon and Shawn Goh, Appendix III: Written Representations, Paper No. No 62, page B365, para 19.
[196] "2018 Edelman Trust Barometer: Global Report", *Edelman Trust Barometer Annual Global Study* (2018), p. 24; Carol Soon and Shawn Goh, Appendix III: Written Representations, Paper No. 62, page B359, para 7.
[197] Lee Min Kok, "Digital news consumption in Singapore on the rise; The Straits Times remains most-read English paper: Nielsen survey", *The Straits Times* (2 November 2017).

falsehoods being propagated. There will be gaps. Fact-checking is a resource-intensive and time-consuming endeavour.[198]

iii. Representors from credible fact-checking organisations voiced their support for Government intervention. For example, Mr Deynychenko from well-known Ukrainian fact-checking organisation, StopFake, expressed support for the Ukrainian Government's decision to ban the broadcasting of a foreign country's television channels in Ukraine and to limit the availability of social media sites from that country. Similarly, Mr Nugroho from Mafindo shared with the Committee that he had recommended more legislation to the Indonesian government, and that he would support the Indonesian government having the power to require technology companies to take down content that is incendiary in nature.

(ii) Representors' Views and Recommendations

359. A considerable number of representors recommended that the Government put in place measures to counter and deter the spread and influence of online falsehoods. They acknowledged that legislation would play a role in doing so. The need for additional measures to safeguard the integrity of elections was also highlighted. On the other hand, some representors proposed voluntary regulation by technology companies, or adopting hands-off approaches altogether. These various positions are summarised below.

(1) Measures to counter the spread of deliberate online falsehoods

360. The Committee acknowledges with appreciation the depth and detail of the specific measures recommended by representors. This section first sets out the objectives sought to be achieved by the recommended measures. It then summarises the details provided by representors on how to operationalise these measures, and the safeguards that they should have.

361. *Specific objectives*. A range of different solutions to stem the spread of online falsehoods were proposed by representors. These solutions sought to achieve the following objectives:

a. **Provide swift access to the facts.** It was proposed that this be done through (i) the tagging of a notification of falsity and the correction to the falsehood, which could also serve to slow the spread of the falsehood itself, and (ii) broadcasting corrections across platforms.

[198] Elmie Nekmat and Carol Soon, "Silver Lining in the Battle against Fake News", *The Straits Times* (2 November 2017).

i. The rationale for this is that the effect of falsehoods is harder to displace the longer they are left unchallenged.[199] Tagging can cause people to be embarrassed to share content that their peers perceive as fake, thereby discouraging people from further sharing it.[200] Tagging would also overcome the limitation often faced by ordinary fact-checking, where corrections do not travel as widely as the falsehood and are overpowered.

ii. Psychological research, including that provided by Dr Ecker, has shown how corrections can be made to overcome cognitive biases and be more effective against falsehoods.[201] Notably, where a falsehood has gone viral, the correction should be "circulated with equal vigor", *i.e.* repeatedly, to reduce the persistence of the falsehood's influence.[202]

iii. There was considerable support for the tagging of corrections. Supporters of tagging, such as Dr George, described it as a means of allowing users to think for themselves, while nudging users to make informed choices.

iv. The Facebook representative noted that it had discontinued flagging posts that had been found by independent fact-checking organisations to be false. According to Facebook, placing a red flag next to an article may entrench deeply held beliefs.[203] Nevertheless, at least one experiment has shown that tagged warnings do lead to a modest reduction in the perceived accuracy of fake news.[204]

v. A few representors, including Ms Yang and Mr Benjamin Ang, highlighted the "Streisand effect", which occurs when the identification (including tagging) or removal of falsehoods instead increases interest in the falsehood, and leads to conspiracy theories about why the content was so treated. It was clarified during the hearing that this could occur whether or not the identification and

[199] Craig Silverman, "Lies, Damn Lies, and Viral Content", Tow Centre for Digital Journalism, Columbia Journalism School (2015), p 152.
[200] "Combating fake news: An agenda for research and action", a conference held at the Harvard Shorenstein Centre on Media, Politics and Public Policy, 17-18 February 2017.
[201] Stephan Lewandowsky et. al, "Misinformation and its correction: Continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106.
[202] Ullrich Ecker et al, "Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction", *Psychonomic Bulletin & Review* (2011) 18(3) 570, p 577.
[203] Tessa Lyons, "Replacing Disputed Flags With Related Articles", *Facebook Newsroom* (20 December 2017).
[204] Gordon Pennycook and David Rand, "The Implied Truth Effect: Attaching warnings to a subset of fake news stories increases perceived accuracy of stories without warnings*"* (8 December 2017), available at <https://papers/ssrn/com/sol3/papers.cfm?abstract_id=3035384>. The same study also found that tagged warnings might lead to an "implied truth effect", where people may believe that items without such a warning must be true. However, as pointed out by Mr Benjamin Ong, this would not apply to regulatory interventions, which do not purport to tag all falsehoods.

removal were by law or voluntary action by online platforms. Ms Yang further clarified that in some situations, identification and removal would be necessary. Mr Ang accepted that the "Streisand effect" could be reduced by the assurance of checks and balances such as a court process in the tagging or removal of falsehoods.

vi. Suggestions on operationalising the tagging feature included (i) notifying users known to have previously clicked on the relevant post that the post had been corrected, and (ii) warning users when they visit websites known to carry misinformation, and providing information on why the website had been identified as unreliable.

b. **Curb exposure to the falsehood.** A considerable number of representors thought there was a role for interventions to curb exposure to the falsehood, by removing the falsehood, shutting down sources of falsehoods, and blocking access to the falsehood or source of falsehoods.

i. A practical concern raised was that it was difficult to completely remove or block access to a piece of online content, due to archival sites and alternative ways to access a blocked site. Nevertheless, removal and access blocking remain the tools used internationally to deal with content such as extremist material, child pornography, and copyright infringements.

ii. Another concern was that it is impossible to eliminate deliberate online falsehoods completely, also known as the "whack a mole" problem. An NTUC representative explained that even when one falsehood was removed, it could surface again in another form. However, national security expert Dr Raska explained that the aim was not to counter every single falsehood, but to prioritise falsehoods that society should be protected from.

c. **Neutralise false amplifiers.** It was recommended that inauthentic accounts that amplify falsehoods, such as those run by bots and trolls or accounts which cannot be traced to real persons, be swiftly shut down. In this regard, Facebook has noted that in the context of information operations, most false amplification on its platform is driven by humans who are coordinated in operating inauthentic accounts.[205]

d. **Discredit sources of falsehoods.** It was suggested that websites known to purvey falsehoods could be tagged to warn visitors who visit the website for information. This could discourage the sharing of falsehoods from the tagged website.

[205] Jen Weedon et al, "Information Operations and Facebook", *Facebook Security* (27 April 2017), p 9.

362. *Platform-neutral.* Representors highlighted that the above measures should apply to all digital platforms, regardless of size and whether they are open or private; they should include platforms other than social media or that may be developed in future.

363. *Operationalising the proposed measures.* Detailed recommendations were made on how to operationalise the proposed measures to stop the spread of online falsehoods. Views were shared on the following issues, which are each addressed in turn further below:

a. who the decision-maker should be;
b. the threshold for invoking the measures;
c. safeguards to prevent abuse and ensure due process; and
d. early warning mechanisms.

364. *Decision-maker.* The potential decision-making bodies identified by representors were as follows:

a. *The Courts.* The Court process could be initiated by the Executive or the online platform. The Courts would then consider the matter, and determine whether or not the intervention would be made. This is similar to the mechanism used under the Protection From Harassment Act against false statements of fact that affect private persons. It was suggested that should speedy action be needed, an urgent *ex parte* application to the Courts could be made. This would involve filing an application with an affidavit.

b. *The Executive.* Associate Professor Goh Yihan, Dean of the SMU School of Law, was a proponent of this option, with a subsequent stage of independent judicial oversight.

i. Associate Professor Goh explained that the judicial process, while important, may not be sufficiently fast to deal with the rapid spread of online falsehoods. Court processes require an application for a court order to be made together with a supporting affidavit. The application must then be served on the person against whom the order is sought. That person can then file an affidavit in reply. The court may require a hearing before coming to a decision. Associate Professor Goh noted that Executive action was also used by the Broadcasting Act to take down certain material.

ii. Representors who were sceptical of Executive action accepted that there were situations, such as those involving public order, national security, and the workings of public institutions, where only the Executive would hold the facts, and where the facts should be backed by the authority of the Executive.

iii. Representors raised concerns about whether Executive action would be credible. There was concern that Executive action could feed fears over the abuse of power. It was also pointed out that Executive directions would not be able to deal with falsehoods spread by the Executive. That said, both Law Dean Associate Professor Goh and law academic Associate Professor Eugene Tan explained that judicial oversight of Executive action would serve a crucial balancing role in ensuring the propriety of the Executive's exercise of discretion.

c. *Independent body.* Directions would be issued by an independent council or ombudsman comprising representatives from different fields of expertise.

    i. It was argued that this option would address concerns over the abuse of power, and benefit from the experience and knowledge of different experts.

    ii. It was suggested that a multi-stakeholder body would be better placed to deal with contentious cases, where there were differing opinions on whether intervention was appropriate. An analogy was drawn to how independent advisory panels have advised the Government on the removal of library materials and films.

d. *Online platforms, with recourse to the Courts.* "Notice and take down" or similar models were mooted by several representors, where the online platform would decide on whether to act upon being notified by users. In unclear cases, the online platform could apply for a Court decision. This is similar to the model used by Germany's Network Enforcement Act. The UK Digital, Culture, Media and Sport Committee examining fake news ("UK DCMS Committee") also appears to be in favour of this model. In the UK Committee Interim Report, it recommended that online platforms should be "liable for content that has been referred to them for takedown by their users, and other content that should have been easy for the tech companies to identify for themselves".[206] However, doubts were expressed about whether online platforms were well-placed to make decisions in the public interest. Political data scientists in Germany, Dr Hegelich and Mr Shahrezaye, noted that Germany's Network Enforcement Act had prompted fears that the social media platforms would over-censor to avoid fines. The Asia Internet Coalition was of the view that, in egregious instances of misinformation, it was the role of the Courts or other relevant official authorities to decide if laws had been

[206] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 60.

broken; further, it was not sensible nor sustainable to mandate Internet intermediaries to make subjective judgments.

365. *Threshold for intervention.* An issue often raised during the hearing was the need for calibration in when and how to intervene. Some representors proposed detailed frameworks, such as Dr Soon and Mr Goh's "5Cs" framework, Associate Professor Chua's 2-by-2 grid, and Dr Liew's "traffic light approach". Some of the representors also proposed adhering to the principle of proportionality. The range of considerations proposed is summarised below:

a. *Nature of potential impact*. This relates to the kind of harm that may result. Representors suggested considering whether the falsehood contravened the public interest, by affecting issues such as societal harmony, electoral processes, public health, public order and security, and economic and financial stability.

b. *Likely magnitude of impact*. This involves assessing the likelihood of harm, the likely reach of a falsehood, and the frequency of its publication or re-publication.

c. *Content.* Falsehoods are found in different kinds of publications. For example, Mr Nimmo highlighted that falsehoods could be found in stories that are completely fabricated, as well as in reports of interviews with people from only one side of a debate. Different kinds of publications may warrant different treatment.

d. *Context.* Dr Soon and Mr Goh recommended considering the content within the country's political, economic and social context. They observed that what is of concern to one country may be regarded differently by another country. The purpose of the publication may also be relevant. For example, while Dr Mathews supported measures to curb exposure and access to sites that propagated online falsehoods, he also was of the view that there should be room to raise falsehoods for clarification. Some representors also cautioned that measures should allow people to discuss and debate online falsehoods.

e. *Surrounding circumstances.* Dr Liew's proposed "traffic light approach" was calibrated according to the severity of the circumstances in which the falsehood was made. He identified three types of circumstances, namely, a situation of normality, a situation of heightened tension (such as an election or riot), or a national state of emergency. Dr Koh proposed establishing an "Election Media Monitoring Commission".

f. *Identity of actor.* There was the view that different types of perpetrators, *e.g.* members of the general public, networked players, and foreign State actors, should be treated differently. A related consideration raised was

the reach of the publisher involved, such as the amount of site traffic received.

g. *Intent.* Representors highlighted that the intent of the publisher should be properly taken into account in deploying countermeasures. Some appeared to be referring only to criminal sanctions when proposing that intent be required.

366. *Safeguards.* There were two main proposals for safeguard mechanisms.

a. One proposal was for a right to appeal against the decision before an independent arbiter. Representors suggested that the appeal could be made to the Courts, an independent committee or ombudsman, or even the President.

b. Another proposal was for an independent advisory body to assist in coming to the decision. This body could help assess the appropriateness of an intervention in situations that are unclear.

367. It was also recommended that the decision-maker explain the reasons for the intervention. This was said to help mitigate conspiracy theories and suspicion, and foster understanding of the values the intervention seeks to protect. It was also suggested that opportunity be given for the target of the intervention to voluntarily remove or correct the falsehood.

368. *Early-warning mechanisms*. Some representors proposed the use of data analytics and real-time monitoring to identify early on falsehoods or online spaces that may require intervention.

(2) Measures to deter online falsehoods

369. To deter and hold perpetrators accountable, two main recommendations were made, namely, to disrupt the financial incentives for online falsehoods, and to impose criminal sanctions.

370. *Disrupt financial incentives*. Representors highlighted the need to remove the financial benefits that purveyors of deliberate online falsehoods receive. First, this would help reduce the numbers of financially-motivated purveyors of deliberate online falsehoods. Their conduct can result in serious harm, even though unintended. Second, as pointed out by law academic Associate Professor Eugene Tan, this would send a clear signal that the deliberate propagation of falsehoods would not be tolerated or profitable, especially for those whose motivations are purely financial. The recommendations made covered both digital advertising revenue, as well as other forms of financial benefits.

371. *Digital advertising revenue.* How digital advertising incentivises the creation and spread of online falsehoods has been described above at [48]-[51]. Technology companies with digital advertising services have taken some measures to address the role their services have played in the problem. These are set out at **Annex F**.

372. *Other forms of financial benefit.* Law academic Associate Professor Eugene Tan advocated for a legislative regime to require perpetrators to disgorge their profits and other financial benefits.

373. *Criminal sanctions.* The recommendations made concerning criminal sanctions may be summarised as follows:

a. *Conduct to be sanctioned.* Those who create or actively spread deliberate online falsehoods should be punished.

b. *Intent.* There must be the requisite intent. Besides those with malicious intent, some representors proposed that those who were grossly negligent or reckless as to the truth should also be punished.

c. *Harm.* Some suggested that there should be demonstrable public harm. Examples of public harm were interference with elections, de-stabilisation of the financial system, causing hatred or inciting seditious sentiments, and severe financial or reputational harm to Singapore or any of Singapore's key institutions.

d. *Nature of sanctions.* For online falsehoods that divide communities, Dr Mathews and Mr Jamari suggested having sentences that seek to rehabilitate and educate.

374. There were also suggestions for how to identify the perpetrators of deliberate online falsehoods notwithstanding their anonymity on the Internet. Examples included requiring online platforms to disclose user information, and private initiatives to trace perpetrators using investigative research methods.

(3) Additional measures to safeguard elections

375. Various representors highlighted the importance of ensuring that there are effective measures to stop the spread of online falsehoods particularly during election periods. Special attention has been given to elections. As described by the representatives of UCMC, elections should be considered a "part of the national critical infrastructure", given that they are a cornerstone of a nation's sovereignty.[207] Dr Thio pointed out that deliberate online falsehoods which attempt to undermine democratic elections rise to the level of a national security threat, akin to attempts to subvert an elected government.

[207] Nataliia Popovych and Oleksiy Makhuhin, Appendix III: Written Representations, Paper No. 54, page B280.

376. Another factor is the intensity of disinformation operations during election periods. Election periods are often vulnerable to information attacks. The Committee received evidence of how, in Indonesia, every election would allegedly create a "big wave of disinformation", and that the number of such information attacks during Indonesian elections has increased significantly over the years.[208]

377. In the same vein, some representors proposed the implementation of additional measures applicable during elections. Mr Dan Shefet, a French lawyer, proposed banning any use of micro-targeting research and techniques during elections. Law academic Associate Professor Eugene Tan suggested requiring political candidates to disclose the amount spent on social media targeting during their campaigns. Dr Koh proposed that an independent body be established during elections, to monitor and take action against content disseminated by foreign entities in our information space. While these measures target the period during elections, representors also recognised or pointed out that foreign interference does not only occur during elections, and that the problem of deliberate online falsehoods goes beyond the issue of elections.

(4) Role of legislation

378. The Committee heard evidence on what role legislation should play and whether new legislation was needed.

379. <u>*Role of legislation.*</u> The evidence showed three points about the role of legislation. *First*, online platforms will not voluntarily undertake the measures proposed at [361] above. During the hearing, Facebook, Twitter and Google (and YouTube) confirmed that they generally will not, as a matter of policy, remove content on the basis that it is false. This also meant that they would not necessarily remove falsehoods on the basis of a mere request from the Government, unless the request was backed by the law.

380. In that regard, a Facebook representative explained that Facebook's policy against assessing falsity stemmed from practical considerations. He said that unlike hate speech, terrorism, or child sexual abuse, the company would also have to provide due process involving evidence to be furnished to show that the content was false. This was not something the company was well-placed to do. Notably, the Facebook representative agreed that there should be an objective process to deal with an online falsehood that was speedy, and that ensured due process.

381. *Second*, according to Dean of the SMU School of Law, Associate Professor Goh Yihan, existing laws are limited in terms of scope, speed and adaptability when applied in the real world. Associate Professor Goh had conducted a detailed

[208] Mafindo, Appendix III: Written Representations, Paper No. 61, page B353.

analysis applying Singapore's existing legislative levers to actual incidents, and identified limitations. In summary:

a. One case analysed was of a viral online Facebook post concerning Hurricane Irma, which claimed that the hurricane had left thousands dead and that the media and authorities were hiding the truth. Associate Professor Goh found that existing legal powers for requiring removal or correction of the falsehood would unlikely apply against the original publisher or Facebook. He noted that criminal laws could apply to punish the perpetrator, but these would not stem the spread of the falsehood.

b. Another example was the #MacronLeaks incident. During the 2017 French Presidential Election, an online forum, 4chan, a well-known conspiracy network, began circulating documents supposedly proving that then-candidate Emmanuel Macron had a secret offshore account to evade tax. This occurred two hours before a televised debate between Macron and his rival. Although the Protection From Harassment Act could allow Mr Macron to apply for an order for a correction to be tagged to the falsehood, Associate Professor Goh was of the view that the court process may not be fast enough to provide an effective remedy within the time needed during an ongoing election. Associate Professor Goh also explained that swifter Executive powers under existing law did not clearly apply to content on online forums.

c. A third case examined was of fabricated stories posted by an online blogger, alleging that the motive behind the murders of nine people in an apartment in Japan was organ trafficking. Associate Professor Goh found that existing legal powers for requiring the removal or correction of the falsehood would not likely apply against the blogger or blog in respect of this falsehood.

382. Other lawyers shared Associate Professor Goh's view. A lawyer and group of SMU law students explained that "there is a gap in the regulatory tools available to deal with the chief mischief of deliberate online falsehoods, which is their near instantaneous dissemination and ease of access."[209] Law academic Associate Professor Eugene Tan was also of the view that existing legislation could be strengthened. Researchers Dr Soon and Mr Goh explained that "[a] current gap in existing legislations is they do not sufficiently address the spread or dissemination of deliberate online falsehoods."[210]

383. Some representors also addressed the issue of whether or not to enact a new statute or to amend existing laws. They said that while there were gaps in the law, a new statute may not be needed, and that additional legal powers could be placed under existing statutes.

[209] Sui Yi Siong et al, Appendix III: Written Representations, Paper No. 130, page B1131, para 4(a).
[210] Carol Soon and Shawn Goh, Appendix III: Written Representations, Paper No. 62, page B38, para 44.

384. There was an opposing view that existing legislation was adequate. The Committee noted that the representors holding this view had either not provided support for their assertion, or had only cited various pieces of potentially relevant legislation, without deeper analysis.

385. *Third*, legal action could send a positive signal about what matters to society. As Dr Ecker put it, legislation served "as a signal reinforcing the view that facts and evidence matter to the society and the leaders of the country".[211]

(5) Voluntary action by technology companies

386. There was the view that self-regulation by online platforms was adequate to deal with the problems posed by online falsehoods.

387. Technology companies including Facebook, Google and Twitter gave evidence on a number of measures to address the proliferation of online falsehoods on their platforms. A non-exhaustive list of these measures is set out at **Annex F**.

(6) A hands-off approach

388. The Committee heard differing views on the "marketplace of ideas" theory. This theory was used by several representors to justify taking a "hands off" approach to online falsehoods. They asserted that interventions in the flow of information were unnecessary, as society would eventually determine the truth through contestation in the "marketplace of ideas". There was also a view that such interventions were instead an interference in the "marketplace of ideas".

389. *Essence and origins of the "marketplace" theory.* The "marketplace of ideas" theory was famously articulated by US Supreme Court Justice Louis Brandeis, who wrote in 1927 that the solution to false speech is more speech; with more speech, the truth would prevail.[212] Constitutional law professor Dr Thio observed that this theory likely originated from American judges and philosophers in the early 20th century, before the digital age. For example, theorist John Stuart Mill had said in 1869 that false speech should be protected so as not to deprive society of "the opportunity of exchanging for truth" and a "clearer perception and livelier impression of truth". Justice Oliver Wendell Holmes in his 1919 decision in *Abrams v United States* had explained the idea as follows:

> "[W]hen men have realized that time has upset many fighting faiths, they may come to believe even more than they believe the very foundations of

[211] Ullrich Ecker, Appendix III: Written Representations, Paper No. 44, page B184, para 19; see also Liew Kai Khiun, Appendix IV: Minutes of Evidence, pages C868-869, para 7531 ("symbolic, political message"); Benjamin Ang, Appendix IV: Minutes of Evidence, page C644, para 5433.

[212] "Harmful Content: The Role of Internet Platform Companies in Regulating Terrorist Incitement and Political Disinformation", *NYU Stern Centre for Business and Human Rights*, p 3.

their own conduct that the ultimate good desired is better reached by free trade in ideas – that the best test of truth is the power of the thought to get itself accepted in the competition of the market, and that truth is the only ground upon which their wishes safely can be carried out. …"

(iii) Observations and Recommendations

390. Developments in the digital realm are outpacing the rules and norms of societies around the world. Actors seeking to create and disseminate online falsehoods find wide space in the online world to take advantage of new and sophisticated digital methods and tools with impunity.

391. Strong action is needed to ensure that the Internet does not remain a "Wild West", as the UK DCMS Committee described it to be, but a realm where people can truly enjoy the freedom and benefits that they do in the offline realm.

392. Legislation cannot be a silver bullet in itself. Like measures to nurture an informed public, strengthen social cohesion, and promote fact-checking, legislation has its limitations. However, effectively disrupting the spread and influence of online falsehoods requires legislation. Legislative measures should adhere to certain criteria, such as being calibrated in their effect and deployment, taking into account the context and circumstances. It is also important that they be accompanied by checks and balances.

(1) Countering online falsehoods

393. *Rationale*. These capabilities are needed because there will be situations where the free circulation of online falsehoods is simply untenable. To summarise the evidence considered by the Committee:

a. Exposure to an online falsehood can influence people in immediate and long-term ways that are difficult to dispel. The continued online circulation of a falsehood can increase its influence.

b. Online falsehoods can be formidable, particularly in their speed and reach. While measures such as education, quality journalism, building social cohesion, and fact-checking are important and necessary, detailed evidence was given on why they are not sufficient.

c. The necessity of these capabilities is underscored by the serious consequences that online falsehoods can have, which includes threatening national sovereignty and security, undermining key public institutions and sowing discord within societies (see Part I above).

394. *Objectives*. In view of the evidence on how online falsehoods influence and spread, capabilities to disrupt their spread and influence should be able to swiftly do the following, as proposed by representors (see [361] above):

a. Identify the online falsehood.
b. Provide access to and increase visibility of corrections, including through tagging functions and use of other platforms with significant reach.
c. Limit or block exposure to the online falsehood.
d. Disrupt the digital amplification of online falsehoods, including through the use of false amplifiers (*e.g.* inauthentic accounts run by bots or trolls) and digital advertising tools.
e. Discredit the sources of online falsehoods.

395. The measures adopted should be platform- and technology-neutral, as some representors pointed out. Methods used by bad actors are constantly evolving; even as Facebook and Twitter clamped down on the abuse of their platforms, bad actors shifted to closed platforms, where their activities were more difficult to tackle. Closed messaging platforms should be covered by the measures. There is a need to ensure that public interest is not harmed.

396. *Principles for implementation*. The above measures should be accompanied by safeguards, to achieve their purpose of protecting, rather than undermining, freedom of expression and the contestation of ideas in the "marketplace". The measures adopted should therefore seek to fulfil the following objectives:

a. The measures will need to achieve the objective of breaking virality by being effective in a matter of hours.
b. The decision-maker should be effective and credible.
c. There should be adequate safeguards in place to ensure due process and the proper exercise of power, and give assurance to the public of the integrity of the decision-making process.
d. The measures should be deployed in a calibrated manner, taking into account the context and circumstances, including potential impact and reach.

(2) Deterring online falsehoods

397. Deterrent measures are necessary. Law Dean Associate Professor Goh explained that the real world consequences of online falsehoods, including across borders, required legislation to deter and punish perpetrators of deliberate online falsehoods. Counsel for the Singapore Press Club, Dr Stanley Lai, and psychologist Dr Ecker highlighted the importance of the signalling effect such measures would have across the different layers of stakeholders.

398. The digital advertising industry has played a key role in incentivising deliberate online falsehoods. The Committee emphasises the responsibility of stakeholders

in the digital advertising ecosystem, such as digital advertising platforms and digital advertisers, to ensure they do not support purveyors of deliberate online falsehoods. While some of the technology companies have adopted de-monetisation policies, these policies do not squarely apply to accounts on the basis that they purvey online falsehoods. Whether enough is being done by these stakeholders is not at all clear.

399. The Committee agrees that criminal sanctions play a role in deterrence and accountability. Intent would be an essential requirement of criminal sanctions. The same applies for conduct – a person who intentionally spreads falsehoods through impersonation should, for example, be penalised more harshly. Criminal sanctions should take into account the fact that the ultimate instigators of online falsehoods may not always be the ones creating or spreading them. Importantly, they should also be adequate to cover online falsehoods that are designed to have serious consequences, such as election interference, public disorder, and the degradation of trust in public institutions.

400. The Government and Parliament should consider the adequacy of existing criminal sanctions. Criminal sanctions for the knowing transmission of falsehoods are provided for in section 45 of the Telecommunications Act. The Committee notes that this does not cover falsehoods conveyed over closed messaging platforms. In such a review, the need to maintain a careful balance in preventing the public interest from being harmed in the use of closed messaging platforms, and at the same time respecting communications that are personal, private, and of limited circulation, needs to be borne in mind.

(3) Additional measures to safeguard elections

401. The Government should also consider what additional measures are needed to safeguard the integrity of our elections from the harm which foreign interference and deliberate online falsehoods can cause today. Evidence was presented to the UK DCMS Committee on companies which target foreign elections, by manipulating social media, engaging in misinformation and disinformation, and doing so in such a way so as not to be identified as the source of the material.[213] A study by the MIT Internet Policy Research Initiative also revealed that electoral regulations today face limitations due to the ease in which authors of social media posts can hide their locations and identities.[214]

402. Elections are of critical importance to a nation. Similar to the UK DCMS Committee, the Committee is concerned as to whether current electoral laws in Singapore are "fit for purpose for the digital age".[215] The UK Committee Interim

[213] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), paras 206-207.
[214] "Dealing with Fake News: Policy and Technical Measures", *MIT Internet Policy Research Initiative*, p 5.
[215] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 45.

Report proceeded to make various recommendations that sought to keep up with new digital means of campaigning, and respond to the use of digital advertising by various actors, not only political parties, to spread disinformation to sway the vote during elections. The measures recommended include mandating digital imprint requirements for all electronic campaigning, increasing the fines for electoral fraud, establishing an advertising code which would apply on social media during election periods, and increasing transparency around digital political advertisements.[216]

403. Some representors were of the view that our current laws are sufficient to deal with the threat of deliberate online falsehoods undermining our elections. They referred to provisions in the Parliamentary Elections Act and the Presidential Elections Act, which prohibit undue influence and the making of false statements about the character of a candidate, and mandate a "Cooling-Off Day". However, the Committee did not receive any detailed analysis on whether our *electoral laws* are sufficiently comprehensive and modernised to combat the sophisticated methods employed by malicious actors today to undermine elections, such as the use of "dark ads", fake accounts, or the infiltration of local social media communities to influence voters. Further, as pointed out by Mr Shaun Lim, a NUS Law student, "despite the statutory imposition of a Cooling-Off Day, it is hardly likely that a foreign agent seeking to influence our election would abide by such niceties and refrain from throwing an explosive rumour into our electoral mix".[217]

404. The Committee notes the various special measures which have been proposed or implemented by both technology companies and other governments to safeguard electoral processes elsewhere:

    a. Technology companies: Twitter, for example, has created a "cross-functional elections task force" in the US to work with federal and state election officials to manage issues that arise during the campaign, verify party candidates' accounts to prevent copycat accounts, and improve its algorithm to stamp out bot accounts targeting election-related content. Google and Facebook are also implementing measures to ensure transparency in political advertisements, by identifying and disclosing the parties who pay for these advertisements.

    b. Governments: Sweden has made plans to set up a government agency to protect its elections from hostile foreign propaganda, which would identify, analyse and respond to external influence campaigns. In France, a new law was proposed in January 2018 which provides for emergency procedures that would empower judges to remove content, close user accounts, or block websites which publish false information during elections.

[216] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), paras 45, 47, 50, 142.
[217] Shaun Lim, Appendix III: Written Representations, Paper No. 133, page B1181, para 41.

405. The Committee is of the view that that the Government should conduct a detailed study on this issue, and consider, amongst other things, whether any of the UK Committee Interim Report recommendations should be adopted in Singapore, and implement the necessary measures, including legislation.

(4) The need for legislation

406. To give effect to the objectives above, legal powers through legislation are necessary, for the following reasons:

a. Legal powers would be needed to compel persons who have published the online falsehood and others, to take necessary actions to combat deliberate online falsehoods.

b. The Committee considered whether the technology companies would voluntarily undertake the above actions. The evidence before the Committee, as well as the technology companies' global track record, suggest that the outcomes are more likely to be achieved if there was a legally valid and binding order. A request by the Government for them to do so may not be enough (save perhaps for neutralising false amplifiers).[218]

407. The Committee notes the in-depth legal analysis provided by Law Dean Associate Professor Goh, which concludes that existing legal powers are inadequate to achieve the above objectives. New legislative powers will be necessary.

408. *Potential limitations of legislation.* Representors identified two potential limitations of legislation, namely, (i) the extraterritorial reach of legislation, and (ii) the ability of legislation to keep up with technology. The Committee also received views on how these limitations could be addressed.

409. *Extraterritorial reach of legislation.* Some representors raised concerns about the effectiveness of laws in dealing with online falsehoods spread from overseas. In that regard, a group comprising a lawyer and law students pointed out that it would hence be important for laws to cover online intermediaries such as Facebook. Online intermediaries can deal with any offending material originating from outside Singapore, as long as they are circulating on their platforms.

410. *Ability to keep up with technology.* There was the view that legislation should be a last resort because technology was constantly evolving, and new and unforeseen challenges may emerge. Notably, a number of representors, including foreign experts from different fields, Dr Farid and Mr Deynychenko, acknowledged this

[218] Google, Facebook, and Twitter said they would not comply with a request by the Government to take down a falsehood, unless backed by a legally valid and binding order.

challenge, but were of the view that strong measures still had to be tried and tested in an iterative process.

411. In particular, Mr Deynychenko emphasised that online falsehoods could be used against any country at any time very quickly; even as technology continued to advance and adversaries continued to adapt, action had to be taken. Dr Shashi agreed that different methods had to be tried, despite the uncertainty that any particular approach would be a silver bullet. In that regard, Germany was reportedly prepared to make improvements to its 2017 Network Enforcement Law, which strengthens measures against illegal online content.

(5) Freedom of expression

412. On the issue of freedom of speech, the Committee makes the following observations:

a. Measures to combat online falsehoods do not necessarily lie in opposition to freedom of speech. In fact, both serve the same ideals.

i. Online falsehoods harm democracy and the genuine contestation of ideas in the "marketplace"; the latter is what the freedom of speech serves to protect. Actions to combat online falsehoods serve to protect these ideals.

ii. No representor gave any convincing reason why falsehoods should be protected by the right to freedom of speech.

b. With regard to concerns that freedom of speech may be affected by countermeasures that are overly broad:

i. This can be addressed by adopting a calibrated approach, as the Committee recommends.

ii. Further, falsehoods are capable of being defined. The law has historically done so, and the Courts regularly do so. Falsehoods concern provable facts, and not opinions, philosophical notions of truth, or moral notions of right and wrong.

c. French expert Dr Limonier explained how in Europe, misinformation campaigns found success partly because attempts to tackle them were labelled a denial of democracy. Czech expert Mr Janda stated that traditional liberal-democratic ideals, such as free speech, critical

journalism, and independent thought have been used by foreign disinformation agents as a shield for their disinformation.[219]

d. The 2018 Reuters Digital Institute Digital News Report found that there is generally public appetite, both globally and locally, for some form of government intervention to stop the spread of false information on the Internet.[220] 61% of respondents across 23 countries, including Singapore, were of the view that the government should do more to separate what is real and what is fake on the Internet. In Singapore, 63% of respondents endorsed this view.[221] This stands in stark contrast to the "hands off" approach advocated by a minority of representors before the Committee.

413. The Committee discussed the above issues extensively with representors from different backgrounds. The Committee heard a spectrum of views. On one hand, there was scholarly evidence about how free speech does not extend to the deliberate spread of falsehoods. On the other, there were views that there should not be any legislative restrictions on expression, even if demonstrably false and harmful, except as a last resort.

414. The Committee considered the varied views and arguments put forward on the following issues:

a. whether the right to freedom of speech protects falsehoods;
b. whether free speech would be curtailed;
c. whether legal action would have a "chilling effect" on speech; and
d. whether legal action would undermine critical thinking.

415. *Whether the right to freedom of speech protects falsehoods.* Several representors, drawing on their background in law, put forward arguments for why falsehoods should not be protected by the right to freedom of speech.

416. German political data scientists Dr Hegelich and Mr Shahrezaye provided a useful framework for understanding the trade-offs between freedoms and the public interest:

a. In society, there is a public sphere and a private sphere. In the private sphere, we can share personal messages, and we generally have the right to say what we want.

[219] Monika Richter, "The Kremlin's Platform for 'Useful Idiots' in the West: An Overview of RT's Editorial Strategy and Evidence of Impact", *European Values Think-Tank* (18 September 2017), p 3.

[220] Nic Newman et al, "Reuters Institute Digital News Report 2018", *Reuters Institute for the Study of Journalism, University of Oxford*, p 10.

[221] Nic Newman et al, "Reuters Institute Digital News Report 2018", *Reuters Institute for the Study of Journalism, University of Oxford*, p 40.

b. In the public sphere, contradictory interests must be integrated. We do not have the automatic right to distribute any message we want. There will always be a trade-off between personal freedom and public interest.

417. Constitutional law professor Dr Thio elaborated on how the balance between freedoms and the public interest is made:

a. In the public sphere, not all speech is worthy of protection. Article 14(2)(a) of the Singapore Constitution balances the right to free speech with other competing interests such as the security of Singapore, public order, and incitement to any offence.

b. Society must ask what purpose the speech serves. Societies differ as to what speech is worthy of protection.

c. Speech that violates the rights of others or undermines a social interest, or both, is not worthy of protection.

d. An important purpose of speech is its key role in democratic society, so that we can have free and open political debate. This is because democracy depends on members of society being informed, not misinformed. Citizens have an interest in receiving information that will enable them to understand public affairs and make informed choices in electing their representatives.

418. The Committee found useful the arguments made by Dr Thio and other representors for why online falsehoods harm democracy, and are not worthy of protection, which are as follows:

a. Online falsehoods harm the earnest search for truth, and cause confusion. They crowd out reliable news and facts. They divert attention from substantive issues.

b. They damage the trust and sense of solidarity and common identity among citizens.

c. They drown out other people, undermining their exercise of free speech.

d. They undermine the process by which citizens engage in public discourse. Their proliferation may discourage people from engaging in civic life.

e. They polarise and divide. They undermine the public square and common domain for citizen interaction that is important for public debate. They undermine the "marketplace of ideas".

419. Dr Thio also pointed to the following observation of the UK House of Lords,[222] which has been cited with approval by the Singapore Court of Appeal:[223]

> "[It] is important always to remember that it is the communication of information not misinformation which is the subject of this liberty. There is no human right to disseminate information that is not true. No public interest is served by publishing or communicating misinformation. The working of a democratic society depends on the members of that society … being informed not misinformed. Misleading people and … purveying as facts statements which are not true is destructive of the democratic society and should form no part of such a society. There is no duty to publish what is not true: there is no interest in being misinformed. These are general propositions going far beyond the mere protection of reputations."

420. *Whether free speech would be curtailed.* It was suggested that it will be difficult to adequately define what constitutes a "deliberate online falsehood", potentially leading to the unfair censorship of legitimate speech. In contrast, constitutional law professor Dr Thio stated that courts regularly have to determine whether a statement of fact was false. Associate Professor Chua also noted that falsehoods could be defined using objective and verifiable data.

421. Definitional issues have not stopped efforts to combat content with vaguer and more subjective definitions. Computer scientist Dr Farid recounted how there were initially doubts over how to define material depicting "child exploitation" or content that promoted violent extremism; still, there was a responsibility to do so.

422. The difficulty perceived by some representors may stem not from determining whether a statement of fact is false, but from determining when to intervene. An example of a potential "grey area" was the "Lisa case" in Germany, where a Russian girl had lied that she had been abducted and raped by men of Middle Eastern origin. The false rape claim was widely reported by media outlets from a foreign country, without mentioning that the German police had found the claim to be false. Mr Nimmo accepted that the rape claim was wholly false. However, he felt the foreign media reports were a "grey area" as the foreign media outlets had not themselves made the false claim; they were reporting what someone else had said, while omitting contrary facts. Mr Nimmo nevertheless agreed that requiring the reports of the false claim to carry a clarification of the facts was a nuanced measure that made sense, and that the "grey area" could be resolved by using different measures for different situations.

423. *Whether legal action would have a "chilling effect"*. This was the concern that laws targeted at false statements may deter more than false speech, as explained by the US Supreme Court in its decision in *New York Times Co v Sullivan.* This is "because of doubt whether [the truth] can be proved in court or fear of the expense

[222] *Reynolds v Times Newspaper* [2001] 2 AC 127.
[223] *Review Publishing Co Ltd and another v Lee Hsien Loong and another appeal* [2010] 1 SLR 52.

of having to do so."[224] The US Supreme Court held that liability for libel could nevertheless be imposed if the false statement was made knowingly or recklessly. The "chilling" concern appeared to relate to laws that imposed liability on a person, and it was unclear if it also related to laws that did not impose liability.

424. Mothership testified that they did not experience a drop in traffic, nor a drop in contributions, comments and engagement on its platform as a result of being covered by the Broadcasting Act licensing regime. This suggested the need for circumspection in assessing the extent of any potential "chilling effect". The prospect of a "chilling effect" should be dealt with through calibration in the powers deployed; the answer cannot be to do nothing at all.

425. *Whether legislation will undermine critical thinking.* Some were concerned that legislation targeting online falsehoods would undermine the ability to think critically. The premise of this view was that legal action risked being a substitute for the ability of people to think for themselves. A different view was given by other representors who regarded legal action as complementary to efforts in media literacy education. The Committee agrees that legislation is complementary to the fostering of critical thinking. Notably, the representors from Ukraine and the Czech Republic shared how their countries were countering foreign disinformation with a multi-pronged approach that combined legal tools to challenge the sources and content of disinformation with a push to improve media literacy and critical thinking.

(6) A hands-off approach is not tenable

426. The notion that contestation in the "free marketplace of ideas" will solve the problem is contradicted by the real and serious consequences that online falsehoods have had around the world. It was also discredited by in-depth analyses of the application of the "marketplace" theory in the real world.

427. *Critique of the "marketplace" theory.* Constitutional law professor Dr Thio, French lawyer Mr Shefet, and other representors critically questioned the "marketplace of ideas" theory. They made the following arguments to show that the theory was not applicable to falsehoods, especially in the digital age.

a. There is no such thing as a completely free "marketplace". Even the "real" marketplace is regulated through consumer protection, anti-trust, and financial regulations.

b. In practice, and as shown throughout history, falsehood frequently prevails over truth with deleterious societal consequences.

[224] *New York Times Co v Sullivan* 376 US 254 (1964).

c. Given that there is no public interest in being misinformed, falsehoods may not belong to the "marketplace" to begin with. As one representor put it, "just as the economic marketplace is negatively affected by the peddling of counterfeit goods, the proliferation of falsehoods [also] damage[s] democracy".[225]

d. The "marketplace of ideas" operates on several assumptions, namely –

i. People are rational. They can and will sift cogent arguments from dubious one. They will equally participate in the democratic process in search of the best approaches towards social problems.
ii. People have free and equal access to the "marketplace".
iii. A broad range of views is available in this "marketplace".
iv. There is authentic discussion, where views can be exchanged.

e. The above assumptions do not necessarily, always hold in today's digital world.

i. The "marketplace of ideas" theory does not accurately describe how people behave. Cognitive biases (heuristic tendencies) operate. People may also be not open to other points of view, because of ideology. These cognitive biases have become accentuated in the digital age, given the overload of information online.

ii. Not everyone has equal access to the "marketplace". The amplification of ideas can be falsified by bots. Some people, through either wealth or success, have greater influence than others. Further, in the context of hate speech, targeted groups may be outnumbered or suffering from historical disadvantages, such that they cannot hold their own in the "marketplace".

iii. The Internet does not ensure that people are exposed to a broad range of views. Algorithms on social media sites and search engines create "echo chambers" and "filter bubbles" that entrench people in ideological silos.

iv. Rather than authentic discussion, there is anonymity online.

v. With algorithms that promote popular content rather than accurate content, the "best" idea that emerges from the "marketplace" may not necessarily be the truth; it may simply be what is popular or commonly shared. What is viral becomes what is the best, but the virality of content should have no bearing on its validity.

[225] Darius Lee, Appendix III: Written Representations, Paper No. 32, page B105, para 5.

428. The Committee notes the views of contemporary US constitutional law expert Professor Noah Feldman, who argued against treating the "marketplace" metaphor as the basic rationale for free speech. He pointed out that the marketplace could fail, and that the classic solution to market failure was regulation.[226]

429. Leaving matters to the "marketplace" is also inconsistent with the evidence further above on the limitations of non-interventionist measures, such as education on media literacy and critical thinking, fact-checking, as well as the inadequacy of voluntary efforts by the industry.

(7) Voluntary action by technology companies is not enough

430. Voluntary efforts by technology companies are unlikely to be able to achieve the results needed. In summary:

a. The technology companies have a policy of not removing content on the basis that it is false. Neither do they shut down purveyors of false content on the basis that the content published is false.

b. While their alternative measures may help improve the overall situation, these measures are not able to swiftly deal with damaging online falsehoods.

c. The track record of the technology companies show that they have not always responded seriously or adequately to the harm that their platforms have contributed to, for example, hate speech in the UK and the state of emergency in Sri Lanka.

d. Fundamentally, there exists a conflict of interests between technology companies' willingness to undertake self-regulation to tackle the problems of online falsehoods and their goal of maximising commercial output. For these reasons, valid questions have been raised as to whether technology companies are best placed to make decisions in the public interest, to adjudicate on what is true or false, beneficial or harmful, for the rest of society, especially in societies which norms differ from the technology companies' own standards. These concerns remain even if technology companies are legally obliged to take on this responsibility (*e.g.* under Germany's Network Enforcement Act).

431. To elaborate, the following points about the approaches of these technology companies may be made:

[226] Noah Feldman, "Is fake news protected by the First Amendment", *Bloomberg View* (24 November 2016), p 2, citing *Schenk v United States* (1919) 249 US 47.

a. *Approach to false content.* Facebook and Google are generally using algorithms to detect content of dubious credibility. As they have a policy of not removing content on the basis that it is false, they instead demote the content in news feeds and search results. In important situations, such as elections, Facebook has hired human fact-checkers to flag specific false content, which will then be demoted in users' news feeds.

b. *Approach to sources of false content.* The major technology companies do not remove sources of false content on the basis that they are producing false content. Facebook Pages that are identified as propagating falsehoods may instead lose their ability to earn revenue, and have the visibility of their content reduced.[227]

432. The following evidence shows that while these measures go some way to tackling the problem, they are far from being an adequate response in the event of the spread of damaging online falsehoods.

a. A study of Facebook's latest change to its algorithms showed that while the engagement rates of many long-established click-bait sites seem to have reduced, newer or less well-known ones have seen thriving engagement with one-off viral false stories.[228] This suggests that algorithms may be over- or under-inclusive in their detection of false content, and are not fool-proof. During the hearing, a Facebook representative maintained that relying on algorithms was the right approach. He nevertheless acknowledged that it still had to be tested, and how well it worked was still not known.

b. Facebook has hired human fact-checkers for Mexico's 2018 Presidential Election. There was reportedly still a "sea of misinformation" on platforms such as Facebook.[229] Further, several Facebook pages identified as spreading fake news remained on Facebook with large followings in the hundreds of thousands and millions.[230]

c. During the Las Vegas shooting in October 2017, divisive hoaxes appeared in Google's top search results. Google used algorithms to demote these hoaxes instead of directly removing them, which meant that despite the ongoing public alarm, it took hours for their visibility to be reduced.[231]

---

[227] Julia Love et al, "In Mexico, fake news creators up their game ahead of election", *Reuters* (29 June 2018); Craig Silverman, "Facebook is about to bring the hammer down on overseas fake news operators", *BuzzFeed* (21 June 2018).
[228] Liam Corcoran, "The most engaged sites on Facebook in April 2018", *NewsWhip* (7 May 2018); Paris Martineau, "Facebook's fake news algorithm seems to be working", *The Future* (10 May 2018).
[229] Ioan Grillo, "Fake News Crosses the Rio Grande", *New York Times* (3 May 2018).
[230] Julia Love et al, "In Mexico, fake news creators up their game ahead of election", *Reuters* (29 June 2018).
[231] Richard Waters, "Facebook and Google help showcase Las Vegas fake news", *Financial Times* (3 October 2017).

d. While Facebook generally has a policy of not removing content that is false, Facebook CEO Mark Zuckerberg stated in July 2018 that Facebook would remove false information that could result in physical harm to people. However, according to Mr Zuckerberg, content such as the denial of the occurrence of the Holocaust and conspiracy theories that falsely claimed school shootings did not happen would still not be removed under Facebook's new policy, unless it amounted to attacking individuals.[232]

e. Facebook's new policy has been criticised as still being inadequate. Some have observed that those who post Holocaust-denial stories online do so with the intent to defame and target Jews, which can amount to causing imminent harm.[233] Parents of victims of school shootings commented that the policy of not removing conspiracy theories claiming that the school shooting did not occur, and instead only demoting the content in news feeds and search results, provided no protection to them at all. They noted that since few people would write about a school shooting which occurred a number of years ago, only the false information appears and is spread, giving increased credence to hateful and dangerous content.[234]

433. <u>*Global track record*</u>. The Committee considered evidence relating to the following aspects of the major technology companies' track records:

a. Their policies and actions in response to problematic content on their platforms that could cause serious harm; and
b. Their attitudes towards their responsibility for negative impacts of their business on society.

434. *Responses to harmful content on their platforms.* The technology companies have faced heavy criticisms for failing to act against harmful content on their platforms. Their failures were attributed to passivity, inadequate due diligence, as well as lax policy standards. Some of these criticisms, including from other governments, are set out below:

United Kingdom

a. Google placed advertisements by UK advertisers on extremist YouTube videos created by supporters of terrorist groups such as ISIS. This enabled these groups to generate revenue from their YouTube sites, at the expense of the UK advertisers. The UK Home Affairs Committee, in its inquiry into *Hate crime: abuse, hate and extremism online*, described it as "shocking that Google failed to perform basic due diligence" to prevent the placement of the online advertisements. This was, to them, "a

[232] Kara Swisher, "Zuckerberg: The Recode Interview", *Recode* (18 July 2018).
[233] Emily Dreyfuss, "Facebook's Fight against Fake News Keeps Raising Questions", *WIRED* (20 July 2018).
[234] Leonard Ponzer and Veronique De La Rosa, "An Open Letter to Mark Zuckerberg", *The Guardian* (25 July 2018).

reflection of the *laissez-faire* approach that many social media companies have taken to moderating extremist content on their platforms."[235]

b. Despite repeated requests over eight months by the UK Home Affairs Committee, YouTube (which is owned by Google) failed to remove YouTube videos that promoted extreme neo-Nazi groups that were proscribed in the UK. This was so even though YouTube accepted that the videos were illegal. The UK Home Affairs Committee found that "[t]he weakness and delays in Google's response to our reports of illegal neo-Nazi propaganda on YouTube were dreadful."[236] The Committee noted how the technology companies were able to swiftly implement technology to remove material that breached copyright, but were in contrast slower to stop the sharing of violent extremist material.[237] They also concluded that the technology companies were "shamefully far from taking sufficient action to tackle illegal and dangerous content, to implement proper community standards or to keep their users safe."[238]

c. The UK Home Affairs Committee also described how Twitter refused to remove a cartoon that the Committee had reported, depicting a group of male, ethnic minority migrants tying up and abusing a semi-naked white woman, while stabbing her baby to death. The cartoon was published with a hashtag #DeportAllMuslims. Although the cartoon has since been removed, the UK Home Affairs Committee observed in its report then that Twitter had refused to take action on the grounds "that it was not in breach of [Twitter's] hateful conduct policy".[239]

d. The policies of technology companies did not prohibit anti-Semitic and Islamophobic content intended to stir up hatred against ethnic minorities.[240] This drew harsh criticism from the UK Home Affairs Committee, which stated: "The biggest companies have been repeatedly urged by Governments, police forces, community leaders and the public, to clean up their act, and to respond quickly and proactively to identify and remove illegal content. They have repeatedly failed to do so. That should not be accepted any longer. Social media is too important to

[235] "Hate crime: abuse, hate and extremism online", *House of Commons Home Affairs Committee* (1 May 2017), para 24.
[236] "Hate crime: abuse, hate and extremism online", *House of Commons Home Affairs Committee* (1 May 2017), para 30
[237] "Hate crime: abuse, hate and extremism online", *House of Commons Home Affairs Committee* (1 May 2017), para 30
[238] "Hate crime: abuse, hate and extremism online", *House of Commons Home Affairs Committee* (1 May 2017), para 25.
[239] "Hate crime: abuse, hate and extremism online", *House of Commons Home Affairs Committee* (1 May 2017), para 12.
[240] "Hate crime: abuse, hate and extremism online", *House of Commons Home Affairs Committee* (1 May 2017), paras 10 – 13.

everyone … to continue with such a lax approach to dangerous content that can wreck lives."[241]

e. The UK DCMS Committee also reported that on the problem of bots and fake accounts, Mr Mike Schroepfer (Chief Technology Officer, Facebook) had acknowledged the scale of this problem on Facebook's platform, but was "evasive about how many fake accounts had been removed" by Facebook.[242]

Sri Lanka

f. Due to fatal anti-Muslim riots in Sri Lanka in March 2018, the Sri Lankan government blocked access to Facebook, WhatsApp and Instagram in an emergency measure to stop the violence. The Sri Lankan government criticised Facebook for failing to prevent its platforms, including WhatsApp and Instagram, from being used to spread hate speech, which had reportedly been fomenting since 2013.[243] It noted that Facebook had taken days to review flagged posts and take down pages. It also highlighted a claim that a highly inflammatory Facebook post calling for the killing of Muslims and using degrading terms was found by Facebook to not breach its community standards.[244]

United States

g. Twitter took 11 months to shut down a fake troll account (later alleged to be linked to a foreign country) that impersonated the Tennessee Republican Party.[245] This was despite repeated requests from the real party. During that time, the fake account gained over 150,000 followers. Further details on what the fake Tennessee Republican Party account had done can be found in **Annex E**.

h. In September 2018, Campaign for Accountability (CfA) published a report,[246] detailing how it was able to buy advertisements on Google's Russian advertising platform targeting US internet users. Throughout the

[241] "Hate crime: abuse, hate and extremism online", *House of Commons Home Affairs Committee* (1 May 2017), para 36.
[242] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 69.
[243] Vindu Goel et al, "In Sri Lanka, Facebook contends with shutdown after mob violence", *The New York Times* (8 March 2018).
[244] Michael Safi, "Sri Lanka accuses Facebook over hate speech after deadly riots", *The Guardian* (14 March 2018).
[245] Kevin Collier, "Twitter was warned repeatedly about this fake account run by a Russian troll farm and refused to take it down", *BuzzFeed* (18 October 2017).
[246] "How to Sow Discord Using Google and $100 (or RUB 6,800)", *Campaign for Accountability, Google Transparency Project* (September 2018).

process, CfA waved "obvious red flags"[247] in an effort to trigger Google's safeguards. CfA used a Russian IP address to access Google's Russian advertising platform, supplied the details of the IRA (a Russian-linked troll farm as described at [206] above) to set up its account, submitted images previously identified to be created by the IRA, and even paid for the advertisements in Russian currency through Russia's largest electronic payment service. Google made no attempts to verify the identity of the account, and approved the advertisements in as few as 24 hours. The advertisements ran on a wide range of websites and YouTube channels, including CNN, The Daily Beast, Huffington Post and the UK's Daily Mail, generating over 20,000 views and more than 200 clicks. According to the report, CfA achieved all this with less than US$100. As was also pointed out in the CfA report, CfA managed to run this successful campaign despite Google stating, in the aftermath of the 2016 US Presidential Election, that "[Google has] a set of strict ads policies including limits on political ad targeting"[248] and despite Google recently stating in August 2018 that it has "invested in robust systems to … identify influence operations launched by foreign governments".[249]

Libya

i. In Libya, members of armed groups use Facebook to boast of their battlefield exploits and rally supporters by sowing division and ethnic hatred. The New York Times also found evidence of military-grade weapons being openly traded on Facebook, where there are pages containing advertisements for machine guns, anti-aircraft guns and artillery shells. It was reported that every armed group in Libya has their own Facebook page. Human traffickers also advertise their success in helping illegal migrants reach Europe, and use their Facebook pages to drum up more business. All of these are happening on Facebook, despite Facebook insisting that it assiduously polices its platform in Libya, it implements policies that prohibit the trading of firearms between individuals, and it "[does not] allow organizations or individuals engaged in human trafficking or organized violence to maintain a presence on Facebook".[250]

Germany

j. According to the German Ministry of Justice, the technology companies fell significantly short of meeting their commitments to remove illegal

[247] "How to Sow Discord Using Google and $100 (or RUB 6,800)", *Campaign for Accountability, Google Transparency Project* (September 2018), p 2.
[248] Elizabeth Dwoskin et al, "Google uncovers Russian-bought ads on YouTube, Gmail and other platforms", *The Washington Post* (9 October 2017).
[249] Kent Walker, "An update on state-sponsored activity", *Google, Safety & Security* (23 August 2018).
[250] Declan Walsh and Suliman Ali Zway, "A Facebook War: Libyans Battle on Streets and on Screens", *New York Times* (4 September 2018).

content, such as hate speech, within 24 hours of being notified.[251] This was one of the main reasons for Germany's enactment of its Network Enforcement Act.[252]

Others

k. The Counter Extremism Project, which Dr Farid worked with, had found that extremist material would stay online for "hours, days and in some cases weeks, gathering thousands and tens of thousands of views."[253] The Counter Extremism Project has publicly declared that "[w]hile big social media platforms acknowledge the existence of radicalising content that violates their stated terms of service, their response to date has followed a familiar pattern utilised in response to other discoveries of abuse, denial, followed by half measures and attempts to spin the issue in the media, and finally, reluctant action when faced with threats to their bottom line or possible regulatory action."[254]

435. *Attitudes towards their responsibility to society.* The Committee heard evidence on the attitudes of some of the major technology companies towards the negative impact their businesses have had on society.

Response to extremist content

a. Dr Farid, Senior Advisor to the Counter Extremism Project in the US, described how the major US-based technology companies had dragged their feet for several years instead of developing or deploying any effective solution to disrupt the global distribution of child pornography, despite being prompted by the then-US Attorney General to do so. They adopted the same attitude when called on to tackle online extremism by government agencies internationally in around 2014. Effective responses began only around 3 years later, after pressure from the EU and US. Dr Farid caveated that his views applied to some of the technology giants more than others.

Response to the problem of "fake news"

b. In November 2016, Facebook CEO Mr Mark Zuckerberg described the notion that fake news had influenced the 2016 US Presidential Election as "a pretty crazy idea", and blamed the problem on users' own personal bias.[255] Facebook's position shifted after evidence of the manipulation of

[251] Explanatory Note by the German Ministry of Justice to the Network Enforcement Act dated 27 March 2017.
[252] Explanatory Note by the German Ministry of Justice to the Network Enforcement Act dated 27 March 2017.
[253] Hany Farid, Appendix IV: Minutes of Evidence, page C633, para 5368.
[254] Statement of the Counter Extremism Project titled "Senate Committee on Commerce, Science and Transportation" (17 January 2018), p 2.
[255] Adrienne Jane Burke, "Facebook Influenced Election? Crazy Idea, Says Zuckerberg", *Techonomy* (11 November 2016).

Facebook users by foreign State-linked actors came to light. In a statement by Mr Zuckerberg in September 2017, he noted that he "care[d] deeply about the democratic process and protecting its integrity" and that "[Facebook was] actively working with the US government on its ongoing investigations into Russian interference [in the US elections]".[256]

c. There have been doubts in the US over whether the technology giants have done a full accounting of Russian interference on their platforms in the 2016 US Presidential Election. In particular, US Senator Mark Warner criticised Twitter for failing to conduct a thorough search of its platform for signs of Russian interference. Twitter had initially informed the US Senate Committee that it had identified only around 200 Russian state-backed accounts.[257] Tellingly, after pressure from US senators, they subsequently raised the number to over 2,000 accounts a month later. Just another few months later, Twitter admitted that 3,814 accounts were being actively managed by Russian state operatives, and 50,258 automated accounts linked to Russia tweeted election-related content during the election period.[258] In January 2018, Facebook also admitted that they were "too slow" to recognise Russian election interference.[259] During the hearing before the Committee, Facebook stated that they were still continuing investigations into the extent of Russian interference on their platform.

d. In the UK, the Chair of the UK DCMS Committee had also stated that "Facebook continues to display a pattern of evasive behaviour – a pattern which has emerged over the course of our inquiry… [t]he company appears to prefer minimal over rigorous scrutiny…"[260] The UK Committee Interim Report also highlighted that "Facebook [had] consistently responded to questions by giving the minimal amount of information possible, and routinely failed to offer information relevant to the inquiry, unless it had been expressly asked for. It provided witnesses who have been unwilling or unable to give full answers to the Committee's questions."[261]

[256] Mark Zuckerberg, Facebook Post (September 2017).
[257] Tony Romm, "A top Democratic senator briefed by Twitter on Russia and the 2016 election called the company's explanation 'frankly inadequate'", *Recode*, (28 September 2017).
[258] Jon Swaine, "Twitter admits far more Russian bots posted on election than it had disclosed", *The Guardian* (20 January 2018).
[259] Alex Hern, "Facebook: we were too slow to recognise our 'corrosive' effect on democracy", *The Guardian* (22 January 2018).
[260] "Facebook still evasive over Cambridge Analytica and fake news: UK lawmakers", *Reuters* (29 June 2018).
[261] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 61.

Response to Facebook's role in the violence in Myanmar

e. In March 2018, the United Nations' Fact-finding Mission on Myanmar announced its interim findings of its investigations, and observed that Facebook had "played a determining role in stirring up hatred against Rohingya Muslims in Myanmar". The chairman of the Mission, Mr Marzuki Darusman, pointed out that Facebook had "substantively contributed to the level of acrimony" amongst the wider public against Rohingya Muslims.[262] The UK DCMS Committee subsequently highlighted this incident to Mr Mike Schroepfer, the Chief Technology Officer at Facebook, and found his response unsatisfactory. Despite Mr Schroepfer describing the situation in Myanmar as "awful", and promising to do more, the UK DCMS Committee pointed out that Mr Schroepfer was unable to provide evidence as to how many fake accounts had been identified and removed from Myanmar, or the amount of revenue Facebook was making from Facebook users in Myanmar.[263]

Responsibility for data privacy – Cambridge Analytica

f. The Committee questioned Facebook on the recent data privacy scandal it was involved in with Cambridge Analytica, a British political consulting firm that used data mining to help political candidates target potential voters. The scandal concerned the access of Facebook user data by an application developer, one Dr Kogan, who had then provided the data to Cambridge Analytica.

g. Facebook knew about the breach in 2015, but did not inform any of those whose data had been illegitimately accessed. It obtained a legal certification from Dr Kogan and Cambridge Analytica that all the illegitimately accessed data had been deleted.

h. In February 2018, Mr Simon Milner, now Facebook's Vice President for Public Policy, Asia Pacific, was questioned by the UK DCMS Committee about whether Facebook had provided any Facebook user data to Cambridge Analytica, and whether Cambridge Analytica held Facebook user data. Notably, Mr Milner did not disclose the data breach committed by Dr Kogan and Cambridge Analytica, even though he knew of it at the time. In the UK Committee Interim Report, the UK DCMS Committee found that Facebook had failed to disclose the existence of this "breach of trust" and its implications.[264]

---

[262] "UN: Facebook has turned into a beast in Myanmar", *BBC* (13 March 2018).
[263] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 79.
[264] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 114.

i. In March 2018, Facebook issued a public statement about the violation of its data policies by Cambridge Analytica and Dr Kogan. This was done after it came to light that Cambridge Analytica had retained the data instead of deleting it. Facebook's CEO Mr Mark Zuckerberg accepted that there had been "a major breach of trust" between Facebook and its users, who expected Facebook to protect their data.[265]

j. However, Facebook then played down the breach. It claimed that Dr Kogan had legitimately gained access to the data, as users who chose to sign up to his application had given their consent. This explanation, however, failed to address whether users had given their consent to their data being passed on to a third party, *i.e.* Cambridge Analytica. It also failed to address how such "consent" applied to persons who were friends of those who had downloaded the application, but had not themselves downloaded the application and had therefore not agreed to the terms of the application.

k. The Committee questioned Mr Milner about the Cambridge Analytica scandal during the hearing. Mr Milner conceded that he should have given a fuller answer to the UK DCMS Committee in February 2018, when asked about Cambridge Analytica. He agreed that a reasonable person could take the view that he had not been full and frank in his answers, and had misled the UK DCMS Committee. He further said that, as he was not involved in those decisions, he could not explain (a) why Facebook did not notify the 50 million affected users in 2015, and (b) why Facebook did not take further steps beyond requiring a legal certification to ensure that the data had been deleted by Cambridge Analytica, even after having been lied to by them.

436. The above evidence supports the view of experts and observers cited by Dr Soon and Mr Goh, who have "expressed doubt that the technology giants would have the will and incentive to self-regulate effectively simply because they are monopolies (or near-monopolies)."[266]

437. <u>*Conflict of interests*</u>. The business imperatives of the online platforms may mean they are fundamentally not positioned to take the actions necessary to fully deal with the problem. They would not take measures that would risk their popularity with users.

a. For example, the largest source of traffic to fake news channels on YouTube reportedly comes from YouTube's "up next" function. This function tends to promote false news channels because it is influenced by user interest. According to a Google representative in a hearing before the UK DCMS Committee, if someone is expressing an interest in a particular

[265] Mark Zuckerberg Facebook post (22 March 2018).
[266] Carol Soon and Shawn Goh, Appendix III: Written Representations, Paper No. 62, page B374, para 36.

type of content, it would be hard for YouTube to insert something opposite to their interests as this would lead to an abandonment of YouTube's service.

b. The UK DCMS Committee also subsequently observed, in the UK Committee Interim Report, that the "business models [of online platforms] rely on revenue coming from the sale of adverts and, because the bottom line is profit, negative emotions (which appear more quickly than positive emotions) will always be prioritised. This makes it possible for negative stories to spread."[267]

438. The Committee notes the Asia Internet Coalition's comment that the European Commission (EC) has decided against legislation in favour of voluntary regulation by technology companies. This is, however, inaccurate. In the document cited by the Asia Internet Coalition, the EC had not made any final recommendation on the issue.[268] Although the High Level Group on fake news and online disinformation did recommend voluntary regulation as a first step, it also recommended that the EC review the effectiveness of voluntary regulation, and consider appropriate regulatory responses "in order to ensure that the actions recommended … are effectively implemented". In doing so, it acknowledged that "the willingness of all parties to adhere to [a voluntary] approach remains to be proven", and "consistent implementation across the whole EU may represent a challenge for all players concerned".[269]

Recommendation 12. The Government should have the powers to swiftly disrupt the spread and influence of online falsehoods.

The objectives to be achieved should be as follows:

a. Provide access to and increase the visibility of corrections, including through tagging functions and use of other platforms with significant reach.
b. Limit or block exposure to the online falsehood.
c. Disrupt the digital amplification of online falsehoods, including through the use of false amplifiers (*e.g.* inauthentic accounts run by bots or trolls), and digital advertising tools.
d. Discredit the sources of online falsehoods.

These capabilities should be able to apply to all relevant platforms regardless of their technological basis. There needs to be careful balance and calibration to prevent the public interest from being harmed, and to at the same time respect

[267] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 70.
[268] European Commission, "Roadmap on fake news and online disinformation" (9 November 2017).
[269] European Commission, "A multi-dimensional approach to disinformation, Report of the independent high level group on fake news and online disinformation" (March 2018), p 31.

communications that are personal, private, and of limited circulation. They should therefore apply both to open and closed platforms.

Legislation will be needed to achieve the above objectives. Such legislation should have the following objectives:

a. The measures will need to achieve the objective of breaking virality by being effective in a matter of hours.
b. The decision-maker should be effective and credible.
c. There should be adequate safeguards in place to ensure due process and the proper exercise of power, and give assurance to the public of the integrity of the decision-making process.
d. The measures should be deployed in a calibrated manner, taking into account the context and circumstances, including potential impact and reach.

Measures provided in the legislation could include: tagging of corrections and notifications, take-down powers and access-blocking, among other measures. This should include judicial oversight where appropriate.

Recommendation 13. The Government should identify the additional measures needed to safeguard election integrity, and implement the necessary measures, including legislation, in view of the issues that have been highlighted in this report.

Recommendation 14. The Government should consider implementing monitoring and early warning mechanisms, to facilitate assessments of when and how to intervene to stop the spread of online falsehoods.

Addressing the provenance of the problem is necessary. The Committee is supportive of measures to ensure deterrence and accountability of perpetrators of deliberate online falsehoods. This include ensuring that digital advertising platforms or digital advertisers are not supporting purveyors of online falsehoods; and imposing punitive measures on the perpetrators of deliberate online falsehoods.

The Committee is not calling for the criminalisation of all online falsehoods. Consistent with the calibrated approach the Committee has recommended, criminal sanctions should be used only against purveyors of online falsehoods that meet a prescribed threshold.

Recommendation 15. The Government should consider powers needed to establish a de-monetisation regime, including through legislation which will:

a. Disrupt the flows of digital advertising revenue to purveyors of online falsehoods. This should take into account the responsibility of different stakeholders in the digital advertising ecosystem.

b. Require the disgorgement of financial benefits by purveyors of online falsehoods. This should cover the "hired guns" who are paid by others to create and spread online falsehoods.

Recommendation 16. Criminal sanctions should be imposed on perpetrators of deliberate online falsehoods. These deterrent measures should be applied only in circumstances that meet certain criteria. There should be the requisite degree of criminal culpability (*i.e.* intent or knowledge), in accordance with established criminal justice principles. There should be a threshold of serious harm such as election interference, public disorder, and the erosion of trust in public institutions.

The Government should ensure these deterrent measures are adequate in scope to cover the range of methods and actors, including the deliberate use of inauthentic accounts or bots, the provision of tools and services to publish falsehoods, and the masterminds behind online falsehoods, who may not always be the ones creating or spreading them.

Whether existing criminal sanctions are adequate to achieve the above should be considered.

b. Adapt online platforms

439. Online platforms, including social media platforms, have revolutionised societies. They have transformed how information is communicated in society, and how we engage with one another, and with society. They have brought greater freedom and benefit to people.

440. However, how online platforms are designed and function have also been integral to making the proliferation of online falsehoods the serious global problem it is today. In this section, the Committee discusses the proposals that sought to tackle this issue.

(i) Rationale and Context

441. How online platforms are designed has a huge influence on society. Online platforms have "an oversize[d] presence in determining how we engage with information and with one another", as Professor Lim Sun Sun put it.[270]

[270] Lim Sun Sun, Appendix III: Written Representations, Paper No. 110, page B1039.

442. The Committee considered evidence that showed three ways in which online platforms do so. *First*, they are said to have negatively impacted the information that enters public discourse. This has been explained above at [184]-[185].

443. *Second*, their design influences human behaviour.

a. Facebook itself has made the claim that its "audience-specific" advertisements can shift the intent of voters.[271] Ms Samantha Bradshaw, from the Oxford Internet Institute, also gave evidence before the UK DCMS Committee on the power of Facebook to "manipulate people's emotions by showing different types of stories to them".[272]

b. With regard to search engines, experiments have shown that nearly undetectable changes to search engine rankings can influence the choice of political candidates by undecided voters. This has been termed the "search engine manipulation effect".[273]

444. It is hence of concern that Google search can present contradictory search information to different users. For example, in January 2018, Google admitted that its algorithms resulted in people getting contradictory information from Google's "featured snippets" when asking about the same thing in different ways.[274] To illustrate, people who searched for "are reptiles good pets" received featured snippets that contradicted those received by people who searched for "are reptiles bad pets". Google has stated that it is exploring solutions to this issue.

445. *Third*, they have had a direct impact on the dissemination of both online falsehoods and quality journalism. An underlying reason for this is that the algorithms of online platforms are designed primarily to maximise user engagement, rather than content of high quality.

a. Even without manipulation, the algorithmic design of online platforms independently plays a role in the distribution of online falsehoods. As mentioned above, YouTube's "up next" function was reportedly a key driver of traffic to fake news YouTube channels. This was also demonstrated by a recent trial by Facebook to change its News Feed in six countries, as described below:

[271] Adam Pasick, "Facebook says it can sway elections after all – for a price," *Quartz* (2 March 2017); Olivia Solon, "Facebook says likely Russia-based group paid for political ads during US election", *The Guardian* (7 September 2017).
[272] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 70.
[273] Robert Epstein and Ronald Robertson, "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections", *Proceedings of the National Academy of Sciences* (4 August 2015).
[274] Danny Sullivan, "A reintroduction to Google's featured snippets", *Google Blog* (30 January 2018).

i. In the trial, Facebook removed news content from the Facebook News Feed, and placed it in a new feed called the Explore Feed. There were hence two different News Feeds: one that was a dedicated place with posts from friends and family, and another a dedicated place for posts from Pages.[275]

ii. According to Facebook, the change was a test response to feedback that users wanted to see more posts from friends and family. According to sceptics, Facebook's real goal was to increase users' time-on-site, to allow it to serve more advertisements in between content, in videos, and elsewhere, thereby earning more advertising revenue.[276]

iii. This change in design had an effect that was described as "downright Orwellian", with numerous news sites seeing traffic to their sites plummet due to the change. The countries reportedly "saw massive reductions in the amount of trusted news content shared on the site – with little corresponding reduction in low-quality, politically inflammatory memes that still spread like wildfire across the network."[277]

iv. During the hearing before the Committee, a Facebook representative acknowledged that the test had not worked, and stated that the trial had ended. He nevertheless described it as a genuine attempt to make the platform better.

b. The algorithms used by online platforms to disseminate information have been exploited to deliberately manipulate public opinion.

i. According to a study by the Oxford Internet Institute's Computational Propaganda Research Project, a powerful tool for manipulating public opinion on social media was junk news that was backed by automation through the platform operators' own dissemination algorithms, as well as bots.[278]

ii. These dissemination algorithms are likely designed to maximise engagement of users with the content on their platforms, according to computer scientist Dr Farid. Facebook described how posts on News Feed were ranked according to relevance to the user, which

[275] Adam Mosseri, "Ending the Explore Feed Test", *Facebook Newsroom* (1 March 2018).
[276] Jonathan Shieber, "Facebook ends its experiment with the alternative 'Explore' news feed", *TechCrunch*, (2 March 2018).
[277] Alex Hern, "We were too slow to recognise our corrosive effect on democracy", *The Guardian* (22 January 2018).
[278] Samuel Wooley and Philip Howard, "Computation Propaganda Worldwide: Executive Summary", *Computational Propaganda Research Project, Oxford Internet Institute, University of Oxford*, p 8. The study also found that there were positive ways algorithms and bots generally could be used to serve the public interest.

was in turn determined by signals of what content the user engaged with. This is consistent with Dr Farid's evidence, as increasing relevance to the user would in turn increase engagement.

iii. There was the argument that users, not algorithms, are responsible for determining the standards of what is engagement-worthy. A Facebook representative described Facebook's algorithm as amplifying human intent, both positive and bad. However, the counter-argument was that these algorithms have been specifically designed to maximise engagement, "because engagement is part of [Facebook's] business model and their growth strategy".[279] Moreover, this also means that posts which "tap into negative, primal emotions like anger or fear ... perform best and so proliferate".[280] By maximising engagement instead of other outcomes, this has left users vulnerable to manipulation. This view was supported by Dr Farid.

c. Other design choices by online platforms have also been exploited to spread online falsehoods.

i. Twitter's application programming interface (API) is provided to developers who want to design Twitter-compatible applications and innovate using Twitter data.[281] While this has reaped creative uses and benefits, this has also led to the generation of a large amount of automated spam. Twitter has since announced that it would impose restrictions on how its API would be used.[282]

ii. An earlier iteration of Facebook's "self-service" targeted advertising model had been designed to allow for anonymous advertising.[283] As advertisements could be aimed at specific and narrowly-defined audiences, they could be easily hidden from public scrutiny (and hence were termed "dark ads"). These features were reportedly abused by foreign-linked disinformation agents to target advertisements at specific audiences.

446. Notably, Facebook's product manager, Mr Samidh Chakrabarti, acknowledged that "at its best, [social media] allows us to express ourselves and take action. At

[279] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 64.
[280] Amanda Taub and Max Fisher, "Facebook Fueled Anti-Refugee Attacks in Germany, New Research Suggests", *New York Times* (21 August 2018).
[281] Testimony of Sean J. Edgett, Acting General Counsel, Twitter, Inc., before the US Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism (31 October 2017), p 6.
[282] Karissa Bell, "Twitter wants to kill spam for good", *Mashable Asia* (25 July 2018).
[283] Donie O'Sullivan, "What Russian trolls could have bought for $100,000 on Facebook", *CNN* (7 September 2017); Seth Fiegerman and Dylan Byers, "Facebook's Russia Problem: What we know", *CNN* (7 September 2017); Siva Vaidhyanathan, "Facebook Wins, Democracy Loses", *New York Times* (8 September 2017).

its worst, it allows people to spread misinformation and corrode democracy." He candidly acknowledged that there was no guarantee that the positives would outweigh the negatives.[284]

(ii) Representors' Views and Recommendations

447. An overarching theme from the evidence was that online platforms needed to adapt and change to address the problem. Some, including the technology companies, proposed making specific adjustments to existing platforms and features. Others envisaged that more fundamental change may be needed.

(1) Accountability and regulation of online platforms

448. A key issue during the hearing concerned the accountability of online platforms for their impact on society. Several representors supported making online platforms more accountable for their social impact.

a. The online platforms today deliberately choose to design their algorithms to optimise engagement as opposed to other outcomes such as credibility of information. Scientists Dr Farid, Dr Hegelich and Mr Shahrezaye emphasised that ultimately, humans decided what outcomes the social media algorithms should optimise.

b. The online platforms play a role in the publication of media and news. Dr Hegelich and Mr Shahrezaye pointed out that the platforms' core business was the distribution of users' content. A group comprising a lawyer and SMU Law students described them as the "modern gatekeepers of information", who ought to assume responsibility for the content they hosted or published.[285] The Singapore Press Club was of the view that the online platforms had become media platforms, as they were publishers of content (although not the source), and derived revenue from that content.

449. Representors from the media and news industries expressed that online platforms enjoy lower standards of responsibility for the information spread on their platforms compared to traditional media companies. An example given was that despite the scope and impact of the distribution algorithms used by the online platforms, they were not required to account for those algorithms in any way.

450. Calls to level the playing field were hence made by the traditional media companies. It was clarified that this did not necessarily mean regulating the online platforms in the exact same way as the regulation of media companies, since both

[284] Alex Hern, "Facebook: we were too slow to recognise our 'corrosive' effect on democracy", *The Guardian* (22 January 2018).
[285] Sui Yi Siong et al., Appendix III: Written Representations, Paper No. 130, page B1141, para 32.

were operationally different. Neither did it mean tipping the playing field in favour of the traditional media companies.

451. During the hearing, the online platforms Facebook, Twitter, and Google gave qualified answers about the extent of their responsibility. The Facebook representative accepted that Facebook had a global responsibility to do what they could to prevent the abuse of their platform "in terms of undermining the integrity of elections."[286] The Google representative said that Google understood its role in the information ecosystem, particularly in "breaking news situations" and elections.[287] The Twitter representative acknowledged that there were real problems and crises, and that Twitter took its impact on society "very seriously."[288]

(2) Specific measures proposed

452. At the outset, the Committee acknowledges the ongoing measures being taken by the major technology companies, which are set out at **Annex F**. Recommendations made by other representors on the specific measures that should be implemented by or on technology companies are set out below.

453. *"Cleaner" and more "enlightened" products and business models*. To some representors, fundamental change was needed. This entailed finding new business models, and designing better products.

a. Dr Farid identified the current advertising revenue-driven business model of the major online platforms as an arguable cause of the proliferation of abusive and harmful content in the online world. He called on companies to offer products that relied on different business models that would lead to healthier online communities. He also called on governments to encourage the development of more honest, safer, more private and more secure infrastructures for our online world.

b. Professor Lim Sun Sun suggested that digital platforms could be more "carefully deliberated with their consequences more fully thought through."[289] She proposed that digital platforms adopt more "enlightened design", which incorporates principles from psychology, behavioural economics, and philosophy. The aim was to "undo the damage and polarization that fake news has inflicted through social media." One example, which was recommended by a few representors, was for technology companies to design and use algorithms that were driven by credibility more than user engagement.

[286] Facebook, Appendix IV: Minutes of Evidence, page C380, para 3399.
[287] Google, Appendix IV: Minutes of Evidence, page C309, para 3655.
[288] Twitter, Appendix IV: Minutes of Evidence, page C406, para 3642.
[289] Lim Sun Sun, Appendix III: Written Representations, Paper No. 110, page B1039.

454. Some interesting new products were mentioned by representors. For example, Mr Zhulkarnain referred to Userfeeds, a Warsaw-based startup that utilised blockchain tokens to provide an economic incentive to rank content well. Mr Nugroho from Mafindo suggested creating a hoax-free search engine, which listed only legitimate sites that had been registered with a media regulator.

455. The evidence of Mr Shefet, however, suggested that products run on a different business model and different principles may face high barriers. He observed that today's technology giants controlled the marketplace, and that the "economic, cultural and psychological power vested in the titans [has created] an entirely new economic paradigm."[290]

456. *Disclose information on algorithms*. Professor Lim Sun Sun and Mr Shahrezaye called for technology companies to be transparent about what their algorithms were designed to do. For Professor Lim Sun Sun, this information would be used to help users understand the information infrastructure around them, to facilitate their ability to think critically about the information they received online. For Mr Shahrezaye, doing so would ensure accountability of the technology companies.

457. *Disclosure of data for research*. Dr Wardle and Dr Bontcheva emphasised the need for greater investigation into the scale and nature of the problem. They called for greater disclosure of data by technology companies to enable researchers to do so, and highlighted that the lack of access to data was currently a serious impediment to monitoring the problem. As Dr Wardle put it, technology companies are "the only ones who can view the scale of the problem [and] without any external access to this data, there is no way to independently audit the scale of the problem, to understand how and when users have interacted with disinformation, and to understand how disinformation moves across platforms".[291] In relation to Facebook, the UK DCMS Committee has also observed that "Facebook has all of the information. Those outside of the company have none of it, unless Facebook chooses to release it".[292]

458. *Develop technology to automate the detection of online disinformation*. To tackle the extremely large volume of disinformation online, representors such as Dr Bontcheva and Dr Wardle recommended developing automated tools to detect online disinformation. Dr Bontcheva and Dr Farid both underscored that current technology was still far from being fully automated, and would require "human-in-the-loop" interventions. It was suggested by Dr Wardle that incentives could be provided for companies to develop such systems to detect online falsehoods.

459. *Prevent online advertising tools from being abused*. The online advertising tools offered by online platforms such as Facebook and Twitter were reportedly used

[290] Dan Shefet, Appendix III: Written Representations, Paper No. 75, page B453.
[291] Claire Wardle, Appendix III: Written Representations, Paper No. 94, page B926.
[292] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 64.

by foreign agents to spread disinformation to interfere in the 2016 US Presidential Election. Proposals were made to address the potential for abuse of these tools.

a. *Transparency.* A solution emphasised by several representors, including Dr Wardle, was to ensure disclosures about whether the content was paid for, and who paid for the advertisement. This is a measure found in a proposed Honest Ads Act in the US. In Dr Wardle's view, this is to ensure that digital and non-digital ads are both held to the "same standard of transparency about who paid for the advertisement".[293] Moreover, this transparency should be required of "all forms of digital advertising", given that disinformation campaigns do not just target political candidates or policy issues overtly, but also cultural issues that may not be obviously "political" on first sight.[294]

   During the hearing, a Facebook representative agreed with Dr Wardle's point that the lack of transparency around promoted content online made it easier for manipulation to occur. He explained that Facebook has taken steps to implement such disclosures in its online advertising tools. For political advertisements, Facebook is or will be requiring Page owners and administrators to provide verification details when engaging in such advertising, including through postal mail.

b. *Foreign currency payments.* During the hearing, the Committee drew Facebook's attention to a proposal raised by US legislators to ban foreign currency payments for political advertisements. A Facebook representative said they would be willing to consider such a proposal, and agreed that Facebook could take steps to ensure that people engaging in political advertising were based in the country concerned.

c. *Prohibit or regulate targeted advertising during elections*. As mentioned earlier, Mr Shefet proposed banning any use of micro-targeting research and techniques during the elections, while law academic Associate Professor Eugene Tan suggested requiring political candidates to disclose the amount spent on social media targeting during their campaigns.

460. *Strengthen user code of conduct or standards*. Various representors also called on technology companies to strengthen their user codes of conduct and community standards.

461. *Ensure authenticity of accounts*. A few suggested requiring that online platforms ensure that all accounts belong to authentic users.

462. Representors also suggested a spectrum of regulatory methods for holding technology companies accountable for finding better solutions. These comprised

[293] Claire Wardle, Appendix III: Written Representations, Paper No. 94, page B928.
[294] Claire Wardle, Appendix III: Written Representations, Paper No. 94, page B928.

(i) reporting requirements, (ii) independent auditing requirements, and (iii) binding regulation.

(iii) Observations and Recommendations

463. Technology companies have a social responsibility to contribute to a clean Internet information ecosystem. In such an ecosystem, users will be made properly aware of the sources behind the information they are exposed to; additionally, accurate and reliable information is prioritised, while false and harmful information cannot thrive. As the evidence suggests, technology companies are in control of the design of their platforms and products, through which they have profited greatly. It follows that they should bear responsibility for preventing their platforms and products from contributing to the creation and proliferation of online falsehoods, which can harm the public interest.

464. While the Committee accepts that technology companies have begun to undertake various initiatives to improve their platforms and products, the Committee agrees with a substantial number of representors that more can and should be done on the part of technology companies. The Committee also notes the 2018 Reuters Digital Institute Digital News Report finding that 71% of respondents across 23 countries, including Singapore, were of the view that technology companies need to do more to separate what is real from what is fake on the Internet.[295] In essence, there is a clear need for technology companies to increase their transparency and improve their accountability.

465. *Increasing transparency*. Transparency is critical to ensure that users of the products or platforms managed by technology companies are fully aware of whom they are exchanging information with online. The goal of such transparency is to educate users on the behaviour and intent of other content providers they encounter online, and reduce the opportunity for malicious actors to hide behind Internet anonymity to carry out abusive activities.

466. Users ought to be provided with sufficient information to know whether they are interacting with accounts belonging to and managed by a real person, or whether they are interacting with accounts run by a bot, or with an account where someone is impersonating another. Impersonation is often used to create an appearance of popularity, to increase the influence of the online falsehoods propagated by inauthentic accounts. This is not a trivial concern; inauthentic accounts operating on a broad scale have the potential to disseminate widespread disinformation.[296]

467. The Committee therefore agrees with representors who have submitted that online platforms should ensure that all accounts belong to authentic users. Technology

[295] Nic Newman et al, "Reuters Institute Digital News Report 2018", *Reuters Institute for the Study of Journalism, University of Oxford*, p 40.
[296] See also: Mark Warner, Potential Policy Proposals for Regulation of Social Media and Technology Firms (draft White Paper), p 8.

companies must implement a robust system to be able to authenticate whether there is a real, identifiable person behind every account. Even if certain bot activities are allowed on their platforms, technology companies should "establish clear marking systems and rules for bots and ensure [the] activities [of bots] cannot be confused with human interactions".[297]

468. Given how digital advertising has facilitated the creation and spread of online falsehoods, there should also be full disclosure on the sponsor identity, amounts spent, and targeting criteria of all forms of digital advertising on the platforms of technology companies.[298] In this regard, the UK Committee Interim Report has recommended that "paid-for political advertising data ... [should be made] publicly accessible, [to] identify the source, explaining who uploaded it, who sponsored it, and its country of origin".[299] There is also a need to ensure the "full disclosure of targeting used as part of advert transparency".[300] The Committee agrees with these recommendations, to enable users to critically assess the information they access online. The Committee also agrees with Dr Wardle that these requirements should apply equally to *all* forms of digital advertising, as false information with serious consequences can and have been peddled by advertisements which are not targeted at political candidates, or a particular election.

469. Transparency is also a necessary precursor to enable technology companies and independent experts to conduct analyses and expose the instances or trends of malicious agents using sophisticated modalities to spread harmful content online, which may otherwise escape the eyes of ordinary users. This requires, as recommended by various representors before the Committee, the appropriate disclosure of the technology companies' platform data, including how its algorithms are designed to operate on its platforms or products.[301]

470. *Improving accountability*. It is imperative for technology companies to accept and acknowledge that they *do* have the capability, and responsibility, to ensure that

---

[297] European Commission Communication, "Tackling online disinformation: a European Approach", *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions* (26 April 2018), p 8; see also: Mark Warner, Potential Policy Proposals for Regulation of Social Media and Technology Firms (draft White Paper), p 6.

[298] European Commission Communication, "Tackling online disinformation: a European Approach", *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions* (26 April 2018), pp 7-8.

[299] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 144.

[300] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 145.

[301] European Commission Communication, "Tackling online disinformation: a European Approach", *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions* (26 April 2018), p 8; and Mark Warner, Potential Policy Proposals for Regulation of Social Media and Technology Firms (draft White Paper), p 10. In July 2018, it was announced that Facebook will partner with a group of academics to establish an independent research commission, where academics will have privacy-preserving access to Facebook data in order to look into "the effects of social media on democracy and elections", see: Richard Nieva, "Social Science One group will study Facebook's effect on elections", *CNET* (11 July 2018).

their platforms or products are not open to abuse and to take proactive steps to respond accordingly. As the UK Committee Interim Report correctly observes, technology companies "cannot hide behind the claim of being merely a 'platform' ... [as] they continually change what is and is not seen on their sites, based on algorithms and human intervention".[302] At a minimum, technology companies should undertake regular voluntary reporting and produce audit reports on the problem of deliberate online falsehoods on their platforms, to be directly accountable to the public on the nature of the problem on their platforms and what they intend to do about it.

471. Being accountable also means devoting resources to design platforms and products which contribute to, rather than undermine, the integrity of our online information ecosystem. The platforms and products of technology companies should not incentivise the spread of online falsehoods. This frequently happens when advertisers are allowed to place advertisements on the sites or accounts of those who spread online falsehoods, or when advertisers known for disseminating false information via advertisements are allowed to use a platform's advertising tools (*e.g.*, a digital advertisement, or an online post which amplification is sponsored) to do so. It is imperative for technology companies to ensure that falsehoods are not spread using their advertising tools, and this requires technology companies to improve their due diligence efforts as well, to effectively combat "professional attempts to hide identity in advert purchasing".[303] In this regard, the EC has recommended that technology companies "significantly improve" its "scrutiny of advertisement placements ... to reduce revenues for purveyors of disinformation".[304]

472. Technology companies also have a responsibility to prevent inauthentic accounts (which cannot be traced to an individual) from being created and used on their platforms. The Committee therefore agrees with the EC's call on technology companies to "intensify and demonstrate the effectiveness of efforts to close fake accounts".[305] In order to do so, technology companies must devote resources to ensure they have the means of authenticating and tracing the real persons running the online accounts on their platforms.

473. The Committee accepts the recommendation by many representors that the algorithms of technology companies today have overtly prioritised engagement over credibility, and that technology companies need to find ways to adapt their

[302] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 57.
[303] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 145.
[304] European Commission Communication, "Tackling online disinformation: a European Approach", *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions* (26 April 2018), p 7.
[305] European Commission Communication, "Tackling online disinformation: a European Approach", *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions* (26 April 2018), p 8.

algorithms to prioritise credible content on their platforms, and prevent an identified falsehood from spreading further. This is important in order to "dilute the visibility of disinformation" on their platforms.[306]

474. There is also a need for technology companies to set and enforce credible standards publicly to govern both their products, and the behaviour of users on their platforms. This also means undertaking prior risk assessments when rolling out new platforms, products and features. In this regard, the UK Committee Interim Report has called for a "professional global Code of Ethics" to be developed by technology companies, in which technology companies would set down in writing what is and what is not acceptable on their platforms, and new products would be tested to ensure that products (*e.g.* new technologies and algorithms) are fit-for-purpose and do not constitute dangers to the users, or to society.[307]

475. In sum, a technology company which is serious about implementing measures to adapt its online platforms to respond to the phenomenon of online falsehoods should seek to achieve the following general objectives:

a. Minimise the amplification of falsehoods on its platform;

b. Avoid providing incentives for people to propagate online falsehoods on its platform;

c. Minimise opportunities for purveyors of online falsehoods to hide behind the anonymity of the Internet;

d. Conduct studies and devote resources to improve and safeguard its products from being misused;

e. Equip users and experts to informatively assess the credibility of the information they are exposed to on its platform, and the surrounding information ecosystem;

f. Undertake regular, voluntary reporting of the scale and nature of the problem of deliberate online falsehoods on its platform; and

g. Establish clear and high standards in which it would hold itself to, and its users, and enforce these standards consistently.

[306] European Commission Communication, "Tackling online disinformation: a European Approach", *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions* (26 April 2018), p 8.
[307] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 89.

476. *Government regulation*. The technology companies did not proffer any solutions beyond ongoing voluntary initiatives that the companies were already undertaking or were going to launch. This begs the question of whether online platforms should be further regulated or whether voluntary efforts on their part would be adequate. This issue has been discussed at [430]-[438] above. Substantial evidence was shared and considered which indicated that the voluntary actions taken by technology companies so far have been inadequate.

477. The Government should therefore use legislation and other forms of regulation to ensure technology companies fulfil their responsibilities and take steps to achieve the objectives set out above. Regulation would avoid technology companies being in the unacceptable position of "marking their own homework".[308] The Committee notes that the UK DCMS Committee has, whilst stressing the need for technology companies to do more, equally recommended that the UK Government consider taking proactive measures – including implementing necessary regulation – to tighten the liabilities of technology companies, regulate the use of external targeting on social media platforms, and enforce transparency requirements on technology companies.[309]

478. In order to ensure that technology companies are making the appropriate adaptations, the Government would require the power and expertise to audit technology companies, to ensure, for example, that their algorithms are indeed operating responsibly.[310] In this regard, the Government should also be empowered to request that technology companies submit the necessary information or data for auditing purposes, for example, detailed information about how their algorithms function and whether they have contributed to the amplification of particular falsehoods. The Government should also consider requiring technology companies to ensure that their algorithms do not contribute to the further spread of an identified online falsehood. The Government should also further study how it can holistically prevent the abuse of personal data on online platforms, which can be used to micro-target manipulative content at users.

479. In light of the evidence presented to the Committee that platforms or products run on a different business model and principles may face high barriers of entry, the Government should explore how it can facilitate start-ups or companies which are dedicated towards developing platforms, products and technologies that are designed to ensure the integrity of our online information ecosystem.

[308] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 65.
[309] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), paras 58-60, 144-145.
[310] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 72.

For online platforms

Recommendation 17. To prevent and mitigate the abuse of their platforms to spread online falsehoods, technology companies should:

a. Take proactive action to prevent and minimise the amplification of online falsehoods on their platforms, including by:
    i. Prioritizing credible content on their platforms, and deprioritizing proven falsehoods to limit their circulation.
    ii. Labelling or shutting down accounts and networks of accounts that are designed to amplify online falsehoods, such as inauthentic accounts engaged in other coordinated activity often seen in online disinformation activities.

   The specific measures undertaken may vary depending on how content is amplified on the platform. For example, on a closed messaging platform (such as WhatsApp, Telegram or WeChat and others), minimising the amplification of an online falsehood may involve prohibiting the forwarding of the online falsehood.

   Digital advertising

b. Ensure that their digital advertising tools and services do not incentivise or otherwise aid the spread of online falsehoods. They should disallow:
    i. The placement of advertisements on sites that propagate online falsehoods.
    ii. The use of their advertising services by sites that propagate online falsehoods.
    iii. Their advertising services, such as targeted advertising tools and boosting of posts, from being used to further amplify online falsehoods.

c. Minimise the ability of bad actors to hide their abuse of digital advertising tools by increasing digital advertising transparency.

   For example, technology companies should ensure that users are able to easily see whether the content is sponsored, the sponsor's identity and country of origin, whether they are part of a targeted audience, and what audience that content is targeted at. Technology companies should also consider creating public registers of political advertisements being run on their platforms. Technology companies would also need to undertake some degree of authentication of the users of their advertising tools, including to address sophisticated attempts at masking true identities.

d. Calibrate or restrict the use of digital advertising tools.

Technology companies should take reasonable steps to detect and bar suspicious actors from using digital advertising tools. They should also consider excluding certain audience categories from being targeted, where targeting such categories would encourage prejudice and bias, and restricting the size of targeted audiences.

User data

e. Prevent user data from being used to manipulate people. There is a need to identify appropriate measures for doing so. One measure technology companies could adopt is to inform users of what their data is being used for.

Strengthen the accountability of users

f. Reduce the opportunity for actors to hide behind Internet anonymity to carry out abuse, and facilitate the identification of offenders, including by:
    i. Conducting authentication of users, to ensure they have been set up by real persons.
    ii. Enabling digital identification and source tracing, to reveal the real persons behind accounts or posts, where appropriate.
    iii. Encourage content creators to digitally sign and verify the content they produce and post.

g. Ensure that their policies for user conduct comply with Singapore's policies and norms, and are consistently enforced.

h. Ensure they have the capability to not only respond to abuse, but also to pre-empt it, by:
    i. Conducting regular risk assessments of aspects of their platforms that may be exploited to spread online falsehoods, especially when rolling out new features and tools.
    ii. Conducting regular mapping of the ongoing and evolving nature and use of online falsehoods on their platforms.

Recommendation 18. To contribute to a cleaner online information ecosystem, and foster an informed public, technology companies should implement measures such as the following:

a. Enable users to meaningfully assess the credibility of the information they receive, including by:

i. Disclosing when content is sponsored, and by whom, especially for all forms of digital advertisements.
ii. Using tags to indicate relevant contextual information, such as whether an account is managed by a bot, or the credibility of the source of information.

b. Enable researchers and experts to find solutions to the problem, by providing them with information on how online falsehoods spread, so that they can better understand disinformation tactics and techniques.

c. Inform users of how the design of their platforms influences the content that they receive.

d. Contribute resources to:
   i. Developing technologies that could advance the integrity of information on the Internet, such as the automated detection of online falsehoods, effective detection of hidden identities behind advert purchasing, blockchain-based tools, and fact-checking apps.
   ii. Strengthening the wider information ecosystem, including fact-checking initiatives and quality journalism.

Recommendation 19. Technology companies should demonstrate their accountability to their users, the public and the Government by being transparent about the nature and extent of the spread of online falsehoods on their platforms, and the effectiveness of their responses. Specifically, technology companies should undertake regular voluntary reporting and independent audits. These should cover the following areas:

a. The scale and nature of the problem of online falsehoods on their platforms, and potential risk areas;
b. How their platforms and products have been used to the spread of online falsehoods;
c. The measures taken to address the problem, and to equip informed users; and
d. How effective these measures have been.

<u>For the Government</u>

Recommendation 20. The Government should consider both legislation and other forms of regulation of technology companies to achieve the objectives stated at Recommendations 17 to 19 above. Legislation would be needed particularly for measures to be taken in response to an online falsehood, since

Facebook, Google, and Twitter have a policy of generally not acting against content on the basis that it is false.

The Government should consider whether there is a need for new areas of regulation, such as of targeted advertising and the use and collection of personal data on online platforms for micro-targeting.

To complement legislation, the Government should consider regulatory approaches such as working with technology companies and other industry stakeholders to develop a voluntary code of practice or guidelines to tackle online falsehoods. Where appropriate, the Government should collaborate with technology companies to develop solutions.

Recommendation 21. The Government should explore how it can facilitate the efforts of start-ups and companies to develop platforms, products and technologies which are designed to ensure the integrity of our online information ecosystem.

(5) Deal with Threats to National Security and Sovereignty

a. Rationale and context

480. The use of deliberate online falsehoods to advance State-sponsored disinformation operations has been described above at Part I(B) and Part I(C)(1). The evidence received by the Committee has revealed the threats that disinformation operations pose to national security and sovereignty. They have led to greater friction, distrust and anger in society, political leaders being influenced, elections being undermined, public protests taking place, and even the loss of territorial sovereignty. The Committee also received evidence on why such disinformation operations have occurred and can be expected to occur in Singapore (see above at [209]-[219]).

481. In light of this evidence, the Committee arranged for and had the benefit of a confidential briefing by a security agency in Singapore. The Committee also received evidence from various representors (who are experts in national security) on how to deal with such threats. It was clear from the evidence received that there is a need to consider implementing specific countermeasures to deal with the threats to national security and sovereignty caused by State-sponsored disinformation operations. The recommendations and views of these expert representors are set out below.

b. Representors' views and recommendations

482. <u>*Identify vulnerabilities*</u>. Several representors, hailing from countries affected by disinformation operations, expressed the importance of identifying vulnerabilities within society – both in the general population, and amongst critical sectors and personnel.

a. *General population:* Dr Berzins, national security expert from Latvia, shared that it was important to constantly perform target audience analysis of the population, to understand what the trigger points are and how susceptible the population is to foreign influence. This was echoed by Mr Janda, national security expert from Czech Republic, who has advocated for the State to have precise and up-to-date knowledge on geopolitical attitudes and general vulnerabilities and grievances of its society, so that it can tailor specific long-term or urgent measures targeting weak spots.

b. *Critical sectors and personnel*: Representatives of UCMC from Ukraine highlighted that the economic, energy, financial, transportation, security and information sectors were vulnerable to disinformation and had to be monitored carefully. In their view, these sectors have to be analysed holistically on a continuous basis, to determine the extent to which they have been infiltrated by hostile agents of a foreign state. Mr Janda also highlighted the importance of conducting internal audits and research to measure the geopolitical attitudes of critical personnel like politicians, and military and police officers. This would allow the Government to identify the level of infiltration of hostile influence in its political and security arenas.[311]

483. <u>*Equip vulnerable targets within society*</u>. Tied closely to the recommendation to identify vulnerabilities is the recommendation to train politicians, diplomats, and high-level State bureaucrats on how to identify and resist disinformation campaigns. Many of these individuals can be vulnerable targets as they often fall into hostile active measures without initially knowing or realising it. It is therefore important for them to be trained and provided with common standards on information security and protocols.[312]

484. <u>*Ensure effective real-time communication*</u>. Dr Shashi (Head, Centre of Excellence for National Security, RSIS) observed that the key to countering disinformation operations is to shore up trust between the people and the Government. In this respect, a number of representors spoke about the need for effective real-time communication.

[311] Jakub Janda, "Full-Scale Democratic Response to Hostile Disinformation Operations", *European Values Think-Tank* (20 June 2016), pp 11, 22.
[312] Jakub Janda, "Full-Scale Democratic Response to Hostile Disinformation Operations", *European Values Think-Tank* (20 June 2016), p 21.

a. It has been suggested that the Government should develop a crisis management plan that works with affected entities to provide a transparent, timely, and accurate response to disinformation operations. It was also suggested that the Government conduct inter-agency scenario planning and mock crisis communication exercises to ensure that these plans stay up to date.[313]

b. Dr Berzins and Mr Benjamin Ang (Senior Fellow, RSIS) have both emphasised the need to inform the public about the disinformation operation, and explain the attacker's interests, motivations, and objectives. Similarly, Mr Nimmo, expert representor from the UK, opined that if a large-scale attack on the information system is detected – either as a short burst, such as false stories surrounding a demonstration or debate, or a major attempt to inflame social tensions – the Government can expose the attack in as much details and with as much attribution as possible, to contribute to overall social awareness of the threat.

485. *Improve public education and Government messaging*. Representors highlighted the importance of educating the public and ensuring the correct messaging in combatting foreign-led disinformation operations.

a. *Government Messaging*: Representors of UCMC shared how it was important for States facing disinformation operations to think of and spread a positive and proactive narrative.

b. *Public Education*: Dr Shashi has suggested that Total Defence (particularly its psychological pillar) be revisited and reviewed. He shared that various countries, such as those in northern Europe, have revisited their equivalent of Total Defence, partly because they recognise that threats today could come from "slow burn" issues arising from disinformation and cyber threats.

Dr Shashi and Mr Ang have also voiced support for other public education measures to fight foreign disinformation operations, such as promoting media literacy, encouraging social norms against sharing information without checking, and including disinformation operations as a type of threat in any revamped National Education or Social Studies syllabus.

486. *Partner non-Governmental entities*. Representors also recommended that non-governmental entities be involved in combatting disinformation operations.

487. One suggestion was to work with think-tanks, academics and NGOs on the following:

[313] Muhammad Faizal bin Abdul Rahman et al, "Countering Fake News: A Survey of Recent Global Initiatives", *S. Rajaratnam School of International Studies, Policy Report* (March 2018), p 19.

a. To come up with short-term and long-term scenarios of political development, during which vulnerabilities could be exploited.[314]

b. To advocate against falsehoods. Dr Shashi shared that in Europe, some of the key advocacy against disinformation has been done by think-tanks, who publicly challenge supporters of disinformation allegedly sponsored by Russia, disclose the substance and vehicles of the disinformation campaign, and systemically build social resilience.

c. To challenge disinformation narratives on a regular basis. It has been suggested that activist think-tanks must engage with disinformation daily, and monitor specific disinformation operations and trends on a weekly basis, to create a list of systemic publishers of disinformation that can be publicly reported.[315]

d. To "watchdog" the media space, as well as politicians and institutions which may be sympathetic to a foreign state, and call out any lobbying for the foreign state. [316]

488. Another suggestion was to work with non-governmental initiatives, which have invested great effort in combatting disinformation operations, and have created vast networks that the Government could then tap into. The diversity of the participants' skills and knowledge in such initiatives will aid in building credible narratives against disinformation. Collaboration with these non-governmental initiatives will also provide a quick response to disinformation campaigns as these initiatives will not be encumbered by bureaucratic demands. Examples of such initiatives include StopFake, First Draft, and the International Fact Checking Network.[317]

489. There was also a suggestion to create an independent body that uses grassroots participation to counter disinformation operations. According to Dr Shashi, this body could: (1) carry out research and fact-checking initiatives, and congregate various experts under its umbrella to wage targeted campaigns against fake news (particularly when organised fake news campaigns are brought to bear against the people); (2) produce content for TV, newspapers and social media to debunk fake news and inform audiences, and (3) offer training to media professionals and other relevant parties.

490. *Engage in international efforts*. Some representors have suggested that Singapore participate in international efforts to combat disinformation operations. One

[314] Jakub Janda, "Full-Scale Democratic Response to Hostile Disinformation Operations", *European Values Think-Tank* (20 June 2016), pp 23-24.
[315] Jakub Janda, "Full-Scale Democratic Response to Hostile Disinformation Operations", *European Values Think-Tank* (20 June 2016), p 18.
[316] Jakub Janda, "Full-Scale Democratic Response to Hostile Disinformation Operations", *European Values Think-Tank* (20 June 2016), p 12.
[317] Muhammad Faizal bin Abdul Rahman et al, "Countering Fake News: A Survey of Recent Global Initiatives", *S. Rajaratnam School of International Studies, Policy Report* (March 2018), pp 17-18.

suggestion was for Singapore to host yearly conferences for international bodies fighting against public opinion and psychology oriented information operations.

491. Another suggestion was for ASEAN to spearhead regional efforts to combat disinformation. The EU set up the EU East StratCom Taskforce in 2015 to combat Russian disinformation, and it serves as a regional mechanism that enables collaboration with a wide network of government officials, experts, journalists, and think-tanks. NATO also includes countering Russian disinformation campaigns as part of its strategic communication activities. It was suggested that ASEAN could study these models with a view to introducing similar strategies customised to Southeast Asia's cultural and political landscape.[318]

492. <u>*Enhance Government capabilities to detect and address disinformation*</u>. The following suggestions were made for how the Government can improve its ability to combat disinformation operations.

a. Define targeted and systematic disinformation operations as a threat to national security and the democratic legal system, and include such a threat in our national security or foreign policy strategies.[319]

b. Audit the Government's official communication mechanisms and information security practices, and enhance these based on the outcome of the audit.

c. Increase resources for cyber forensics and intelligence-gathering agencies to enhance their capabilities, such as their investigative capabilities to trace the origins of falsehoods.

d. Establish a multi-disciplinary disinformation analysis team, headed by a leader trusted by key members of the Cabinet. This team will comprise foreign policy, national security, communication and media experts, and homeland security professionals. This includes trusted professionals from the Foreign Ministry, the Defence Ministry, the Interior Ministry, the Army, the Police, and all national intelligence services. This team must be able to play at least four key roles:[320]

i. It must react in real-time to developing disinformation cases with a potentially significant impact on the public or national security.

ii. It must follow regular media coverage and, in cases of serious disinformation, advise the relevant State agencies to publish additional

[318] Muhammad Faizal Bin Abdul Rahman et al, "Countering Fake News: A Survey of Recent Global Initiatives", *S. Rajaratnam School of International Studies, Policy Report* (March 2018), p 17.
[319] Jakub Janda, "Full-Scale Democratic Response to Hostile Disinformation Operations", *European Values Think-Tank* (20 June 2016), p 8.
[320] Jakub Janda, "Full-Scale Democratic Response to Hostile Disinformation Operations", *European Values Think-Tank* (20 June 2016), pp 14-15.

information in real-time. It should not create "its own propaganda" or censor.

iii. It must publish a regular overview of disinformation trends and how they are perceived by politically-neutral State security apparatuses. A nationwide disinformation threat scale could be deployed, similar to the terrorist threat scale.

iv. It must conduct research on the topic and coordinate with similar allied teams. It can categorise disinformation so that State professionals can distinguish which pose potential danger and a stronger impact.

493. *Remove sources of foreign disinformation*. Representors also shared how other countries facing the threat of disinformation operations were taking measures to remove the sources of disinformation.

a. Dr Berzins shared that it is important to develop a system to monitor social media and make social media providers responsible for removing disinformation as quickly as possible, like what was done in Germany.

b. Representatives of UCMC shared that Ukraine enacted a law on quotas for Ukrainian-language content and music on TV and radio, and prohibited social media networks owned or linked to a foreign State. This caused certain foreign-owned websites to lose their dominant position among Ukrainian users.

c. Mr Deynychenko also shared how Ukraine installed special technology in certain territories to weaken the TV signals from a foreign State, and strengthened Ukrainian TV and radio signals.

494. It was also suggested that Singapore should work with overseas partners to help take down servers that are part of a disinformation operation against Singapore.

495. *Deter and penalise sources of amplification*. Removing foreign sources of disinformation may not always be plausible. It is therefore important to also deter and punish those who amplify deliberate online falsehoods. Mr Deynychenko shared that the Ukrainian Parliament is developing a law to effectively punish media organisations for spreading fabrications. Such punishment would include fines and the removal of broadcasting rights.

496. Similarly, a suggestion was also made to impose stiffer penalties for Singapore-based companies that knowingly facilitate the generation and dissemination of disinformation.

c. Observations and Recommendations

497. The use of deliberate online falsehoods to advance State-sponsored disinformation operations is a clear threat to Singapore's national sovereignty and security. The Committee observes that the various other countermeasures mentioned earlier – *i.e.* nurturing an informed public, reinforcing social cohesion and trust, promoting fact-checking, and disrupting falsehoods – would all form a necessary part of the response to State-sponsored disinformation operations. The eradication of both the influence, and presence, of online falsehoods in our information ecosystem would make it harder for State-sponsored disinformation operations to successfully spread in Singapore. Beyond that, there is also a need to consider whether certain countermeasures which specifically address the threat of State-sponsored disinformation operations as proposed or highlighted by expert representors above should be implemented, given the grave threats such attacks pose to our national security and sovereignty.

498. A clear theme which has emerged from the expert evidence received by the Committee was that, in the face of a threat to national sovereignty or security, the "*visible hand of particularly the state is needed*".[321] Almost all of the proposals or experience highlighted by expert representors above require the initiative or involvement of the Government. This is understandable given that there is presently no comprehensive international agreement regulating the use of cyber operations by States. In the circumstances, States have to take the initiative to defend their national interests.

499. Governments have successfully employed the "visible hand" of the State to counter disinformation operations. According to one commentator, it was the determined and coordinated efforts of the French Government which allowed France to successfully withstand the alleged onslaught of State-sponsored disinformation and interference during the 2017 French Presidential Election.[322] Prior to the 2017 French Presidential Election, a number of key French government agencies had reportedly co-operated to implement a unified strategy to counter anticipated disinformation operations or cyberattacks from hostile sources, engaging in pre-emptive measures such as offering practical advice to presidential candidates, and reducing the use of vulnerable technological products for and during the election. This reportedly ensured that certain presidential candidates and also the general public were well-prepared for such attacks, minimising the impact these attacks eventually had on French voters. All in all, the French experience showed that governments which are "determined ... and organised enough" can successfully preserve their State's national sovereignty and security even in the age of cyberspace.[323]

[321] Liew Kai Khiun, Appendix III: Written Representations, Paper No. 46, page B190, para 3.
[322] Jonathan Eyal, "How France fought off influence ops in the last election", *The Straits Times* (2 July 2018).
[323] Jonathan Eyal, "How France fought off influence ops in the last election", *The Straits Times* (2 July 2018).

500. In the same vein, the UK Committee Interim Report has also recommended the importance of deeper collaboration within Government to counter the threat of disinformation operations. The UK Committee Interim Report has suggested that Government departments "should be working together, [and] sharing data, intelligence and expert knowledge".[324]

501. The Committee is of the view that given the threats posed by State-sponsored disinformation, the Government should closely study the specific countermeasures proposed or highlighted by the expert representors above, and outline a unified strategy for countering State-sponsored disinformation operations which seek to threaten our national security and sovereignty.

> **Recommendation 22.** The Government should study the specific countermeasures proposed by expert representors, and come up with a national-level strategy and coordinated approach for countering State-sponsored disinformation operations.

[324] "Disinformation and 'fake news': Interim Report", *UK House of Commons Digital, Culture, Media and Sport Committee* (29 July 2018), para 200.

(III) SUMMARY

502. The Committee's mandate is to consider and give its views on a serious challenge that Singapore and many other countries face – the phenomenon of deliberate online falsehoods. Dealing with this phenomenon has required melding the technical study of hard evidence, with a critical understanding of the ideals and values that should inform solutions for this country.

503. This public inquiry hence not only involved testing the evidence presented; it also involved open debates over fundamental issues, such as what democratic public discourse requires, the purpose of freedom of expression in a democratic society, and the role of the media and journalism. This Report and its recommendations are a result of these public debates, the contributions of people who generously gave their time, effort and expertise, and other illuminating international and local research.

504. It is the hope of this Committee that the findings in this Report will help shape effective responses by the Government, and spur other stakeholders to action. We also hope that the Committee's findings will be a resource for the Singapore public to more fully understand a phenomenon that could appear in their everyday lives, causing harm well before it is noticed.

Understanding the Phenomenon of Deliberate Online Falsehoods

505. The power of a falsehood has been demonstrated time and again. Falsehoods influence people's emotions, decisions and actions. They can especially arouse anger, fear, and mistrust, and drive people to harmful decisions and actions. Long before the digital revolution, foreign States have found falsehoods a useful tool to manipulate the public and weaken target countries.

506. **Deliberate online falsehoods are, however, a unique phenomenon of an unprecedented scale.** This phenomenon is made possible by the Internet and digital technology. It is unique because of the accessibility of its underlying methods to a far wider range of actors, the ease with which an online falsehood can reach into everyday lives, the scale and strength of its impact on society, and the greater difficulties faced in combatting it, as explained below.

507. **The phenomenon of deliberate online falsehoods is pervasive.** Deliberate online falsehoods have gained currency from multiple flanks – from organised foreign disinformation campaigns, to the rough–and-ready tactics of disgruntled or thrill-seeking persons, to a single WhatsApp message by an angry individual in the midst of racial tensions. With the Internet and digital technology, deliberate online falsehoods are being used to successfully advance not only the geopolitical agendas of States, but also to illegitimately promote the everyday causes, ideologies, politics and prejudices of civilians – with serious consequences.

508. It is striking that the same digital tools and techniques used by States in modern disinformation warfare are so easily accessible to average civilians. The Internet has democratised information; it has also democratised its weaponisation.

509. **Deliberate online falsehoods can take effect in ways that are not visible, until too late.** The impact of online falsehoods is not always immediate. "Low level" falsehoods can be slowly cumulative in their impact, eventually leading to more serious crises. It was said in evidence that exposure over time to falsehoods mixed with partisan views on social media can skew world views. One example given was of a man who drove a van into a crowd outside a London mosque, after being radicalised online by anti-Muslim conspiracy theories and hate speech. On a similar note, a recent study suggested that anti-refugee attacks in German towns were facilitated by the exposure over time of individuals to fear and anger in online "echo chambers".

510. **Deliberate online falsehoods can threaten a nation's security – they can damage the social fabric of a nation, cause the loss of lives, and harm democratic institutions and free speech**. A key pattern observed is that deliberate online falsehoods tend to follow the fault lines of a given society; they play on what people are afraid of or are angry about. These fault lines can be ideological, political, or identity-based, and can change with the times. Such falsehoods led an American to fire a gun in a pizza restaurant in Washington, DC. They provoked massive rallies during elections in Indonesia, and encouraged anti-immigrant demonstrations in Europe. They have had horrific consequences – such as instigating angry mobs to burn down temples and monasteries in Indonesia, and to murder amidst communal violence in India.

511. Evidence from other countries also shows how falsehoods "rely on the strength of the weak". Malicious actors have instigated internal opposition by targeting falsehoods against public institutions, and using falsehoods to feed the narrative of a broken social contract. By eroding trust in public institutions, deliberate online falsehoods impede constructive policy-making, and undermine the ability of public institutions to effectively protect the country from threats and crises.

512. Falsehoods can derail the contestation of ideas in the "marketplace" by making people too angry to understand each other, and too intimidated to express differing views. They can crowd out alternative perspectives and reliable facts. When falsehoods are spread in the "marketplace", the chances of the truth prevailing are weak. By causing citizens to be misinformed, they negate citizens' right to genuine freedom of expression.

513. Falsehoods can erode people's trust in the sources of information that previously helped them make sense of the world. They can diminish the role of public institutions and the mainstream media as traditional sources of authoritative information. The proliferation of falsehoods can, according to research, make

people stop trusting facts in general, and disengage from public debate. Without shared facts, democratic discourse will have weak foundations to stand on.

514. When spread during elections, falsehoods deprive citizens of their right to informed political participation. By casting doubt on the legitimacy of an electoral outcome, they undermine people's assurance of a representative government. There is some evidence that falsehoods can sway voting behaviour, although the overall evidence on voting impact is so far unclear. While there is so far no clear evidence that election outcomes have been affected, evidence nevertheless shows that public opinion has been influenced. Public confidence in the electoral process has also been affected.

515. **Deliberate online falsehoods are used to violate national sovereignty; they are an attractive weapon in information warfare, which can be continuous and covert.** One of the most egregious consequences of deliberate online falsehoods is its use to violate a country's sovereignty. With digital technology, information warfare can be carried out continuously, to advance a State's geopolitical interests whether in war or peacetime. Deliberate online falsehoods are an attractive weapon for doing so. They can effectively de-stabilise a country at a disproportionately low cost. They are harder to detect, and easier to disavow.

516. The threat is heightened when the objectives of foreign State and local actors align. Disinformation operations often exploit local concerns and grievances, disguising foreign interference as an organic local movement. Countermeasures overseas to stem foreign interference have hence met with local opposition.

517. The methods used by some States have been ingenious in their relative simplicity. Shielded by online anonymity, State disinformation agents have used fake social media accounts to infiltrate local communities, and turn people against each other. Digital advertising tools, bots, troll accounts and "click-bait" have been used to amplify falsehoods about opposing sides of an issue, and create conflict in the physical world.

Risks to Singapore

518. **Singapore is a target of hostile information campaigns.** Evidence points to foreign State-sponsored information campaigns having been carried out against Singapore. The Committee received evidence clearly suggesting that online news articles and social media have been used to influence Singaporeans and legitimise another State's international actions. Some of this evidence was received in private hearings of the Committee. The series of cyber-attacks against Singapore, including the recent hacking of SingHealth's databases, are also an indicator; both disinformation and cyber-attacks are part of a spectrum of non-military tools commonly used in information warfare.

519. **Developing disinformation capabilities in the region can be used against Singapore.** Neighbouring countries are seeing a rise in the use of disinformation capabilities, such as for-profit syndicates, bot armies and "data-driven political consultants" with expertise in using data analytics to micro-target messages at susceptible people. The techniques used can be sharpened for the local context, and turned against Singapore.

520. **Singapore's social cohesion can be harmed by online falsehoods that play on local and regional fault lines; they can take the form of a sustained campaign or "low level" everyday falsehoods.** The realities of Singapore's diverse social landscape create wide opportunities for falsehoods to undermine Singapore's social cohesion. Survey findings presented to the Committee show that Singapore is not a race-blind society, and differences do matter. Any source of difference, including racial, ideological differences and social inequalities, can be exploited, turning cracks into chasms.

521. Fault lines run across national borders. Tensions in the region can spill over into Singapore. An example given was of Islamophobic online trolling over local media reports on the crisis faced by Muslims in Myanmar's Rakhine state. The trolling appeared to come from foreign user accounts, and triggered backlash from accounts that appeared to belong to local Muslims.

522. "Low level" everyday online falsehoods are also a risk to Singapore's social cohesion. It was explained that while emotions may not be high initially, such falsehoods can gradually raise tensions. An example given was of the false story spread by a news website called "The Real Singapore" that falsely portrayed a Filipino family as disrespectful and disruptive of local religious customs, provoking xenophobic online comments. The possibility of a sustained campaign, especially in sudden crises and times of heightened tension, also remains.

The Difficulties of Combatting Deliberate Online Falsehoods

523. **Deliberate online falsehoods have the upper hand over facts; the playing field is far from equal.** Research was presented to the Committee, which showed that falsehoods are already generally stronger in influence than the facts, even when offline. Due to mental and psychological tendencies, falsehoods tend to continue influencing a person, even after the person receives the facts. These human tendencies can affect people from all educational backgrounds and world views, including the well-educated and the middle ground.

524. The inequality between falsehood and fact has been widened by the Internet and digital technology. An online falsehood can speed from an obscure YouTube channel to the Facebook page of a prominent media outlet in a day or less, cascading through a multitude of online sites and accounts, gathering credibility and influence with more clicks and shares. Even though corrections can also tap on social media and digital technology, they have generally been shown to be

much weaker in reach compared with falsehoods. Online falsehoods have been shown to be faster in speed, and wider in reach, than corrections; their reach may be 70% greater, according to one study. Still, corrections remain important.

525. A further advantage is gained by abusing easily accessible digital tools. It is not difficult to create a convincing online falsehood that would take precious hours to verify. Even a fake video or audio clip, known as a "deep-fake", can be made for less than US$100. On social media, online falsehoods can be rapidly amplified using fake accounts run by automated bots and human trolls. They can be targeted at susceptible audiences using digital advertising tools. Crude "click-bait" can capitalise on social media algorithms to drive up user engagement with online disinformation outlets.

526. **Sophisticated techniques and tools will keep improving.** States are writing a serious playbook on online disinformation. The techniques seen in the 2016 US Election are not only being used again in the upcoming US Mid-Term Election, but also by various actors against other countries. Greater effort appears to have been made to avoid detection.

527. Online disinformation is becoming professionalised and commercialised. One can buy bot armies, click farms, and petition signatures, and hire people to manipulate votes, and even instigate a street protest. "Cyber propagandists" operate at a high level of sophistication, and are determined to succeed. For example, according to cyber-security firms, they prepare supporting background and side stories for key false narratives to fool more informed readers. The prospect of profits would incentivise the development of methods that are even harder to counter.

528. **Even as deliberate online falsehoods are becoming harder to combat, social resilience to falsehoods has tended to lessen.** Some people tend to be more susceptible to falsehoods in the online realm. With the flood of information online, people tend to rely more on mental shortcuts and "skimming", rather than proper processing of the information received. They tend to be attracted to online "echo chambers", which have been found by researchers to encourage intolerance to facts that contradict their world views, and to be a primary driver of the spread of online falsehoods.

529. The quality of information and discourse online is variable. The online news landscape is seeing a proliferation of sources that do not apply standards of professional journalism. Political discourse is increasingly carried out on social media, despite a fundamental misfit in design, as explained to the Committee by some representors. Social media often facilitates political discourse that is emotionally-driven and convenient, rather than reasoned and considered.

530. **In conclusion, the phenomenon of deliberate online falsehoods is a real and serious problem here and around the world.** It has been said that the US was caught off-guard in the 2016 US Election despite early warning signs, because of

a belief in its society's resilience. Singapore should seek to avoid this scenario. While the difficulties may be daunting, the Committee believes that Singapore should confront these challenges and try to deal with them.

Responding to the Phenomenon of Deliberate Online Falsehoods

531. To adapt to and combat this global "new normal", Singapore's response should be guided by the core values and aspirations of our society. Ultimately, what is desired is a public that is informed and respects the facts, a society that is cohesive and resilient, and a people whose sovereignty and freedom are safeguarded.

Key Principles

532. **It is clear to the Committee that the response must be multi-pronged.** The phenomenon has many facets to address – such as the capacity of individuals and those who support the integrity of information in society (such as journalists and fact-checkers), the trust that makes a society resilient and cohesive, the falsehoods themselves and those responsible, as well as the supporting digital ecosystem, particularly the role of technology companies.

533. **Responses must address the asymmetry between the growing power of technology and the capacity of society and countries.** The phenomenon and its problems demonstrate a growing gap between the power of technological developments and the capacity of societies and governments to deal with them. The evidence has showed how, in the online realm, the phenomenon of deliberate online falsehoods is gaining strength faster than laws and norms can keep up.

534. **Legislative and non-legislative measures are required; there is no silver bullet.** For Singapore to keep pace, it is vital to build up the capacity of individuals, adapt society, and strengthen the capabilities and powers of the Government. These aspects are mutually reinforcing, and each is equally important.

535. The Committee is convinced that Government intervention is necessary. The notion that the problem can be dealt with through the free flow of information in a "marketplace of ideas" wrongly assumes that the playing field is equal. The evidence has shown that deliberate online falsehoods have the upper hand by far, due to psychological, technological, and social factors. While building the capacity of individuals and other stakeholders through non-legislative measures is crucial, these alone are insufficient to deal with the strength and serious consequences of deliberate online falsehoods.

536. The Committee finds that new legislation is needed to provide the necessary scope, speed and adaptability to deal with the realities of the phenomenon. Notably, among the key first steps taken by countries such as Ukraine, the Czech Republic and France was to review their legislation to ensure they were adequate

to deal with the phenomenon, and to make the necessary changes. A similar exercise was conducted and presented to the Committee by Law Dean, Associate Professor Goh Yihan.

537. The Committee considered concerns about the potential limitations of legislation to deal with the constraints of national borders and the rapid advancements of technology (at [408]-[411]). These are real challenges that the Government should seek to deal with by making improvements through an iterative process, rather than passively waiting for a perfect solution that may never be found.

538. **Government intervention requires calibration.** Falsehoods can appear in a broad spectrum of circumstances – from deliberately fabricated content to satire and parodies. They can also have varying degrees of impact – from causing minor confusion to threatening national security and dividing societies. Government intervention should be calibrated in a manner that takes these factors into consideration, especially given the potential for real-world impact and consequences.

539. The Committee gave detailed consideration to concerns about the impact of Government intervention on free speech (see [412]-[425]). These are valid and important concerns. They should be addressed through calibrated interventions and legal and institutional safeguards, not by foregoing actions that are needed to protect society. Measures to combat deliberate online falsehoods and the right to free speech both serve the same democratic ideals.

Specific Recommendations

540. The Committee recommends measures to achieve the following objectives:

a. Nurture an informed public;
b. Reinforce social cohesion and trust;
c. Promote fact-checking;
d. Disrupt online falsehoods; and
e. Deal with threats to national security and sovereignty.

(A) Nurture an Informed Public

(i) Public Education

541. Public education is an essential long-term measure to ensure that citizens are well-informed, able to discern truth from falsehood, and able to interrogate information sources effectively and critically. Public education has to be broad-based, to include the "political economy of falsehoods", moral and civic education, and critical thinking skills especially amongst students.

542. Public education also has to reach all segments of society, not just the young and educated, but those who are less educated and less Internet-savvy as well. This therefore requires public education to be conducted on the appropriate medium, to effectively reach the entire population. With these in mind, the Committee makes the following recommendations.

543. **Recommendation 1.** To ensure that public education efforts have the necessary scope and scale, the Government should consider putting in place a national framework to coordinate and guide public education initiatives. This framework should have the following elements:

a. An expanded, broad-based curriculum for schools that would include:
   i. a component specifically on the motivations and agendas of disinformation agents and their techniques and strategies;
   ii. moral and civic education, to foster active and constructive public discourse and responsible online behaviour; and
   iii. imparting critical thinking skills creatively.

   This curriculum should be regularly updated with the latest research and knowledge about the problem of online falsehoods.

b. A framework of desired skills and outcomes to:
   i. guide public education efforts in building information and media literacy among Singaporeans. This framework should similarly be informed by research on the problem of online falsehoods; and
   ii. coordinate ministry actions, including overarching outreach, to ensure coverage of all segments of society.

544. **Recommendation 2.** The Government should consider encouraging and providing the necessary support for innovative and ground-up campaigns or initiatives for public education, to widen effective outreach beyond Government-led initiatives.

(ii) Quality Journalism

545. Quality journalism ensures that society has trusted sources that publish information in a manner that is accurate, informative, purposeful, playing an important role in allowing citizens to be well-informed. To ensure quality journalism, journalists of all stripes need to be trained to conduct effective fact-checking and engage in accurate reporting. Both the mainstream and alternative media should be held to the same journalistic standards of intellectual integrity and factual accuracy.

546. In this regard, greater dialogue between Government and news platforms, including those which solely operate online, will also be beneficial. With these in mind, the Committee makes the following recommendations.

547. **Recommendation 3.** News organisations, technology companies and institutes of higher learning should consider ways to ramp up the training of journalists of all backgrounds, especially in techniques for ensuring accuracy in a new and rapidly evolving digital news environment.

548. **Recommendation 4.** Journalists should also proactively find ways to update their skills in digital fact-checking, and arm themselves with knowledge of how online falsehoods and disinformation campaigns work.

549. **Recommendation 5.** Both the mainstream media and the alternative news platforms should hold themselves to the same professional standards of journalism, ensuring there is fairness, accuracy and integrity in reporting.

550. **Recommendation 6.** The Government should consider how it can support the objectives in Recommendations 3 to 5.

(B) Reinforce social cohesion and trust

551. Trust holds a country and society together in the face of attempts to divide. It is such trust that disinformation agents seek to erode. The impact of online falsehoods is often a symptom of underlying societal conditions and fault lines, which need to be bridged through strengthening trust, whether between different groups of people, or between people and public institutions.

(i) Trust among people and communities

552. While the Committee's present inquiry relates specifically to deliberate online falsehoods, and not to social cohesion generally, the Committee accepts that strengthening trust in society is an important means of bridging fault lines which can be exploited by perpetrators of deliberate online falsehoods.

553. The Committee also recognises that the social harmony experienced in Singapore today has been achieved by taking an active approach towards fostering multi-culturalism and multi-racialism; and is always a work-in-progress. Numerous platforms have been established to respond quickly to racial and religious tensions; ground-up initiatives have also been critical to address issues that may divide communities. These efforts will need to evolve to address specific areas relating to deliberate online falsehoods. With these in mind, the Committee makes the following recommendations.

554. **Recommendation 7.** Organisations and initiatives for the promotion of social cohesion, both old and new, should consider providing clarifications and information on distortions and falsehoods affecting social cohesion. In doing so, they should consider adopting the following principles recommended by representors, where relevant:

a. Employ people-to-people interaction and communication.
b. Create "safe spaces" for exchanging views and perspectives on sensitive issues.
c. Serve as voices of influence in society, to cultivate a strong core of people who are less susceptible to deliberate online falsehoods.
d. Mediate honest discussion among differing groups.
e. Reach into and across "echo chambers".

555. **Recommendation 8.** The Government should consider supporting or conducting research to understand society's vulnerabilities.

(ii) Maintain trust in public institutions

556. The Committee's present inquiry relates specifically to deliberate online falsehoods, and not to the conduct of governance generally. That said, the Committee agrees that strong trust in public institutions is essential in making it harder for deliberate online falsehoods to take effect. Pre-emptive and responsive measures should therefore be taken to ensure public institutions remain trusted in Singapore. With these in mind, the Committee makes the following recommendations.

557. **Recommendation 9.** Public institutions should, wherever possible, provide information to the public in response to online falsehoods in a timely manner. They should also seek to pre-empt vulnerabilities and put out information in advance, where appropriate, to inoculate the public. They should ensure that they communicate with the public in clear and comprehensible terms.

558. **Recommendation 10.** Existing efforts should be reviewed, to consider whether they are adequate to achieve the following:

a. *Transparency.* Swiftly communicating information in response to online falsehoods, the reasons for any Government action against online falsehoods, and the reasons for decisions to not disclose information to the public.

b. *Participation and communication.* Engaging the public on Government strategies against online disinformation operations.

c. *Accountability.* Assuring the public of the integrity of the information the Government puts forward concerning public institutions.

(C) Promote fact-checking

559. The Committee is of the view that credible fact-checking initiatives can provide speedy debunks of falsehoods, which constitute a trusted mechanism for people to have access to reliable facts when deliberate online falsehoods are spread. It also helps in creating an ecosystem where accuracy and veracity are valued.

560. The Committee received evidence on how a fact-checking coalition can help pull together valuable resources from otherwise competing media organisations, and tap on the expertise of partners from different industries to fact-check the falsehoods quickly and accurately.

561. The Committee however, notes the concerns raised by several representors on the limitations of fact-checking initiatives. Additionally, the issue of how involved the Government should be in fact-checking remains to be further explored.

562. That said, the Committee is of the view that ultimately, whether a fact-checking coalition will be trusted and relied upon by people depends on its credibility and its effectiveness, and that what is critical is that any fact-checking initiative by committed to presenting the truth to the public. With these in mind, the Committee makes the following recommendations.

563. **Recommendation 11.** There is a role for trusted fact-checking initiatives in combatting deliberate online falsehoods. Different media organisations and partners from other industries should consider establishing a fact-checking coalition in Singapore to debunk falsehoods swiftly and credibly, or providing relevant support to such credible fact-checking initiatives as appropriate. There are differing views on the role, if any, that the Government can play in supporting fact-checking initiatives. Thus, this aspect needs to be further considered.

(D) Disrupt Online Falsehoods

564. Developments in the digital realm are outpacing the rules and norms of societies around the world. Actors seeking to create and disseminate online falsehoods find wide space in the online world to take advantage of new and sophisticated digital methods and tools with impunity.

565. Strong action is needed to ensure that the Internet does not remain a "Wild West", as the UK DCMS described it to be, but a realm where people can truly enjoy the freedom and benefits that they do in the offline realm. To achieve this, the Committee proposes a range of calibrated measures to tackle the problem head-on – by disrupting online falsehoods and changing the digital ecosystem which sustains them. These measures have the aim of (i) stemming the spread of online falsehoods to mitigate the harm caused, (ii) deterring their creation and spread, and (iii) preventing the abuse of digital tools and platforms to do so.

566. The Committee has not received any evidence that shows that the technology companies can or will effectively deal with the problem, without adequate legislation. During the hearing, representatives from Facebook, Twitter and Google/YouTube confirmed that they will not generally, as a matter of policy and absent legislation, remove content on their platforms on the basis that it is false. This is despite the fact that the spread of falsehoods on their platforms have threatened national security, undermined public institutions, and even caused the loss of lives.

(i) Counter and deter the spread of online falsehoods

567. The ability to swiftly stem the spread and influence of online falsehoods is vital. Exposure to falsehoods can influence people in ways that are difficult to dispel. Allowing a falsehood to circulate can increase its influence.

568. **Recommendation 12.** The Government should have the powers to swiftly disrupt the spread and influence of online falsehoods.

The objectives to be achieved should be as follows:

a. Provide access to and increase the visibility of corrections, including through tagging functions and the use of other platforms with significant reach.
b. Limit or block exposure to the online falsehood.
c. Disrupt the digital amplification of online falsehoods, including through the use of false amplifiers (*e.g.* inauthentic accounts run by bots or trolls), and digital advertising tools.
d. Discredit the sources of online falsehoods.

These capabilities should be able to apply to all relevant platforms regardless of their technological basis. There needs to be careful balance and calibration to prevent the public interest from being harmed, and to at the same time respect communications that are personal, private, and of limited circulation. They should therefore apply both to open and closed platforms.

Legislation will be needed to achieve the above objectives. Such legislation should have the following objectives:

e. The measures will need to achieve the objective of breaking virality by being effective in a matter of hours.
f. The decision-maker should be effective and credible.
g. There should be adequate safeguards in place to ensure due process and the proper exercise of power, and give assurance to the public of the integrity of the decision-making process.

h. The measures should be deployed in a calibrated manner, taking into account the context and circumstances, including potential impact and reach.

Measures provided in the legislation could include: tagging of corrections and notifications, take-down powers and access-blocking, among other measures. This should include judicial oversight where appropriate.

569. **Recommendation 13.** The Government should identify the additional measures needed to safeguard election integrity, and implement the necessary measures, including legislation, in view of the issues that have been highlighted in this report.

570. **Recommendation 14.** The Government should consider implementing monitoring and early warning mechanisms, to facilitate assessments of when and how to intervene to stop the spread of online falsehoods.

571. Addressing the provenance of the problem is necessary. The Committee is supportive of measures to ensure deterrence and accountability of perpetrators of deliberate online falsehoods. This include ensuring that digital advertising platforms or digital advertisers are not supporting purveyors of online falsehoods; and imposing punitive measures on the perpetrators of deliberate online falsehoods.

572. The Committee is not calling for the criminalisation of all online falsehoods. Consistent with the calibrated approach the Committee has recommended, criminal sanctions should be used only against purveyors of online falsehoods that meet a prescribed threshold.

573. **Recommendation 15.** The Government should consider powers needed to establish a de-monetisation regime, including through legislation which will:

a. Disrupt the flows of digital advertising revenue to purveyors of online falsehoods. This should take into account the responsibility of different stakeholders in the digital advertising ecosystem.

b. Require the disgorgement of financial benefits by purveyors of online falsehoods. This should cover the "hired guns" who are paid by others to create and spread online falsehoods.

574. **Recommendation 16.** Criminal sanctions should be imposed on perpetrators of deliberate online falsehoods. These deterrent measures should be applied only in circumstances that meet certain criteria. There should be the requisite degree of criminal culpability (*i.e.* intent or knowledge), in accordance with established criminal justice principles. There should be a threshold of serious harm such as election interference, public disorder, and the erosion of trust in public institutions.

The Government should ensure these deterrent measures are adequate in scope to cover the range of methods and actors, including the deliberate use of inauthentic accounts or bots, the provision of tools and services to publish falsehoods, and the masterminds behind online falsehoods, who may not always be the ones creating or spreading them.

Whether existing criminal sanctions are adequate to achieve the above should be considered.

(ii) Adapt online platforms

For Online Platforms

575. The Committee received evidence on how the design of online platforms by technology companies has a direct impact on the dissemination of both online falsehoods and quality journalism; and that the content allowed on these platforms can influence human behaviour considerably. Online platforms have been exploited to manipulate public opinion or spread online falsehoods – through their algorithms, the use of inauthentic accounts, and their digital advertising services.

576. Given the amount of control technology companies have over the design of their platforms, through which they have profited greatly, the Committee is of the view that technology companies need to do more. Where appropriate, the Government needs to have in place appropriate legislation, to ensure that technology companies contribute to a clean Internet ecosystem.

577. There is a clear need for technology companies to increase their transparency and improve their accountability. With these in mind, the Committee makes the following recommendations.

578. **Recommendation 17.** To prevent and mitigate the abuse of their platforms to spread online falsehoods, technology companies should:

a. Take proactive action to prevent and minimise the amplification of online falsehoods on their platforms, including by:
    i. Prioritizing credible content on their platforms, and deprioritizing proven falsehoods to limit their circulation.
    ii. Labelling or shutting down accounts and networks of accounts that are designed to amplify online falsehoods, such as inauthentic accounts engaged in other coordinated activity often seen in online disinformation activities.

The specific measures undertaken may vary depending on how content is amplified on the platform. For example, on a closed messaging platform (such as WhatsApp, Telegram or WeChat and others), minimising the

amplification of an online falsehood may involve prohibiting the forwarding of the online falsehood.

Digital advertising

b. Ensure that their digital advertising tools and services do not incentivise or otherwise aid the spread of online falsehoods. They should disallow:
    i. The placement of advertisements on sites that propagate online falsehoods.
    ii. The use of their advertising services by sites that propagate online falsehoods.
    iii. Their advertising services, such as targeted advertising tools and boosting of posts, from being used to further amplify online falsehoods.

c. Minimise the ability of bad actors to hide their abuse of digital advertising tools by increasing digital advertising transparency.

    For example, technology companies should ensure that users are able to easily see whether the content is sponsored, the sponsor's identity and country of origin, whether they are part of a targeted audience, and what audience that content is targeted at. Technology companies should also consider creating public registers of political advertisements being run on their platforms. Technology companies would also need to undertake some degree of authentication of the users of their advertising tools, including to address sophisticated attempts at masking true identities.

d. Calibrate or restrict the use of digital advertising tools.

    Technology companies should take reasonable steps to detect and bar suspicious actors from using digital advertising tools. They should also consider excluding certain audience categories from being targeted where targeting such categories would encourage prejudice and bias, and restricting the size of targeted audiences.

User data

e. Prevent user data from being used to manipulate people. There is a need to identify appropriate measures for doing so. One measure technology companies could adopt is to inform users of what their data is being used for.

Strengthen the accountability of users

f. Reduce the opportunity for actors to hide behind Internet anonymity to carry out abuse, and facilitate the identification of offenders, including by:
    i. Conducting authentication of users, to ensure they have been set up by real persons.
    ii. Enabling digital identification and source tracing, to reveal the real persons behind accounts or posts, where appropriate.
    iii. Encourage content creators to digitally sign and verify the content they produce and post.

g. Ensure that their policies for user conduct comply with Singapore's policies and norms, and are consistently enforced.

h. Ensure they have the capability to not only respond to abuse, but also to pre-empt it, by:

    i. Conducting regular risk assessments of aspects of their platforms that may be exploited to spread online falsehoods, especially when rolling out new features and tools.
    ii. Conducting regular mapping of the ongoing and evolving nature and use of online falsehoods on their platforms.

579. **Recommendation 18.** To contribute to a cleaner online information ecosystem, and foster an informed public, technology companies should implement measures such as the following:

a. Enable users to meaningfully assess the credibility of the information they receive, including by:
    i. Disclosing when content is sponsored, and by whom, especially for all forms of digital advertisements.
    ii. Using tags to indicate relevant contextual information, such as whether an account is managed by a bot, or the credibility of the source of information.

b. Enable researchers and experts to find solutions to the problem, by providing them with information on how online falsehoods spread, so that they can better understand disinformation tactics and techniques.

c. Inform users of how the design of their platforms influences the content that they receive.

d. Contribute resources to:
    i. Developing technologies that could advance the integrity of information on the Internet, such as the automated detection of

online falsehoods, effective detection of hidden identities behind advert purchasing, blockchain-based tools, and fact-checking applications.

ii. Strengthening the wider information ecosystem, including fact-checking initiatives and quality journalism.

580. **Recommendation 19.** Technology companies should demonstrate their accountability to their users, the public and the Government by being transparent about the nature and extent of the spread of online falsehoods on their platforms, and the effectiveness of their responses. Specifically, technology companies should undertake regular voluntary reporting and independent audits. These should cover the following areas:

a. The scale and nature of the problem of online falsehoods on their platforms, and potential risk areas;
b. How their platforms and products have been used to the spread of online falsehoods;
c. The measures taken to address the problem, and to equip informed users; and
d. How effective these measures have been.

For the Government

581. **Recommendation 20.** The Government should consider both legislation and other forms of regulation of technology companies to achieve the objectives stated at Recommendations 17 to 19 above. Legislation would be needed particularly for measures to be taken in response to an online falsehood, since Facebook, Google, and Twitter have a policy of generally not acting against content on the basis that it is false.

The Government should consider whether there is a need for new areas of regulation, such as of targeted advertising and the use and collection of personal data on online platforms for micro-targeting.

To complement legislation, the Government should consider regulatory approaches such as working with technology companies and other industry stakeholders to develop a voluntary code of practice or guidelines to tackle online falsehoods. Where appropriate, the Government should collaborate with technology companies to develop solutions.

582. **Recommendation 21.** The Government should explore how it can facilitate the efforts of start-ups and companies to develop platforms, products and technologies which are designed to ensure the integrity of our online information ecosystem.

(E) Deal with Threats to National Security and Sovereignty

583. To safeguard our sovereignty and security, the threats which deliberate online falsehoods pose in the form of State-sponsored disinformation operations must be effectively dealt with.

584. A clear theme which has emerged from the expert evidence received by the Committee was that, in the face of a threat to national sovereignty or security, the "*visible hand of particularly the state is needed*". The coordinated efforts of various agencies within the French Government during the 2017 French Presidential Election proved that a determined and organised government can preserve a State's national security and sovereignty in the face of State-sponsored disinformation operations. With these in mind, the Committee makes the following recommendation.

585. **Recommendation 22.** The Government should study the specific countermeasures proposed by expert representors, and come up with a national-level strategy and coordinated approach for countering State-sponsored disinformation operations.

ANNEX A: ACTORS WHO USE FALSEHOODS AND THEIR OBJECTIVES

1. **Mr Septiaji Eko Nugroho**, founder of the Indonesian Anti-Hoax Community or MAFINDO, and **Associate Professor Eugene Tan** from the SMU School of Law recognised that actors who create and/or spread deliberate online falsehoods do so mainly for political/ideological reasons, or for economic/financial reasons. **Dr Liew Kai Khiun** was of the view that that there are ongoing efforts around the world to create and spread deliberate online falsehoods for economic, political or criminal purposes.

2. Political objectives: The **editors of Channel NewsAsia** and **Mr Benjamin Ang** agreed that deliberate online falsehoods aim to achieve a variety of political objectives. These include attacking public institutions and individuals, sowing discord amongst racial and religious communities, exploiting fault-lines, undermining public institutions, interfering in elections as well as other democratic processes, and weakening countries.

3. **Ms Myla Pilao**, Director for Technology Marketing at Trend Micro, shared a similar view. She explained that politically motivated disinformation campaigns are generally designed to get people to change their political belief or opinion. Such campaigns will aim to destabilise target countries during major political events like national elections, or discredit personalities such as politicians, influencers, or even journalists that oppose the perpetrators' intended outcomes.

4. Representors such as **students from SMU and NUS** concurred that disinformation campaigns aim to erode social cohesion, sow discord, destroy trust either between communities, within communities, between communities and authorities and also in mainstream media.

5. Financial objectives: Ms Pilao also explained that campaigns motivated by financial gain cause individuals or groups to suffer public shaming just so that the campaign operators can line their pockets. Even businesses risk damage to their corporate reputation, as someone can launch a viral smear campaign against their owner, flagship product, or service.

(1) Foreign State Actors

6. Representors shared how foreign State actors have spread falsehoods to achieve various political objectives, and that they can do so in the form of disinformation campaigns.

7. **Mr Ruslan Deynychenko**, one of the founders of StopFake, was of the view that foreign disinformation campaigns aim to weaken a country, reduce its ability to resist foreign aggression, change its foreign policy, and create conditions for its inclusion in a foreign country's sphere of influence.

8. Various representors from Eastern European countries referred to the alleged use of disinformation campaigns by a particular foreign State, to achieve certain political objectives. For instance, **Mr Jakub Janda**, the Head of the Kremlin Watch Program and Director of the European Values think-tank in the Czech Republic, opined that this foreign State has three primary interests in Europe, which inform its use of disinformation operations: (1) the strengthening of its political allies, (2) undermining trust towards democratic politicians and institutions and legitimising extremists and disinformation projects, and (3) undermining public support for EU and NATO membership.

9. In terms of modalities, **Mr Nicholas Fang** described how influence operations can be instigated by larger, more powerful nations who have at their disposal a full range of information tools. These can comprise a compliant national media, well-manned and well-resourced Internet manipulation capabilities, and even fake civil society institutions that can be used to reinforce the official government positions and lend credence to their views. This then manifests itself as a veritable tsunami of fake news, influence and information operations that can swing opinion both within the target state and externally as well, increasing the pressure on the target.

a. Advance or undermine policies within target State

10. **Mr Jakub Janda** cited the example of the Czech Republic, where one-quarter to one-third of the Czech population believed that Ukraine had a fascist government, as a consequence of disinformation attributed to a foreign State. He noted that this made it almost impossible for the Czech government to aid Ukraine, such as by providing humanitarian aid. He explained that a foreign power could target the weak points within a society (for example, level of support for leaving the EU or NATO) and support local extremists; this in turn could damage the foreign policy options of the targeted country.

11. Similarly, **Dr Elmie Nekmat**, an Assistant Professor at NUS's Department of Communications and New Media, observed that disinformation campaigns are sometimes aimed at influencing public debates on domestic policies. Dr Elmie noted that between 2015 and 2017, 9,097 posts linked to an agency with links to a foreign State, were found to have manipulated Americans' opinions about pipelines, fossil fuels, fracking, and climate change. These posts adopted conservative positions, supported activist groups to stir up tensions and skewed public policy debates in the US.

b. Discredit public institutions and/or leaders

12. Disinformation expert **Mr Ben Nimmo** gave a few examples of how disinformation was used to discredit public institutions and leaders.

13. The first example was the "Lisa" case, where a girl of a foreign ethnicity falsely claimed that she was kidnapped and assaulted by men of Middle-Eastern descent

in Germany. Even though the German police investigated the girl's claims and found that she had fabricated them, the media of a foreign State continued to publicise the girl's claims, and alleged that the German police were part of a cover-up. This led to anti-immigration demonstrations in Germany. The false claims of a cover-up were even echoed by the Foreign Minister of the foreign State, which prompted the German Foreign Ministry to intervene.

14. The second example was about the Internet Research Agency (IRA), a "troll factory" with links to a foreign State. Mr Nimmo noted that the IRA put a lot of effort into widening the divide between the Black Lives Matter movement and the police. The purpose, according to Mr Nimmo, was two-fold – first, to widen the divide between the Africa-American community and the police; and second, to attack the institution of the police. The posts by the IRA pushed both sides. The IRA ran accounts in favour of the Black Lives Matter movement, as well as accounts in favour of the police and the right to shoot. Mr Nimmo noted that one of the very first deliberate fakes published by the IRA was on 13 December 2014, and it was a fake video which purported to show the moment an African-American woman was shot by a policeman in Atlanta, Georgia.

15. Mr Nimmo also noted that in the context of the 2016 US Presidential Election, the IRA posted content that lauded then-candidate Donald Trump and demonised candidate Hillary Clinton. One such account on Twitter, named "Jenna Abrams", amassed over 70,000 followers and was quoted by dozens of high-profile media outlets. Another such Twitter account had over 130,000 followers and its posts were retweeted by senior members of the Trump campaign. Mr Nimmo observed that the purpose of such accounts was to create division and handicap one candidate in the election.

16. In relation to foreign-sourced deliberate online misinformation which was spread during the 2016 US Presidential Election, **Mr Andrew Loh** noted that such misinformation has given rise to concerns about the integrity of public institutions and democratic processes in the US and other countries. He said on a broader level, this shows that even nations as powerful and resourceful as the US are not immune to attempts to sabotage democratic processes and institutions. This view was shared by **Dr Liew Kai Khiun**.

17. In Ukraine, **Mr Ruslan Deynychenko** described how news sources from a foreign State had spread disinformation that tens of thousands of Ukrainians had to seek asylum in a foreign State due to persecution by the Ukrainian government. He said disinformation about atrocities committed by the Ukrainian government against its own citizens (who ethnically originated from a foreign State), such as the murders of children, pregnant women, and the torture and rape of the civilian population, ended up motivating citizens of that foreign State to fight Ukrainian forces in Eastern Ukraine.

18. **Ms Nataliia Popovych and Mr Oleksiy Makhuhin** also described how disinformation from a foreign State had targeted the Ukrainian Armed Forces. For example, the State spread the claim that the Ukrainian Army's leadership is weak and that Ukrainian President and his generals are traitors. As a result, 62% of the media coverage of the Ukrainian military leadership was negative, and trust in the Army deteriorated as a result.

19. **Mr Jakub Janda** noted that disinformation operations often have the goal of undermining public trust towards democratic institutions, and causing the public to lose trust in institutions like the free media and democratic political parties. **Dr Janis Berzins** said that one of the objectives to be achieved by influence operations was to incite mass panic and to create a loss of confidence in key government institutions. **Dr Shashi Jayakumar** said that spreading rumours to discredit politicians, and playing up themes like the negative portrayal of immigration policy have been aimed at undermining public trust towards democracy, and systematically influencing populations to become less trusting of mainstream, established news networks and more trusting of fringe news sources (backed by foreign powers) and conspiracy narratives.

c. Achieve election outcome or sway opinion

20. The **National Council of Churches of Singapore** submitted that "fake news" can be used by a foreign government to interfere with the domestic affairs or elections of another country without the inherent repercussions of other means of domestic interference.

21. **Dr Elmie Nekmat** observed that disinformation campaigns tend to be strategically aimed at influencing outcomes by steering public discourse and altering public opinion, and tend to be orchestrated by foreign players with multi-million dollar funded operations.

22. **Dr Hany Farid**, in the context of discussing the technology involved in manipulating online material, also highlighted that there are signs that technology has been developing and the possibility of fakes being used in upcoming national and state elections.

d. Sow discord

23. Representors shared how foreign State actors have sought to sow discord in various societies.

24. **Mr Ruslan Deynychenko** described how foreign-sourced propaganda has targeted the divisions between Ukrainian nationals speaking a foreign language and the Ukrainian-speaking Ukrainians, by looking for and playing up examples of past conflicts. **Ms Nataliia Popovych** and **Mr Oleksiy Makhuhin**,

representors of UCMC, noted that disinformation campaigns had targeted groups in Ukraine based on basic and rooted characteristics like nationality, age, sex, church, race, language and income.

25. **Dr Shashi Jayakumar**, Head of the Centre of Excellence for National Security at RSIS, noted that rumours and untruths carried by bots and fake ads had supported and inflamed all sides of the political spectrum in the US. He quoted an observation that foreign-linked bots and trolls did not care about the causes that their falsehoods promoted, as long as they "foment[ed] division and chaos." [1] He said that some disinformation campaigns were used not so much to strengthen any one cause, but to create dissension and undermine the resilience of the polity. **Mr Ben Nimmo** echoed this perspective by noting that the efforts to create dissension would come in the form of posts being published both in favour of and against a certain policy. Specifically, they targeted existing divisive issues, such as race, LGBT rights, gun control and immigration.

26. **Dr Carol Soon** and **Mr Shawn Goh** also noted how deliberate online falsehoods were able to stoke already high tensions during the election period in the US, by targeting citizens who rallied behind different party lines. A representative of **QSearch**, a social media analytics company, shared how African-American martial arts instructors were paid by a foreign State to promote self-defence classes to African-Americans in swing states, and these footages from those classes were then used as "proof" that the African-American community was arming itself. Another example given by **Mr Ben Nimmo** was of a video that was shared by users, falsely purporting to show an African-American woman being shot by a policeman in Atlanta, Georgia. According to Mr Ben Nimmo, the video was spread by a foreign troll factory, and its purpose was to widen the divide between the African-American community and the police, as well as to undermine the institution of the police.

27. Mr Ben Nimmo and Dr Shashi Jayakumar both described how foreign disinformation trolls even succeeded in pitting two communities against each other in Texas, by using different Facebook groups to organise both a protest and counter-protest to take place at the same time and place, which eventually happened.

28. Such examples could be found in the United Kingdom as well. **Mr Thiruprakassh S/O Suppiah** shared how, soon after the London terror attacks of 2017, social bots controlled by foreigners spread a post containing a picture of a Muslim woman, claiming that she was walking past a dying man. It transpired that the picture was taken out of context and he noted that the post, which contained hashtags such as "BanIslam", was spread with the intention of turning public opinion against the Muslim community.

[1] Shashi Jayakumar, Appendix III: Written Representations, Paper No.59, page B330.

e. Cause alarm

29. **Mr Ben Nimmo** explained how deliberate online falsehoods can be used by State actors to spread alarm amongst the public.

30. For example, he described a claim by a foreign State that it had an electronic bomb that can disable a US warship. This was subsequently picked up by Western media outlets and spread widely. It was eventually revealed that the claim was false. Another example given by Mr Nimmo is the report of the Donbass News Agency dated 4 Jan 2017, which said that the US was sending 3,600 tanks against a foreign State. This was false, with a headline that was deliberately alarmist.

(2) Non-State Actors

31. Various representors discussed how falsehoods could be spread by both local and foreign non-State actors. They identified several objectives for such actions.

a. Advance or undermine policies

32. Similar to the motivations of State actors, local and foreign non-State actors can spread falsehoods to pursue certain policy agendas. Senior Research Fellow from the Institute of Policy Studies, **Dr Carol Soon**, and Research Assistant **Mr Shawn Goh** gave examples of how deliberate online falsehoods around the world often "mirror the cracks and fissures that pervade each country."[2] Those involving local actors included alt-right communities in France and Germany, who spread anti-immigrant falsehoods that exploited divides between citizens and the immigrant population resulting from Europe's immigration crisis.

33. Dr Soon and Mr Goh noted that anti-immigration falsehoods have focused on the impetus to preserve the traditional French character and nation, and identified migrants as a threat to the French way of life. An example of an anti-immigration falsehood was a report by an influential French far-right opinion website on 5 April 2017, claiming that the Breton lighthouse in Paris would be demolished to provide housing for migrants. In terms of reach, the website received about 1.6 million engagements from Facebook, Twitter and other social media platforms in a 2-month period, between 5 March and 5 May 2017.

34. In relation to Germany, political data scientists with the Technical University of Munich, **Dr Simon Hegelich** and **Mr Morteza Shahrezaye** explained how online attempts at manipulating public opinion may have affected the public debate about the refugee situation in Germany. They observed that people from the political right in Germany had used "all kinds of online manipulation techniques"[3] to create

[2] Carol Soon and Shawn Goh, Appendix III: Written Representations, Paper No. 62, page B359, para 8.
[3] Simon Hegelich and Morteza Shahrezaye, Appendix III: Written Representations, Paper No. 74, page B443.

a negative trend whereby social media platforms were flooded with negative comments. While cautioning that they could not prove this, they raised the possibility that politicians may have been taken in by the false impression of public opinion online, and be led to make decisions based on this false impression.

b. Discredit public institutions and/or leaders

35. **Mr Ben Nimmo** explained that the "alt-right" movement in the United States had similarly spread false stories as well, driven by the political desire to harm the then-candidate Hillary Clinton's campaign and boost Donald Trump's campaign. During the hearing, Mr Nimmo delved deeper into the alt-right's motivations, explaining that these online activists believed that they could gain more power and more influence by spreading false stories online. He stated that there was indicative evidence that a lot of the people who were sharing some of the false stories about Hillary Clinton did not actually believe them, but hoped other people would.

36. Mr Nimmo described how the alt-right in the US were behind a number of major false narratives during the 2016 US Presidential Election campaign. This included "Pizza-gate", which was a conspiracy theory claiming that Hillary Clinton had been complicit in a paedophile ring managed from a pizza restaurant in Washington, DC. The conspiracy theory prompted an American citizen to bring a gun to the pizza restaurant to investigate the claims. Another example given by Mr Nimmo was the claim that Hillary Clinton's adviser, Sidney Blumenthal, had blamed her for the death of US diplomats in Benghazi, Libya. The truth was that Blumenthal had merely been referring to a Newsweek article that had made that claim.

37. Mr Nimmo's written representation also enclosed an investigative article he had authored about false stories circulating during Catalonia's contested independence referendum on 1 October 2017. That event exemplified the spread of falsehoods in a highly tense situation to turn people against opposing groups. Falsehoods were spread to promote and oppose both sides, namely, the Spanish police seeking to block voting in what Spain claimed to be an illegal referendum, and Catalonians seeking to vote on independence.

38. While some images sought to play up police violence, others attempted to play down police violence and force, or to play up violence by demonstrators. For example, one online article claimed that an old woman who had been forcibly carried away by riot police was a supporter of a Basque separatist regarded by some to be a terrorist, when she was not. The false story had sought to dilute sympathy for the old woman. There were also claims that images were fake, when they were in fact genuine. Another Facebook post accused demonstrators of attacking a policeman, using a falsely captioned image. Mr Nimmo observed that the many fake images of police violence had undermined genuine evidence of the use of force by the Spanish police.

39. Mr Nimmo also discussed how there was a large-scale fake news propaganda campaign in South Africa, aided by a PR agency, which was described as a "hateful and divisive campaign to divide South Africa along the lines of race".[4] The African Network of Centres for Investigative Reporting estimated that "the network of fake news produced at least 220,000 tweets and hundreds of Facebook posts to confuse the public between July 2016 and July 2017".[5]

40. RSIS Associate Research Fellow, **Ms Jennifer Yang Hui** explained that in Indonesia, domestic politically-motivated misinformation campaigns, also termed online hoax campaigns or "e-hoax" campaigns, "presented the greatest concern to the nation's stability."[6]

41. One of the objectives of these hoax campaigns was to undermine the credibility of political figures. These campaigns were prevalent during elections, and sought to achieve a particular election outcome. Ms Yang described how they had affected the standing of electoral candidates in several high-profile elections since 2012. For example, during the 2012 gubernatorial election in Jakarta, political candidates Mr Joko Widodo and Mr Basuki Tjahaja Purnama (popularly known as "Ahok"), faced "black campaigns" that sought to paint them as communists, foreigners, proselytisers and so on. The campaigns intensified during the 2014 Indonesian Presidential Election, when Mr Widodo successfully ran for President. According to Ms Yang, the campaigns also polarised public opinion against "Ahok" during the 2017 gubernatorial election in Jakarta, which he lost.

42. Ms Yang highlighted that the hoax campaigns commonly used sectarian and racist falsehoods, and played on ethnic and religious sentiments. For example, during election campaigns in 2012 and 2014, online hoaxes cast doubt on Mr Widodo's Javanese Muslim identity, and falsely claimed that he and his family were Chinese Christians. Ms Yang explained that such characterisations in Indonesia could have the effect of dissuading voters. She stated that such falsehoods that were organised along racial, cultural and religious lines were designed to elicit emotions.

43. Ms Yang observed that false claims about a candidate's political affiliations would be a tactic of choice of "fake news mills" in future elections. She cited a survey as showing that although Mr Widodo is dominating online conversation as the most high-profile candidate for the 2019 presidential election, his name is also negatively linked to communism. "Ahok" and a Chinese lady mayoral candidate during regional elections in West Kalimantan in February 2017 had also been falsely labelled as communist. According to Ms Yang, such false labels of communist affiliation have been a "go-to method for political gain"[7] for decades in Indonesia.

---

[4] Ben Nimmo, WR, Appendix, "#ElectionWatch: American Bots in South Africa", p 3.
[5] Ben Nimmo, WR, Appendix, "#ElectionWatch: American Bots in South Africa", p 3.
[6] Jennifer Yang Hui, Appendix III: Written Representations, Paper No. 82, page B505, para 6.
[7] Jennifer Yang Hui, Appendix III: Written Representations, Paper No. 82, page B507, para 12.

44. In the same vein, a **group of SMU students** noted that character assassination may be a possible motivation for falsehoods in the context of politics. They explained that one could utilise the spread of falsehoods to target politicians or a political party and undermine its integrity.

c. Turn one group against another

45. **Dr Cherian George** noted that hate propaganda that seeks to vilify one group in society always involves disinformation, and is in fact a political strategy. He said that the messages in such disinformation campaigns vary in their degrees of falsehood and provocativeness, and that many of the statements, when viewed in isolation, may be factual and seemingly innocuous. They are used to foster solidarity and maintain a community that is in a constant state of anxiety and fear. Complementary messages, which may or may not be truthful, take the next step of directing that fear against a target group. He observed that once these ways of thinking are deeply entrenched, it does not take much to tip the balance towards the promotion of intolerance and hate towards a certain group.

46. Dr Cherian also noted that hate campaigns involve a division of labour. The leaders often keep their hands clean, as they can issue clear signals to their followers via silent assent or subtle "dog whistles" and yet evade legal accountability. The followers know what these leaders mean, even if it is not explicit, because of complementary messages from others in the network, who make more explicitly hateful remarks. He also said that hate networks may include think-tanks and experts who pump out pseudo-intellectual and pseudo-scientific arguments to support the movement's grand narrative, as well as media owned by the organisations' own outlets and sympathetic mainstream media.

47. **Dr Alan Chong**, Associate Professor at the S. Rajaratnam School of International Studies, highlighted how propaganda has been used to divide a target population by sowing doubt and tension. He stated that there could be "any number of ethnic, religious or ideological features susceptible to such campaigns of paralysis."[8] An example he gave was a serious number of incidents in India in 2012, where false images of Muslims being attacked were spread online, increasing panic and leading to imitation attacks. It led to tens of thousands leaving the cities of Bangalore, Pune and Chennai.

48. Dr Chong elaborated that the example showed the lethality of "spontaneous, untutored 'citizen journalism' that can be unleashed through social media," [9] and how innocent citizen journalists passing on what they thought was a public security warning could lead to mob-level panic that severely damaged multi-ethnic and multi-religious societies.

[8] Alan Chong, Appendix III: Written Representations, Paper No. 91, page B906, para 16.
[9] Alan Chong, Appendix III: Written Representations, Paper No. 91, page B901, para 7.

49. **Ms Jennifer Yang** noted that an objective of hoaxes in Indonesia was to turn people against the ethnic Chinese. She said that Islamists in Indonesia had conflated China's economic and political rise internationally with the position of ethnic Chinese in Indonesia. This had been described by the *South China Morning Post* as "producing a toxic mash that threatens to undermine social stability in the country."[10] For example, after four Chinese nationals were arrested for planting bacteria-contaminated chili seeds, falsehoods were spread by the media accusing China of deploying biological weapons against Indonesia. This prompted the Chinese embassy in Indonesia to express concern over the online anti-Chinese sentiment that followed.

50. **Dr Carol Soon** and **Mr Shawn Goh** described how there have been anti-Islam falsehoods in various countries, which have linked Muslims and Islam with terrorism and instability. For example, in the aftermath of the terrorist attack in Paris in 2017, a far-right political leader in the UK posted a video on Twitter, and described it as showing Muslims celebrating the attack.[11] It was in fact a video of people celebrating a cricket match victory in Pakistan. The video gained nearly 500,000 views in a matter of hours.

51. **Dr Mathew Mathews**, Senior Research Fellow at the Institute of Policy Studies, described how deliberate online falsehoods were spread by *The Real Singapore* website to inflame racial and religious tensions in Singapore. One example was a post in 2015, claiming that a Filipino family had complained about some Singaporeans playing musical instruments during the annual Thaipusam procession, and that this had led to a commotion between Hindu participants and the police. It turned out that there had been no such complaint by a Filipino family.

52. Dr Mathews recounted how he personally witnessed how quickly netizens took to the story, without questioning the veracity of the facts, and how they made comments maligning Filipinos. In his view, the distorted article undoubtedly would have shaped the opinions of some Singaporeans towards immigrants, Hindus, and an important religious festival in Singapore. During the hearing, he further explained that this was an example of how an event was used to stoke existing antagonistic views shown by a website toward certain groups, and increased the up-take of those views.

53. Entrepreneur **Mr Hazrul Jamari** described how falsehoods have been spread among the Malay community in Singapore to pit one group against another. Examples included content and videos from Syria on Facebook and WhatsApp, pitting Shias against the Sunnis. He was of the view that tensions between the Sunnis and Shias had been worsened by the spread of these online posts.

[10] Jennifer Yang Hui, Appendix III: Written Representations, Paper No. 82, page B507, para 11.
[11] Matt Novak, "This Video of 'Muslims Celebrating the Paris Terror Attack' Is Totally Fake", *Gizmodo* (21 April 2017).

54. NUS law under-graduates **Mr Cheah Wenjie and Mr Chester Su** stated that online falsehoods could be "used to spread racially and religiously contentious viewpoints or teachings,"[12] which may have a negative impact on Singapore's social harmony.

d. Promote or oppose policies or ideological beliefs

55. **Ms Yvonne Wong** identified the "most lethal"[13] objective of online falsehoods to be the desire to influence others to share similar views such as ideology on politics, economics, religion, nationalist, environment, culture and terrorism. This was because the motivational drive of the actors was high, their actions were deliberate, and they reinforced their narratives across time.

56. *Ideologically driven individuals.* **Dr Shashi Jayakumar** shared how disinformation campaigns against a country are sometimes carried out by individuals who feel a strong ideological impetus – they feel the target country has gone down the wrong path (*e.g.* in respect of multiculturalism, or immigration and refugee policy) and feel that by spreading deliberate online falsehoods, they are part of a legitimate resistance that would bring that country to the right path again.

57. **Dr Elmie Nekmat** observed that an analysis of nearly 17 million Twitter posts shared within 10 days of the 2017 French Presidential election showed that the user accounts that engaged with "MacronLeaks" mostly belonged to foreigners with pre-existing interest in alt-right topics and alternative mews media, rather than French users with diverse political views.

58. Lawyer **Mr Darius Lee** was of the view that the local case of *The Real Singapore* could be characterised as one where falsehoods were fabricated for ideological ends by certain individuals, whether directly or otherwise, in order to promote and stoke feelings of xenophobia and racism.

59. Mr Lee gave another example of a falsehood spread by an individual for political or ideological ends. This was the case of a man who edited a picture of a news article about the Singapore Court of Appeal's decision concerning the leaders of City Harvest Church. The edit suggested that the Court had ruled in favour of the accused persons because one of them was represented by a lawyer who was also a Member of Parliament from the People's Action Party.

60. *Ideologically driven organisations.* Nanyang Polytechnic lecturer **Mr Zheng Liren** and his students described how certain organisations had the objective of convincing their readers to hold a certain political belief. These organisations used falsehoods that made their readers feel personally threatened by another group. They sought to obtain power and influence by manipulating and polarising those who did not hold the same political beliefs. An example of this was the Brexit

[12] Cheah Wenjie and Chester Su, Appendix III: Written Representations, Paper No. 132, page B1156.
[13] Yvonne Wong, Appendix III: Written Representations, Paper No. 11, page B23.

referendum in the United Kingdom, where falsehoods were widely spread to fuel xenophobia and anti-immigrant sentiments. They observed that people sought to "gratify" ideological beliefs by getting others to concur. This took the form of "echo chambers" of newspapers and media sites that tended to be "extremely polar and lacking dialectic." [14]

61. Lawyer **Mr Zhulkarnain Abdul Rahim** raised the need to be mindful of the power of lobbyists with political or commercial objectives, who may fund the spread of falsehoods online through social media influencers. He referred to the example of conspiracy videos that circulated on social media in the aftermath of the shooting at a high school in Parkland, Florida, spreading the falsehood that a 17-year old survivor, David Hogg, was a "crisis actor". This was apparently to shore up support for gun rights.

62. **NGO Monitor** highlighted how political advocacy non-governmental organisations (NGOs) create an exaggerated or controversial image of expertise by distorting international and human rights practices. NGO Monitor also appended a report they prepared on how Human Rights Watch's campaigns and publications reflect consistent bias, false and contradictory statements, irrelevant evidence and inappropriate methodologies – all in a bid to support an ideological conclusion.

63. A similar view was shared by the **PAP Policy Forum (PPF)**, which specifically described how the Human Rights Watch's report "Kill the Chickens and Scare the Monkeys" presented facts in a selective manner to create a false and misleading impression of the Singapore Government. The PPF voiced concerns about the lack of transparency behind Human Rights Watch's funding, its links to the US foreign policy establishment, and whether these might affect its agenda and operations.

64. **Dr Shashi Jayakumar** also shared about the activities of an online "army" of content creators based in an Asian country, whose role is to promote their government's policies and attack criticisms of those policies, both within and outside their own country.

e. Reap financial gain

65. *Individuals or entities that spread politically charged or sensational falsehoods for financial gains.* **Mr Ben Nimmo** spoke about the economy of falsehoods, how falsehoods were designed to appeal to people in order to attract "clicks" which was revenue generating, and how these falsehoods usually have political implications.

66. On the economy of falsehoods, Mr Nimmo explained that the aim of falsehoods created for money was to attract internet users to advertisements by using

[14] Zheng Liren et al., Appendix III: Written Representations, Paper No. 60, page B342.

sensational, emotional or divisive content, also known as "click-bait." He elaborated that these purveyors of falsehoods used the advertising systems of Google and Facebook, for the sole purpose of generating advertising clicks, as these clicks would generate income for them.

67. According to Mr Nimmo, these people would see what stories spread best, then try to replicate them. He explained how some of what he termed "fake news merchants" began by writing positive stories about Hillary Clinton, and realised that no one was sharing them. They then began writing negative stories about Hillary Clinton, and realised those stories were being spread a lot more. He described this as a "black market economy."

68. For example, Mr Nimmo shared that during the 2016 US Election, one Mr Paul Horner claimed to make US$10,000 per month writing false stories which were politically charged. Mr Nimmo, **Mr Norman Vasu** and other representors also shared about how teenagers in a small Macedonian town had created fabricated and highly partisan "news" stories during the US Presidential elections to earn money from advertising. **Mr Carlos Nicholas Fernandes** wrote that one of the teenagers reportedly earned $16,000 in ad-revenue from two pro-Trump websites, which is many times the average monthly salary in Macedonia (*i.e.* $371).

69. **Dr Shashi Jayakumar** observed that there were individual "consultants" and private sector entities that specialised in hacking or interfering with elections with the aim of achieving the desired election result for their client, and would charge a fee for it. He described how one Andreas Sepulveda, who was very connected with leaders in Latin American countries, would rig elections in Latin America for the highest bidder, such as a politician who wanted to get elected. Sepulveda's methods included hacking, smear campaigns, and technical disinformation and subversion. Dr Shashi said that Sepulveda was a well-known South American case, and that smart persons who do the kind of work Sepulveda does would not be as flamboyant or high profile as him. He said they will be more discrete in their methods but will be known to the people who matter.

70. Dr Shashi also noted that there appears to exist a growing shadow market for methods to influence target populations and outcomes in nations, using methods like those offered by Cambridge Analytica, which is reported to have profiled and micro-targeted the US electorate during the 2016 US Presidential Election.

71. **Mr Septiaji Eko Nugroho** described how a person in Indonesia running several websites that spread disinformation had claimed, on live TV, that he could earn 300-500 million rupiah (approximately US$20,000 – US$35,000) per month from advertisements on his websites. This person had said that he did not care about the nature of the information spread through his websites, and that he will produce any information – including falsehoods – so long as he can "clickbait" people.

72. Similarly, **Ms Jennifer Yang** shared how "fake news factories" were proliferating in Indonesia. The Indonesian National Police have reportedly found that there were many such organisations seeking monetary gains in exchange for creating online falsehoods. The Indonesian Ministry of Communication and Information reportedly found around 800,000 websites that disseminated fake news.

73. Ms Yang and Dr Shashi also highlighted the example of the Saracen Cyber Team in Indonesia, an online-based syndicate that created many social media accounts to spread hate speech for clients that are willing to pay for them. During the hearing, Ms Yang explained how the organisation began in 2014, during the Presidential Election between current President, Mr Widodo, and his rival, Mr Prabowo Subianto. The founder of Saracen was a supporter of Mr Subianto. He would hack into, and take over rivals' social media accounts. Subsequently, he would put up content that denigrated race and religion, as a way of penalising them for supporting the political opposition. Over time, he began selling the social media accounts for money. From investigations, it appeared that some of the accounts were being used to falsely portray Mr Widodo as having a certain ethnic lineage or political leaning. Dr Shashi pointed out that this business was a lucrative one, as one estimate suggested that a single popular post on Saracen could rake in Rp 100 million (US$7,500) because of the wide reach of the site.

74. Another financially-driven group identified by Ms Yang in Indonesia were online influencers who promoted businesses and political causes. She described them as common in Indonesia, and was of the view that they indirectly contributed to sensationalised information. She explained that these online influencers comprised "buzzers" and "micro-celebrities". "Buzzers" were Twitter users with more than 2,000 followers, who were paid to send short, personalised messages to potential customers during rush hour, when they would be caught in traffic and be absorbed in using their smartphones. "Micro-celebrities" were social media celebrities who used online platforms to attract attention to their political causes. According to Ms Yang, these online influencers were hired by political candidates during the 2017 Jakarta gubernatorial elections, and tended to promote messages that benefited their financiers rather than factually accurate information.

75. Locally, **Mr Zhulkarnain Abdul Rahim** noted how online publishers were driven to earn advertising dollars by preferring low quality content over journalistic and verified content. He referred to the experience of the founders of the *The Real Singapore* website, who had been charged under the Sedition Act. He noted reports that they had earned almost S$500,000 in online advertising earnings. In the four months before the site closed down, they were earning about S$42,000 each month, and at their peak, earned almost S$55,000 in one month. He also observed that to readers of the site, the site's lucrative aspect in terms of advertising dollars was not readily apparent.

76. Mr Zhulkarnain elaborated on the motivations of the founders of the site. In an interview, they had claimed their original objective was to bring more freedom of

speech to Singaporeans. However, one of them admitted that making money eventually became a key focus of his, especially with the pressures of paying for his university studies when his parents were unable to pay the whole sum.

77. **Mr Darius Lee** also referred to the case of *The Real Singapore*. He noted the observation of the District Judge who had decided on the sentence of the website's co-founder, that at the heart of the case was "the exploitation of [feelings of xenophobia and racism] purely for financial gain."[15] Similarly, Ms Soon and Mr Goh also pointed out how *The Real Singapore* had used sensational articles to draw readership, for financial gain.

78. *Corporations and businesses.* In a literature review conducted by **Ms Carol Soon** and **Mr Shawn Goh,** they observed that companies may also have a financial interest in spreading deliberate falsehoods. For example, the tobacco industry has published research to counter scientific evidence that linked smoking to cancer. This sought to sow confusion about the truth, and could have public health consequences.

79. **Mr Sui Yi Siong** and students from the SMU School of Law explained that the publication of deliberate falsehoods for financial profit "has a long and storied past", and was known as "yellow journalism" in the 19th century. Supermarket tabloids had "long trafficked in a mix of partially true and outright false stories."[16] In this regard, they also referred to how the tobacco industry has challenged scientific evidence on the link between smoking and lung cancer as an example of financing by large corporations of suspect research to ensure their dominance. Another example was fossil fuel manufacturers funding research attributing climate change to natural causes rather than human activities or carbon emissions.

80. Representatives from **TrendMicro** also submitted that it was not uncommon for companies to seek to undermine their competitors using hoaxes and smear campaigns. An example given was how a company can spread false negative comments about a competitor to rake in more businesses, with several cases of this nature found in New Zealand.

81. The distortion of data by businesses was flagged by student and writer **Mr Jev Akshay**. Mr Akshay gave the example of a British study that appeared to show that smokers had a higher survival rate when compared to non-smokers. However, this turned out to be because the non-smokers selected were significantly older on average and therefore more likely to pass away during the duration of the study. Mr Akshay stated that this was an example of the deliberate omission or hiding of key variables by news sources, so that the data could be brought in line with their own agendas. He described this as enriching corporate executives at the expense of the consumer.

[15] Darius Lee, Appendix III: Written Representations, Paper No. 32, page B108, para 17.
[16] Sui Yi Siong et al., Appendix III: Written Representations, Paper No. 130, page B1134, para 13.

82. Mr Akshay also singled out news networks and agencies, including citizen journalists, who sought to increase their profits, ratings and profile through sensational reporting. This meant shifting "their focus towards capitalising on the pathos of the public so as to incite a reaction from them," and used "an altered reality" to make a story more interesting.[17] Similarly, **Mr Zheng Liren** and a group of students from Nanyang Polytechnic, citing a research study, described how headlines often used "clickbait" and buzzwords to appeal to the emotions of readers in order to drive traffic towards their publications, in order to gain advertising revenue. They also identified a hunger for power as an incentive for the use of falsehoods to exploit audiences and the truth. Citing an article titled *'Fake news' – why people believe it and what can be done to counter it*, they described how mass media publishers were incentivised to compromise the truth in order to draw income, and stated that the "economics of social media favour gossip, novelty, speed and 'shareability,' not truth."[18] **Dr Cherian George** noted that some media run stories based on hate propaganda solely with audience numbers in mind, purely for commercial benefit and recklessly disregarding whether the content is true or not. He observed that shareability or clickworthiness is prized over trustworthiness.

83. **Ms Yvonne Wong's** observation was that spreading malicious falsehoods for monetary rewards was becoming increasingly more common in businesses with intense competition. **Mr Raja Mohan** noted that private entities can now attain financial benefits from the online advertisements posted on sites on which they post online falsehoods. **Mr Nicholas Fang** and **Dr Lim Sun Sun** were of the view that the current business models of social media platforms, small or alternative media sites, and advertisers are partly the cause for the situation we are in today.

f. Mischief and other objectives

84. **Mr Ben Nimmo** noted that many fakes online were originally efforts at mischief-making. He gave the example of a forged letter, purporting to expose connections between Britain's GCHQ intelligence agency and the Obama administration to spy on then-candidate Mr Donald Trump. This forged letter was first posted to 4chan in June 2017. Even though it was exposed as a fake by some users, others suggested sending it to broadcasters anyway "for the lulz" (*i.e.* for entertainment). This forged letter continued to circulate online as a genuine document well into 2018, despite being debunked multiple times.

85. **Ms Soon and Mr Goh** similarly identified mischief as one of the motivators of falsehoods in Singapore. An example of this was the fake announcement on the passing of Mr Lee Kuan Yew by a young Singaporean student, who wanted to show how easy it was to perpetuate a hoax. There was also a case of a Singaporean man who had doctored the headline of a news article relating to the City Harvest Church case because he was dissatisfied with the outcome of the trial.

[17] Jev Akshay, Appendix III: Written Representations, Paper No. 50, page B223.
[18] Zheng Liren et al., Appendix III: Written Representations, Paper No. 60, page B342.

86. **Ms Yvonne Wong** observed that some may spread falsehoods for the thrill of the ability to influence others. She referred to the example of a website called check4spam.com, which published online conversations on spam circulating in different countries. According to Ms Wong, some of their sentiments revealed that the creators of the spam wanted to "show off their ingenuity"[19] that had caused the wide spread of the spam they had created.

87. **Ms Myla Pilao** gave the example of an individual from London who set up a fake restaurant at the back of his home called "The Shed" as part of a social experiment. He promoted "The Shed" online, using various digital technologies such as click farms, click-baits, fake pictures and fake reviews. Within six months, "The Shed" became the top-ranking restaurant on TripAdvisor.

88. Some falsehoods have been spread to injure the reputation of others. **Mr Ngoh Wang Long** highlighted his personal experience as a subject of "an inaccurate account of events" on Facebook and WhatsApp. According to him, the intention was to "teach [him] a lesson."[20] This led to insults being directed at him. **Ms Jennifer Yang** also noted that many ordinary Indonesians who share fake content online may be doing so due to intrinsic motivation such as genuine belief in the content as well as enjoyment of the content itself.

89. People of all age groups can share falsehoods for a variety of reasons, mischief or otherwise. **Mr Goh Sin Teck** noted that many senior citizens in Singapore sometimes spread falsehoods via social media because they cannot tell whether the news they received is true or false, and believe that it is likely to be true if it came from someone they know. Student **Mr Zubin Jain** drew on his experience as a teenager to share how deliberate falsehoods were spread by teenagers to generate profit or attention. His own motivation for having posted false information in the past was to alleviate boredom.

g. Combined objectives

90. **Ms Jennifer Yang** explained how a single actor who spreads disinformation could have both political and financial objectives. For example, the founder of Saracen had started his online disinformation activities to achieve the political objective of undermining a political candidate. He later sought to achieve financial objectives as well, by receiving monies from undisclosed "high profile individuals" suspected of paying his Saracen Cyber Team to create and disseminate fake news.

91. **Dr Cherian George** also said that some media organisations run stories based on hate propaganda for commercial benefit, noting that even though they have no ideological links to those generating this propaganda, they are united by a common

[19] Yvonne Wong, Appendix III: Written Representations, Paper No. 11, page B23.
[20] Ngoh Wang Long, Appendix III: Written Representations, Paper No. 128, page B1113.

methodology of preying on people's fears and prejudices with simplistic depictions of the world.

92. Similarly, **Mr Ben Nimmo** highlighted that commercial content could have political consequences. Mr Nimmo gave the example of a commercial botnet made up of fake Donald Trump supporters. He noted that the motivation of this botnet was commercial because it was sharing news through a URL shortener which, which clicked, takes one to a paid advertising site. The creators of the botnets however targeted accounts of Donald Trump's supporters because those who supported Donald Trump online were seen as more likely to share inventions or fake news; and so, was a better source of clicks and revenue.

93. The convergence of financial and political objectives could also be seen from syndicates willing to spread political falsehoods for money. This led Ms Yang to observe that "[f]ake news is a spectrum of phenomena … [that] can range from online disinformation campaigns by foreign states to other more benign, but still fictitious content circulating on social media. Far from being static categories, therefore, the fact that fake news represent a range of phenomena means that the categories can and do conflate with one another."[21]

(3) Alignment of Different Actors

94. Several representors shared about how the objectives of different *types of actors* (*i.e.* foreign State actors; and local non-State actors) can overlap in the spread of deliberate online falsehoods, which can cause a falsehood to be amplified further.

95. **Dr Michael Raska** observed that a sophisticated State actor can employ non-State actors as proxies in cyber space and information operations. **Mr Ruslan Deynychenko** gave the example of how right-wing extremist groups from Poland and Ukraine created provocations, burned flags of the neighbouring country, and desecrated monuments and military cemeteries. An investigation subsequently revealed that these groups were organised and financed by a foreign State, in an attempt to instigate conflict between Ukraine and Poland.

96. In the same vein, **Mr Benjamin Ang** noted that State actors can use non-State actors to spread falsehoods. Such non-State actors can include the State-sponsored media of a foreign country, business or clan associations (especially if their members have business in the foreign country), NGOs that may be infiltrated by the foreign country, political parties that may have the same view or have been infiltrated by the foreign country, academics who may be agents of influence for the foreign country, as well as organised or volunteer groups of civilians.

97. **Mr Ben Nimmo** explained how two different groups during the US Presidential Election – the American alt-right and foreign disinformation operatives – had

[21] Jennifer Yang Hui, Appendix III: Written Representations, Paper No. 82, B505, para 4.

allegedly shared the same objective, *i.e.* to denigrate Hillary Clinton, and swing the election in favour of then-candidate Donald Trump. This alignment of objectives was exploited by the foreign disinformation operatives, who, in Mr Nimmo's words, "very successfully"[22] infiltrated the alt-right, and masqueraded as genuine alt-right Americans. These foreign operatives used the same modus operandi as the alt-right – making up false stories about a political candidate, *i.e.* Hillary Clinton, to weaken her, and used the same networks as the real news. In fact, according to Mr Nimmo, even before the 2016 US Presidential Election, these foreign operatives had already exploited the political agendas of the alt-right to pit them against other opposing groups.

98. **Dr Claire Wardle**, Executive Director of First Draft, explained how the objectives of different actors had overlapped in the creation, production and distribution of a false article titled "Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement."[23] The article was created by an unidentified person, and published on a self-proclaimed fantasy news site WTOE5News in July 2016. WTOE5News was part of a network of 43 fake news sites, which earned digital advertising income by generating readership. The article was shared on Facebook by someone working for this network of fake news sites. It was then re-shared by different groups of people, namely, (i) those who sought to amplify the reach of the article to make profit, (ii) Donald Trump supporters, (iii) other forces who had an interest in Donald Trump winning, *e.g.* trolls linked to a foreign State, and (iv) Hillary Clinton's supporters, to show how easily Donald Trump supporters could be fooled.

99. Likewise, **Dr Kevin Limonier**, Associate Professor of the French Institute of Geopolitics and Associate Researcher, Castex Chair of Cyberstrategy, described how the alignment of different actors could facilitate the spread of allegedly false propaganda. Dr Limonier had carried out a preliminary mapping of the "galaxy" of Twitter users who relayed content from two foreign newspapers in France. The mapping showed that their content was able to reach a politically varied audience, comprising not only the French nationalist far-right, but also users sharing different political opinions and of different political leanings.

100. **Mr Jakub Janda** highlighted how national security threats become especially urgent when there is alignment of the following three interests: the domestic economic interest of those who systematically publish disinformation, the domestic political interest of those who share the same views as a foreign state, and the geopolitical interests of the foreign state. He said, for example, that in order to achieve its goals in the Czech Republic, a foreign State had used extremist and fringe politicians in the Czech Republic to help share and spread propaganda and disinformation which was in favour of the foreign State.

[22] Ben Nimmo, Appendix III: Written Representations, Paper No. 36, page B141, para 25.
[23] Craig Silverman and Jeremy Singer-Vine, "The True Story Behind the Biggest Fake News Hit of the Election", *BuzzFeed* (16 December 2016).

ANNEX B: USE OF DIGITAL TECHNOLOGIES TO SPREAD ONLINE FALSEHOODS

1. Disinformation expert **Mr Ben Nimmo** noted that the spread of digital publishing technologies has made it easier to *create* false stories, the Internet has made it easier to *publish* fake stories, and social media has made it easier to *spread* false stories.

2. **Ms Nataliia Popovych** and **Mr Oleksiy Makhuhin**, representatives from the Ukraine Crisis Media Centre, noted that while state-sponsored propaganda and disinformation operations have been in existence for a long time, the difference today is the ease, efficiency, and low cost of such efforts. **Dr Shashi Jayakumar** also noted that the strategies and methods used for mass persuasion are not new, but that the propagation of disinformation now uses new technological tools. Likewise, **Dr Norman Vasu** said that while fake news is not new, the challenge today stems from the fact that information today moves far more rapidly, comes at a greater volume, and reaches more people than ever before. A similar view was expressed by the **PAP Policy Forum**, the **editors of Channel NewsAsia**, **NGO Monitor**, the **National Council of Churches of Singapore**, a group of **SMU students** and **Mr Nicholas Fang**.

3. **Mothership** noted that it is technology that directly drives the information superhighways, and not people. **Associate Professor Eugene Tan** observed that deliberate falsehoods have a "viral" effect because of the relative ease and affordability with which they are transmitted. He said that digital technology has become the viable and preferred proxy for the transmission of deliberate online falsehoods, and that it will continue to be so for the foreseeable future. A similar point was made by **Mr Nicholas Fang**, who warned of a social media "arms race", where all sides seek to outdo one another in developing more and more sophisticated tools that will grant them access to the levers that can influence human behaviours. **Dr Janis Berzins** explained that technology has facilitated the almost absolute freedom of information and led to social media becoming one of the most important sources of information. These conditions, according to Dr Berzins, facilitate the use of information as part of modern warfare.

4. **Associate Professor Eugene Tan** noted that technology, with its ease, speed, and difficulty in tracing, would *exacerbate* the problem of deliberate online falsehoods.

(1) <u>Amplification and targeting of online falsehoods</u>

5. **Mr Benjamin Ang** referenced an RSIS research paper titled "Countering Fake News: A Survey of Recent Global Initiatives", which stated that the impact of fake news is amplified through (a) internet platforms, which publish content with significantly lower cost, wider reach, and rapid circulation, (b) social media, which enables more people and groups of various persuasions to interact even as

they consume, produce and recirculate content, and (c) artificial intelligence agents that automate the work of human propagators.

6. Various representors gave evidence on how online falsehoods are being amplified today.

a. Easy amplification

7. *Social media as a source of information.* Various representors shared that falsehoods can be spread easily on social media today given that it has become a very popular source of information.

8. **Ms Myla Pilao**, Director for Technology Marketing at Trend Micro, shared that most information distribution today goes through the platforms such as Facebook, YouTube, and WhatsApp. The level of engagement on these platforms is very high because users often turn to social media first to get information, rather than go to the original source. Similarly, **Dr Claire Wardle**, the Executive Director of First Draft, noted that news feeds, rather than news websites, are often people's direct connection to the news. **Ms Nataliia Popovych** and **Mr Oleksiy Makhuhin**, also observed that more and more people rely on the Internet and social media as their primary source of news and information.

9. Closer to home, **Mr Chua Jun Hao**, an accountancy student from NTU, highlighted that 85% of Singaporeans get their news online, with the majority getting their news from social media. Similarly, representors from the **Singapore Corporate Counsels Association and the Singapore Press Club (SCCA/SPC)** noted that Singapore has a high mobile phone penetration rate, and cited a report from the Business Times showing that (a) 70% of Singaporeans are active social media users on mobile phones, more than double the global average of 34%, and (b) more than 3 in 4 Singaporeans use social media.[1] **Associate Professor Eugene Tan** noted that social media platforms have now become an important source of news for digital natives like students in institutions of higher learning.

10. **Mr Nicholas Fang** noted that the high levels of Internet penetration and media consumption can be exploited to quickly and widely seed and spread fake news. He said these effects can be amplified if promulgated via closed communication channels such as WhatsApp and Telegram, which are difficult to regulate due to security protocols. **Mothership** also noted that disinformation could spread via WhatsApp chats, and that both Facebook and Google have massively amplified deliberate online falsehoods through crawlers and algorithms that failed to discern fact from fiction.

[1] Singapore Press Club and Singapore Corporate Counsel Association, Appendix III: Written Representations, Paper No. 155, page B1364, para 2.4.

11. **Dr Cherian George**, a Professor of Media Studies and researcher of hate speech, observed that social media platforms are currently "too hospitable to disinformation".[2] Similarly, lawyer **Mr Dan Shefet** agreed that the Internet has created a political ecosystem in which the extreme, the incendiary, and the polarising tend to prevail over the considered, the rational, and the consensus-seeking.

12. **Associate Professor Eugene Tan** also observed that while it would be disingenuous to attribute the rise of deliberate online falsehoods to the advent of social media platforms, the fact that their technology and their platforms can be manipulated and be a threat to democracy cannot be taken lightly. **Dr Carol Soon** and **Mr Shawn Goh** noted that YouTube, Facebook, and WhatsApp – the three most commonly used media platforms in Singapore – are the three more common platforms through which deliberate online falsehoods are disseminated and circulated. Similarly, **Ms Jennifer Yang** said that Facebook, Twitter and Instagram are the most commonly used platforms to spread hoaxes in Indonesia.

13. *Reach of social media platforms.* **Mr Ben Nimmo** noted that the number of platforms and channels by which falsehoods can be spread has increased radically. He highlighted that on Facebook alone, the number of active monthly users grew from 100 million in the third quarter of 2008 to over 2 billion in the fourth quarter of 2017. In a similar vein, **Mr Sui Yi Siong** and students from the SMU School of Law pointed out that on Facebook, an individual can share a post with 5,000 people, and on WhatsApp, each group can have up to 256 members with whom information can be shared.

14. **Ms Myla Pilao** also said that out of the 7.5 billion people worldwide as of 2017, 3.5 billion are Internet users, and 3.03 billion are social media users. She also shared that social media platforms are the most prolific way to distribute falsehoods, because of the large number of users that are actually in there; and that Facebook, being the highest number of active users (2.2 billion) as of January 2018, makes it the most likely platform of choice to launch fake news campaigns.

15. *Speed of dissemination on social media.* **Dr Shashi Jayakumar** noted that the circulation of disinformation is very potent, more so than former times. **Tisane Labs**, an organisation that develops and distributes artificial intelligence software, observed that social media channels optimise delivery of news by demand, using subscription models and automatically promoting popular posts. This allows the viral spread of content that the public finds interesting and relevant, much faster than traditional media ever dreamed of. The editors of **ChannelNewsAsia** referenced a report by MIT on how fake news tends to spread further, faster, and deeper than real news. They said that creating fake news becomes more commercially lucrative as a result.

[2] Cherian George, Appendix III: Written Representations, Paper No. 47, page B201, para 23.

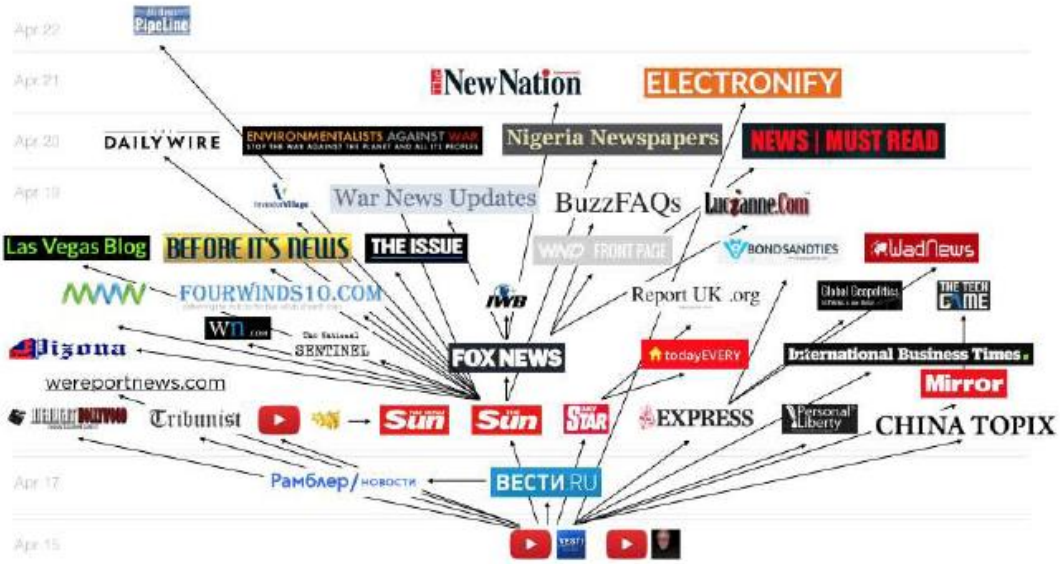
16. On an individual level, **Dr Thio Li-ann** pointed out that anyone who receives information or misinformation can with the click of a button forward it to a large number of recipients, allowing news containing misinformation to go viral and exacerbate the harm caused. **Dr Gillian Koh** also pointed out that communication of any material can spread with much greater ease and speed than before today, as one merely needs a computer or a smart phone to do so.

17. **The National Council of Churches in Singapore** recognised that the social media revolution has enabled the rapid mass dissemination of "fake news" in a way that poses new and serious challenges. **Mr Sui Yi Siong** shared his concern about online falsehoods and the speed of dissemination. He said that as a young person raised in social media, he understood how quickly information can pass and how damaging it can be in that very short period of time it is disseminated.

18. *Private Messaging Applications.* **Ms Jennifer Yang** observed that smartphone-based private chat groups are becoming an important source of information; and by extension, fake news, for many ordinary Indonesians. She quoted a study by the Reuters Institute for the Study of Journalism[3] that WhatsApp has become one of the prevailing ways people discover and discuss news.

19. *Mainstream media.* **Dr Claire Wardle** also recognised that mainstream media may be agents in amplifying (intentionally or not) fabricated or misleading content. On this point, **Mr Ben Nimmo** referred to the "electronic bomb" false story example, where a hoax was spread about a foreign State being able to disable a US warship using electronic jamming. This story was picked up by the mainstream media in various countries, and they all became potent amplifiers of the falsehood. For example, the Fox News report on this was shared over 27,000 times, and a similar report by The Sun was shared over 10,000 times. He used the diagram below to show how mainstream media was used to amplify the false allegation:[4]

[3] Jennifer Yang, Appendix III: Written Representations, Paper No. 82, page B509, para 19.
[4] Ben Nimmo, "Russia's Fake 'Electronic Bomb'", *Digital Forensic Research Lab* (8 May 2017).



b. False Amplification

20. Political data scientists with the Technical University of Munich, **Dr Simon Hegelich** and **Mr Morteza Shahrezaye** commented that in every political discourse, there are manipulative attempts with social bots, trolls, and hyperactive users to create the impression that a particular opinion is popular or unpopular online. They pointed out that as a result of these manipulative attempts, others might fall for these wrong impressions, comment on them, and make them even more popular. Algorithms by social media platforms then pick up on these trends and further amplify this content. As a result, according to Dr Hegelich and Mr Shahrezaye, anyone who is monitoring what is going on on social media might get a wrong impression and make bad decisions. Similarly, **Dr Claire Wardle** shared that agents of disinformation use many forms of manufactured or "false" amplification today, from automated bot networks to groups of people paid to act as bots (cyborgs).

21. **Dr Liew Kai Khiun** noted that "internet trolling" today involves institutionally supported acts of using both human agents (usually operating under multiple online fake accounts) as Internet trolls and Internet bots (software applications that run automated tasks) in infiltrating, inflaming, and overwhelming existing national discussions on social media platforms with intent to sow discord and sway public opinion. **Dr Elmie Nekmat**, an Assistant Professor at NUS's Department of Communications and New Media, observed that 'cyber armies' and 'web brigades' comprising fake accounts, bots, and trolls in social media do three things: (1) induce virality of online falsehoods by 'sharing' disinformation within and across different social media channels, (2) produce faulty perceptions of majority opinion surrounding issues affecting society, and (3) create the illusion

of majority support that can spur actual individual support through a bandwagon effect.

22. **Mr Ben Nimmo** also noted that disinformation agents use multiple platforms, both overt and covert, in a coordinated campaign, to create the impression of a spontaneous movement to cover what is actually an orchestrated campaign. One example of such co-ordinated campaigns was the "Morgan Freeman case". In this case, American actor Morgan Freeman was the voice and face of a video saying that a foreign country, using online activity, had attacked the US during the 2016 presidential campaign. A group of independent activists from the foreign country then launched a counter information operation using the hashtag "#StopMorganLie", which originated from a website run by trolls. The hashtag was then picked up by internet trolls, and amplified by both bots and also the diplomatic missions of the foreign country, before being picked up by the mainstream media of the foreign country. The mainstream media of the foreign country then claimed that the counter-movement was a big Twitter outcry, even though the total Twitter traffic in English in relation to the hashtag was only about 1000 posts. It was later discovered that the accounts of the activists were actually controlled by a troll factory in the foreign country. According to Mr Nimmo, this was a case where "each of these outlets claims to be a separate institution...however, their independence is a façade: on this evidence, they work together to promote a common narrative".[5]

23. *Fake social media accounts.* Representors noted that it is not difficult to set up fake social media accounts, which comprise both troll and bot accounts. RSIS Associate Research Fellow, **Ms Jennifer Yang** described how setting up accounts on Facebook or Twitter was not difficult because the verification procedures are not rigorous. Similarly, **Mr Septiaji Eko Nugroho**, founder of the Indonesian Anti-Hoax Community or MAFINDO, noted that it is not difficult to set up anonymous accounts on the Internet and social media, and that this is exploited by the disinformation ecosystem. **Mr Chui Jian Wei** observed that social media accounts can be created at no cost and without checks on the actual identity of the creator of the account.

24. **Mr Ben Nimmo** elaborated on how fake social media accounts were used by a foreign country to spread disinformation in the US. He said that one troll factory managed at least 3,814 troll accounts and 50,258 bot accounts on Twitter, and 1.4 million Americans are known to have interacted with these accounts in some way. This troll factory also ran at least 470 accounts and spent $100,000 on advertising on Facebook, reaching at least 126 million Americans. **Dr Carol Soon** and **Mr Shawn Goh** also noted that approximately 29 million Americans were directly exposed to 80,000 posts from 120 fake foreign-backed pages.

[5] Ben Nimmo, "Russia's full spectrum propaganda", *Digital Forensic Research Lab* (24 January 2018).

25. Mr Nimmo also noted that fake social media accounts are foot-soldiers in information warfare – they can be used to amplify messaging and force hashtags into the trending lists, or they can even be used to intimidate or block other users. He said that bots and troll accounts can work together to create divisive messages or political messages to push out to the larger ecosystem. For example, during the 2017 French Presidential Election, the #Macronleaks hashtag was used to guide Twitter users to false claims that the emails showed evidence of his offshore accounts, tax evasion and a slew of other nefarious activities. The hashtag was amplified through a network of trolls and bots driven by the alt-right in the US. It reached 47,000 tweets in just three and a half hours after the initial tweet.

26. **Dr Shashi Jayakumar** elaborated on the fake account of one "Jenna Abrams" on Twitter. She appeared to be a normal, likeable, all-American girl, who had right wing or far-right views and a large number of Twitter followers. She induced many people to listen to her and become her follower on Twitter. At one point she had over 70,000 followers, and was quoted by the New York Times, the Washington Post, Breitbart, and other high-profile media outlets. She was able to move sentiment and opinion. However, after the 2016 US Presidential Elections, researchers discovered that "Jenna Abrams" was not real and was, in fact, a creation of a foreign troll factory. Dr Shashi noted that while part of the process involved automation and artificial intelligence, there was human agency at the back end. He said that while "Jenna Abrams" was one such account, the suspicion of people who really know is that there are many more.

27. *Troll accounts.* **Mr Ben Nimmo** explained that in troll accounts, users masquerade as a member of the target population, and try to infiltrate the target population by interacting with genuine members of the community. These troll accounts would be set up across various social media platforms, like Facebook, Instagram, Twitter and WhatsApp. They would interact with leading members of the community by tagging them in posts and hoping that they retweet or share or amplify the post. This would validate the identity of the fake account, and pave the way for other users to interact with the fake account.

28. According to Mr Nimmo, the troll accounts try to build a following in that community. They start off with innocuous posts with very heart-warming and positive messaging, before introducing biased and false information to influence the community or steer it in a particular political direction. Mr Nimmo observed that in the US, such accounts were so successful that they were effectively the spokespeople for the alt-right movement.

29. **Dr Shashi Jayakumar** described how there were fake Facebook groups apparently created in support of Donald Trump which were almost entirely populated by bots, and which leveraged on existing ideological filter bubbles and echo chambers to attract real fans. According to some researchers, many Donald Trump fans were emboldened to declare their support for him by the artificially

created perception of a swell in support for him, and as these originally fake Facebook groups swelled with real accounts, the fake accounts withered away.

30. **Dr Elmie Nekmat** described an example of how troll farms engaged in "audience development" on social media platforms, with operations that began with a few dozen people eventually reaching 150 million people through Facebook and Instagram.

31. *Bots and botnets.* **Dr Kevin Limonier**, Associate Professor of the French Institute of Geopolitics and Associate Researcher, Castex Chair of Cyberstrategy, referred to bots as a kind of "mass information weapon". They create a fake buzz or audience around a particular piece of content, as people are more likely to believe content that they see has been shared and ready by many before them.

32. According to Dr Limonier, bots usually have these elements: (a) their account names are usually random strings of letters and numbers, (b) their profile pictures are usually taken from royalty-free image banks and used simultaneously by many accounts, and (c) their behaviours are monotask – they "like", tweet, or follow other accounts, but rarely all at the same time.

33. **Mr Ben Nimmo** explained that bot accounts are mostly or entirely automated and are used as amplifiers. He quoted a well-known bot user in the American alt-right as saying that if 1000 bots make one retweet each, making a total of 1000 retweets, genuine users will look at it and think it is a credible tweet and are more likely to share the tweet.

34. Mr Nimmo also highlighted the phenomenon of commercial botnets – a network of thousands of bots available for hire. He said that even if a small group cannot create its own botnet, it can rent one. He noted that bot activity is not limited to incidents in the US – there have been reports of troll factory activities in Mexico, Venezuela, and in the dispute between Saudi Arabia and Qatar. **Mr Raja Mohan** cited an article which stated that, from 2013-2017 every third website visitor was a bot attack.

35. Mr Nimmo also observed how the use of bots and botnets can change depending on who they have been hired to serve. For example, during the German election campaign in September 2017, a botnet which had formerly retweeted Russian-language commercial content (such as advertisements for cars, Bitcoin and plastic windows) began retweeting posts supporting the anti-migrant Alternative für Deutschland party. Similarly, during the African National Congress leadership contest in South Africa, a botnet which posed as American and had largely posted commercial content began posting political South African messaging. Mr Nimmo noted that a Russian bot herder interviewed by BuzzFeed news claimed to have made his commercial botnet available to the far right in Germany "for free (mutually beneficial)". This illustrates the murky crossover between bots which

are created for political purposes and bots which are created for commercial uses, and then hired out or otherwise made available to political users.

36. Mr Nimmo also recounted his personal experience facing attacks by botnets. His team had published an article, identifying and explaining the botnet that was used to harass those who research disinformation conducted by a specific foreign country. Those controlling the botnet then created a meme saying that Ben Nimmo was dead, copied it to all his colleagues, and used bots to retweet it 23,000 times to each person. Additionally, the bots targeted anyone who posted the name of Mr Nimmo's team and the words "bot attack", retweeting the post 23,000 times in the first minute. Mr Nimmo managed to kill the bot net by copying the Twitter unit that takes down bots in a post containing the name of his team and the worlds "bot attack". This caused the bots to retweet the post and thus, identify themselves to Twitter. According to Mr Nimmo, this exposed the sheer scale of botnets available for intimidation, as there were over 100,000 faceless bots and tens of thousands of more sophisticated bots involved in the attack.

37. **Mr Ruslan Deynychenko** described how there are networks of bots on social networks allowed in the Ukraine, calling for massive anti-government riots. These bots speak from an ultra-patriotic Ukrainian perspective, criticising authorities for not implementing reforms, for ineffective anti-corruption measures, and lost opportunities. Often, the problems were exaggerated and the achievements were ignored. It was later discovered that citizens of a foreign State were behind these networks of bots.

38. **Dr Kevin Limonier** noted that bots were used during the Russian demonstrations in 2011-2012. Large amounts of messages were published on Twitter and a Russian social media network to discredit, harass, and disorganise the protest movement. He referenced a study that showed that of the 2,000 bots established during the demonstrations, only 5 to 6 accounts are still active today. These surviving accounts mostly published advertising, and do not cover political topics, which shows that someone bought the services of a company to manage the bots during the demonstrations.

39. The written representation from the **National Council of Churches** also recognised the prevalent use of bots in social media today. The representation discussed how bots can mass-send content, re-tweet selected items, and even follow each other – creating the false impression of the popularity of a particular profile. It concluded by stating that bots are responsible for spreading much of the fake news found on social media.

40. *Difficulties of detecting bots.* Representors also shared how it is increasingly difficult to tell if an account is run by a human or bot. **Mr Ben Nimmo** noted that some accounts appear to be cyborgs, which automatically repost the selected accounts, but occasionally make their own posts to appear human. Similarly, **Dr Kevin Limonier** said it was reasonable to assume the existence of sophisticated

bots who are effectively impossible to identify strictly because they do not have codes in common with other bots and often imitate accounts managed by human operators. **Mr Nicholas Fang** also noted that bot creators can blend automation with human curation, where humans post new comments, along with automated retweets, to create the impression that the accounts are used by real people. **Ms Myla Pilao** explained that an ordinary person usually cannot tell if a post online was created by a human or a bot. **Dr Kalina Bontcheva**, a professor of text analysis at the University of Sheffield, said that the key challenge in detecting spam bots is the fact that political bot accounts and fake news sites have a short lifespan, with new bots emerging quickly.

41. *Detailed accounts of methods of false amplification.* **Mr Ben Nimmo** explained that disinformation campaigns often try to generate a very high number of tweets from a very low number of users, in order to create the impression of spontaneous human activity. He used the analogy of shepherds, sheep dogs, and sheep to illustrate how up to 50,000 tweets can be generated in an hour from probably a group of no more than six people.

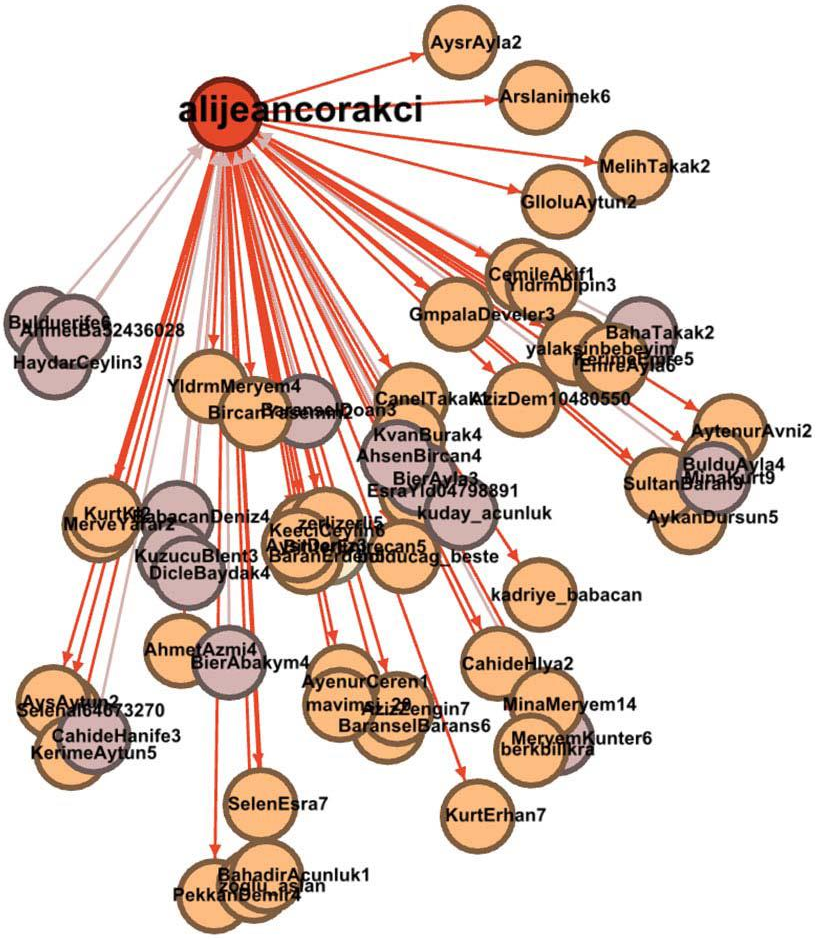
    a. The ones launching the campaign are the shepherds. These are a small number of accounts run by humans, with a low number of retweets but high levels of individual content. These accounts launch a hashtag or meme simultaneously to get a particular message out.

    b. Next in the line, to amplify what the shepherds have put out, are the sheep dogs. These are a larger number of accounts, with high levels of activity and retweets. This means they may be either very active humans or partially automated accounts. These accounts will retweet what the shepherds have posted, and also create their own tweets using the hashtags and memes posted by the shepherds.

    c. Finally, to amplify the information on a larger scale are the sheep, a medium-sized network of bots which will retweet everything the shepherds and sheep dogs have done.

42. **Ms Myla Pilao** also offered a similar explanation for how false amplification occurs. She said that Twitter accounts can be categorised as "gurus" or "sect followers". "Gurus" are often followed by a large number of "sect followers" who actively repost and retweet the original posters' messages. She described several aspects of these "sect followers" which appear to be suspicious. First, they have almost identical tweets or posts. Second, they post almost 24 hours a day. Third, they post posts about the same topic, and hardly have posts of other topics. Fourth, while each of these accounts tend to have different profiles, their registration date and their activities are almost the same, to the extent that one can conclude that they belong to the same network. She said that in a "fake news" campaign, bot accounts retweet a single post at around the same time using the same hashtag. Each bot has the same group of followers, and groups of bots will usually follow

and retweet the posts of the same "gurus" as well. She illustrated this point with the following diagram:[6]



c. Targeted advertising

43. Various representors shared about how targeted advertising can be an influential and effective amplification tool.

44. **Associate Professor Eugene Tan** noted that the fundamental business model of social media platforms like Facebook, Twitter, YouTube, Instagram, and others (like Google) is to enable prospective advertisers to use the treasure trove of data they have and their laser-like ability to sell advertisements (including political messages) to the platform users. He noted that with targeted political advertising on social media, it is possible for political parties and election candidates to communicate directly to each voter on social media platforms and on specific issues they are concerned about and even to tell them what they want to hear.

[6] Trend Micro Inc, Appendix III: Written Representations, Paper No. 86, page B872.

45. Similarly, **Mr Dan Shefet** noted that the business model of technology companies was to divide society into many sub-sections, which allows advertisers to target people in a very efficient manner.

46. **Mr Nicholas Fang**, the founder and managing director of Black Dot, cautioned that potential future threats will likely include measures like narrow-casting or micro-targeting individuals through social media and other online channels. Artificial intelligence programmes will use online behaviour to ascertain personal tendencies and characteristics, which can then be used to determine how to target different groups of people with tailored messaging. Mr Fang noted that this has proven effective especially in times of elections or when the public vote is being contested.

47. The lack of transparency in targeted advertising was highlighted by **Mr Charles Richard Kriel**, Specialist Advisor to the UK Digital, Culture, Media and Sport Select Committee, who wrote about how users are targeted with posts that contain highly customised misinformation designed for each user's personality type and previously expressed preferences, also known as "dark posts". According to Mr Kriel, these "dark posts" are notoriously difficult to discover or regulate, as by design, they can only be seen by the targeted users. These "dark posts" are often used to spread falsehoods about critical issues.

d. Social media algorithms

48. **Dr Claire Wardle** submitted that social media algorithms are designed to encourage people to seek out, consume and engage with information that supports their worldview. These algorithms help promote popular posts to intended and specific users, thereby amplifying certain messages to target groups. In this regard, **Dr Elmie Nekmat** referenced a nine-nation study, conducted between 2015 and 2017, which showed that the wide spread of disinformation on social media aimed at influencing public opinion is largely supported by Facebook and Twitter's algorithms.

49. Various representors also raised concerns of the use of these algorithms by social media companies. **Professor Hany Farid**, a professor and chairman of computer science in Dartmouth College, observed that these algorithms are programmed to optimise engagement by the users. He suggested that as a result, "clickbait" content is being optimised by the algorithms, instead of material that may be more trustworthy. He noted that if social media companies' algorithms are just optimising engagement, then they are vulnerable to manipulation. **Dr Lim Sun Sun** noted that the structures and algorithms by which technology companies sort and share information, and forge connections between media consumers, are still largely opaque and proprietary. She said this makes it difficult for media consumers to be conscious of hidden biases in the news and information they come into contact with. A similar point was made by **Mr Warren Fernandez**.

50. **Dr Kevin Limonier** noted that using social networks to spread disinformation is at the core of a foreign State's informational strategy, as these networks make it possible for the discourse and arguments of the foreign State's media outlets to be propagated efficiently to the masses. He said that to enlarge their audiences, the foreign State's media outlets would use the algorithms of social networks to "dope" the visibility of their content. For example, they would entice readers onto their sites using "clickbait" and funny pictures, and then rely on the social media networks to spread the information widely. Dr Limonier also said that on social media networks like Facebooks, users may be in an "algorithm jail", where they only see one point of view, i.e. the news and content the algorithm judges they might like.

51. Furthermore, **Dr Shashi Jayakumar** noted that algorithms can now harvest enormous amounts of information about us from social media platforms. He said that these algorithms can record, analyse and anticipate our preferences and sometimes needs even better than we do. He referenced an article that showed that with ten Facebook "likes" as inputs, an algorithm can predict a subject's other preferences better than the average work colleague, with 70 "likes", better than a friend, with 150 likes, better than a family member, and with 300 likes, better than a spouse. Similarly, **Associate Professor Eugene Tan** noted that algorithms have powered not just the speed but also the precision and relative impact of communication, in particular, boosting the intended reach and effect of deliberately targeted falsehoods.

52. **Mr Zhulkarnain** cited a quote describing the work of social media algorithms in the spread of falsehoods. In relation to the false, conspiracy video suggesting that a survivor of the Parkland, Florida shooting was but a "crisis actor", a commentator had this to say: "*It takes a special sort of heartlessness to create a conspiracy video about a teenage survivor of one of the deadliest school shootings in US history. But it takes a literally heartless algorithm to ensure that thousands, or even millions, of people see it.*"[7] This was after the conspiracy video was briefly pushed to the top of YouTube's Trending section, significantly increasing its visibility online.

e. Online falsehoods cascade over different platforms

53. **Mr Jakub Janda,** the Head of the Kremlin Watch Program and Director of the European Values think-tank in the Czech Republic, observed that false information can be spread through several means, such as social media, mainstream media, knowing third parties and unwitting persons.

54. **Dr Claire Wardle** referred to the example of the false article titled "Pope Francis Shocks World, Endorses Donald Trump for President, Releases Statement". The article was created by an unidentified person, and published on a self-proclaimed

[7] Zhulkarnain Abdul Rahim, Appendix III: Written Representations, Paper No. 80, page B479, para 18.

fantasy news site WTOE5News in July 2016. WTOE5News was part of a network of 43 fake news sites, which earned digital advertising income by generating readership. The article was shared on Facebook by someone working for this network of fake news sites. It was then re-shared by different groups of people, namely, (i) those who sought to amplify the reach of the article to make profit, (ii) Donald Trump's supporters, (iii) other forces who had an interest in Donald Trump winning, e.g. trolls linked to a foreign State, and (iv) Hillary Clinton's supporters, to show how easily Donald Trump's supporters could be fooled.

55. When this example was presented to the **representative from Twitter, Ms Mary Reen**, she expressed the view that such traveling of information across platforms was something that was not unusual.

56. Locally, **Mr Raja Mohan** observed that falsehoods from other sources can spread through the sharing of messages through closed platforms like WhatsApp by members of the same community. Mr Mohan gave the example of a fake message in Mandarin or another Chinese dialect that could be easily spread around within an elderly Chinese neighbourhood WhatsApp group.

f. Amplification is key in the systematic spread of falsehoods

57. **Ms Myla Pilao** explained that any successful campaign to spread falsehoods must have the following three elements – (1) motivation, (2) use of social media, and (3) access to the requisite tools and services, to amplify the falsehoods. Similarly, **Mr Ben Nimmo** noted that successful falsehoods have four components: (1) they have an instant emotional appeal, (2) they claim authority by referring to an unimpeachable source, (3) they have an insertion point into the information space, and (4) they have an amplification network which passes them on to a broader public.

58. Ms Pilao also described an eight-step "Public Opinion Cycle", which she said cyber propagandists are likely to follow when they want to change the public's opinion on a chosen topic. Some of the key steps include the following:

    a. Delivery. This involves spreading the falsehood across traditional and social media. Various technological tools and services to do so will be used at this stage to spread the falsehood quickly.

    b. Exploitation. This involves controlled target promotion among small but active groups of supporters on social media networks by running polls or putting up fake studies.

    c. Persistence. This involves convincing the target to actively promote the key falsehood until it goes viral. This is to reach critical mass in terms of supporter volume to ensure that the key falsehood has maximum visibility.

d. Sustainment. This involves measures taken to keep the falsehood going while anticipating and reacting to changes in sentiment. Tools and services to market deliberate online falsehoods can be used here as well, such as buying an advertorial on a major news site to give the falsehood credibility.

(2) Creation of low cost and high impact online falsehoods

a. Information Is Shared Without Verification of Content or Source

59. **Mr Ben Nimmo** explained that, given the scale of peer-to-peer interactions enabled by social media, disinformation agents can bypass traditional editorial verification and spread falsehoods unchecked today. Similarly, **Dr Ullrich Ecker**, an associate professor in the School of Psychological Science at the University of Western Australia, observed that although there is a lack of editorial gate-keeping or commitment to journalism ethics and standards in citizen journalism, blogs and social media posts are seen by many as trustworthy sources of information.

60. **Dr Thio Li-Ann**, a law professor from NUS, also observed that anyone with access to the Internet can now be a citizen journalist. She noted that such persons are not subject to the rigors of checking mechanisms and editorial oversight in ensuring the veracity of information. She said that where material is published anonymously or under a *nom de guerre*, recklessness or negligence may be the order of the day, instead of responsible and accountable journalism. Similarly, **Dr Gillian Koh** observed that the identities of those who publish or circulate information can remain anonymous or masked behind pseudonyms.

61. **Professor Gerald Steinberg**, a professor of political science and president of NGO Monitor Research Institute, lamented how false allegations made by powerful non-governmental organisations are published in the mainstream media as well as on social media without verifying the accuracy of the allegations. He said such allegations are almost instantly circulated on social media platforms, and amplified by tens of thousands of accounts (both real and fake) without any effort made to evaluate the accuracy of the allegations.

62. **Dr Claire Wardle** also highlighted that people can misunderstand the nature of a piece of literature. For example, people often do not realize that satire is actually satire, especially when they are reading on a social feed. During the 2017 French Presidential Election, CrossCheck, a fact-checking project, found that people were disseminating falsehoods masquerading as satire in order to avoid fact-checks.

b. Consumer-friendly Tools to Create Audio-Visual Content Online Are Readily Available

63. **Mr Ben Nimmo** noted that modern editorial techniques have made it much easier for malicious actors to create and spread false or misleading content, ranging from photoshopped images to doctored videos which can make a speaker appear to say something they did not. **Dr Liew Kai Khiun** said that because of the democratisation of the media, everyone has the skills to create their own platforms, to manipulate and doctor information, and to put on different perspectives in a rapidly increasingly compressed time.

64. **Professor Hany Farid** explained that even relatively unskilled users can now manipulate and distort visual media, given the wide availability of sophisticated image and video editing applications that permit editing in ways that are very difficult to detect, whether visually or with current image analysis and visual media forensic tools. **Dr Ullrich Ecker** also observed that the development of sophisticated image and video editing software will make it more and more difficult to differentiate real news from fake news.

65. **Mr Chan Yun Hsing Ronald** noted that fake news will get more convincing with improving software technologies. He said that it is currently possible to doctor photos to professionally show celebrities' headshots affixed to scandalous or compromising photo composures, which is known as "deepfake".

66. **Mr Carlos Nicholas Fernandes**, a technology entrepreneur, elaborated further on "deepfakes". He said that there is free and readily available technology such as "FakeApp", which can be used to create "deepfakes". The New York Times reported that creating a "deepfake" cost the writer less than US$100. Mr Fernandes also noted that researchers at the University of Washington had created a fake video of former President Barack Obama using very advanced technology, and said that it was a matter of time before commoditised fake video technology becomes as advanced.

67. In a similar vein, **Mr Teymoor Nabili**, a freelance journalist, said that the growing sophistication of artificial intelligence and machine learning technologies have enabled new techniques like "laser phishing" and "FakeApp" software, which can convincingly simulate actual people, whether friends or leaders, to deliver messages that are unrelated to the apparent sender. He also said that many such software, like Adobe's "Project Voco" are being developed and presented as amusing, easy-to-use consumer products, with scant recognition of the negative potential they inherently possess.

68. Representors provided specific examples of how real-world impact has been caused by photographs and videos which were digitally edited. **Ms Jennifer Yang** described how a doctored video of a speech by then-incumbent Jakarta governor Basuki Tjahaja Purnama (popularly known as "Ahok") was used to mobilise

opinion against him. The speech Ahok delivered was received well by listeners initially, with some noting that his comments were frank and straight-forward. However, a freelance academic, Buni Yani, edited the speech by removing parts of it, which changed the meaning of what Ahok said. Buni Yani then uploaded the edited speech to Facebook with the caption "is this blaspheming Islam?" This was used to mobilise opposition to Ahok and culminated in a protest movement on 2 December 2016. **Ms Myla Pilao** also pointed out that a doctored photo of burning teepees and a caption that sternly criticised the police for setting a protest group's camp on fire caused a misguided uproar on social media.

c. Online Platforms Can Be Created at Low Cost

69. **Mr Ben Nimmo** testified that carrying out disinformation operations is not expensive, and does not require a high level of technological expertise – all that is needed is a building with computers and internet connections, appropriate VPN masking, fake phone numbers to create accounts, and enough people to do the job. He also noted that the relatively low cost of creating an online platform has made it far easier for purveyors of falsehoods to look like traditional reporting outlets, without adhering to traditional editorial standards. For example, a website was created to mimic a genuine South African news site, and spread the false claim that South African President Jacob Zuma had resigned. This triggered a brief spike in the value of the South African rand.

70. **Ms Myla Pilao** concurred on this point, noting that to obtain the same reach, spreading fake news costs significantly less than posting legitimate advertisement or paid content. **Tisane Labs** also highlighted that it costs virtually nothing to publish a post that would rival the influence of traditional media. This removes barriers to participation by actors that, in the past, would not have been able to exercise influence on public opinion.

(3) <u>Market for Online Disinformation Tools and Services</u>

a. Tools

71. **Mr Ben Nimmo** explained that commercial groups are creating tens of thousands of bots that are available for hire, which can be used to spread falsehoods. He said people can buy 10,000 followers, or 1000 retweets, or 500 likes for Bitcoins or through online transactions. He gave the example of Devumi, a US company which sells a range of bot services. These bots are well-presented and look exactly like human users. According to Mr Nimmo, these Devumi bots were used to amplify the tweets of a South African political activist during the selection of the new African National Congress leader.

72. Similarly, **Ms Myla Pilao** explained that underground markets offer automation bots to amplify the popularity of a fake news story. She also said that cyber

propaganda campaigns will use do-it-yourself tools to automatically spam social media users. She said these tools require only a low level of programming, so it is cheap to buy, easy to set up, and the results are immediate.

b. Services

73. **Ms Myla Pilao** described some of the various tools and services available to disinformation agents. She said that such tools can be legitimate advertising and content marketing tools, or illegitimate tools and legitimate tools that are being abused.

a. Content Marketing Services. For as little as US$15-30, a fake news operator can obtain 500-1,000 word articles from content marketing service providers. Ms Pilao noted that these are great tools for people who do not have the time to create content, or convincing content.

b. Analytics Services. Ms Pilao also mentioned "public opinion monitoring systems", which can survey, research on and influence opinions in prominent forums and social media for US$1,850-4,175, depending on the number of key words identified.

c. Social Media Promotion Services. Such services rely on the popularity of the social media account used to trigger a word-of mouth effect on the account's followers. This can cost between US$0.16 and US$180,000 depending on how many followers the account has. Ms Pilao explained that providers of such services scan networks with more influencers in the target market, and inject posts onto this network, to flood the network with the information desired.

d. Content Takedown Services. Some fake news operators take down content that can have an effect opposite to what they desire. A content takedown provider called Yage Times reportedly earned US$7.9 million in 2011 alone for a single operation.

e. Vote Manipulation and Click Farm Services. Fake news operators who wish to influence the outcomes of polls or elections rely on these for US$4,925-14,524.[8] Ms Pilao shared how an individual tricked people into believing his shed is a top-rated restaurant on TripAdvisor using click farms. Click farms employ either bots or (in more underprivileged areas) actual workers to click like or dislike, or make similar reviews and comments to influence others. She commented that click farms are effective because there are not enough controls and standards in place to prevent this type of behaviour.

[8] Trend Micro Inc, Appendix III: Written Representations, Paper No. 86, page B863.

f. Crowdsourcing Services. Fake news operators can crowdsource for likes or dislikes, depending on their desired outcomes, for as little as US$1. Ms Pilao explained that, as with social media promotion services, platforms containing those who share the targeted interest are flooded with the disinformation, such as, for example, through advertorials and search engine results.

g. Content Distribution Services. In some countries, bogus and even legitimate news outfits can serve as platforms for fake news. Making fake content appear on legitimate news sites without appearing as advertorials costs a premium (more than US$20,000).

74. Ms Pilao also included two tables in her written representation which reflected the various services available for disinformation campaigns and the costs involved. The tables showed, for example, that buying one social media "like" would cost US$0.04, 1,000 WeChat likes would cost US$0.19, 500 retweets would cost US$2 or so, and 1 million Instagram likes would cost US$18. She shared that using such services, it would only cost approximately US$200,000 to cause a street protest in the US over a potentially inflammatory issue. The tables are enclosed as follows:[9]

[9] Trend Micro Inc, Appendix III: Written Representations, Paper No. 86, pages B866-867.

Tool/Service	Price (US$)
Vote that bypasses an IP address check, a Captcha, or simple authentication	0.02–0.03
1 social media like	0.04
Vote that bypasses social media authentication	0.04–0.05
Vote that bypasses detailed online registration	0.05–0.07
Vote that bypasses SMS confirmation/more complex authentication	0.09
40 WeChat article views	0.16
1K WeChat likes	0.19
like4u crowdsourcing service (5 minutes–24 hours)	0.20–80
100 Instagram subscribers, friends, likes, or video views	0.23–0.40
100 VK group subscribers	0.30
10K video views on Youku, LeTV, Sohu, Tencent, or iQiYi	0.32–3
100 Twitter followers, likes, or retweets	0.34
100 YouTube subscribers	0.66
1K video views	0.89–4
100 YouTube video dislikes	2
100 YouTube likes/video views/full video views	2/0.30/0.23
500 retweets	2–130
100 YouTube video comments/10 comments/1K YouTube views	3
1K Instagram subscribers	3–15

Tool/Service	Price (US$)
VTope (with 2M users) crowdsourcing coupon worth 10K points	8–21
1K group joins	11
VK user spamming (1K personal messages) service	14
500–800-word fake news article	15
500 WeChat followers	16
10K site visitors	17
1K high-quality fan page likes	17–31
1M Instagram likes	18
500 Instagram followers/1 month Facebook auto-like service subscription	25
1–1.5K-word fake news article	30
1 month Twitter promotion service subscription	30–150
1K votes on Weibo	32
8 comments per day for 1 month	45
1K high-quality channel subscribers/"Fake" content takedown service	50
5K votes on Weibo	55
VTope (with 2M users) crowdsourcing coupon worth 50K points	62
5K Weibo followers	66
Content distribution on a provincial news site	72
Content distribution on a national news site/ site on the Baidu news feed	116

Tool/Service	Price (US$)
5K WeChat followers	103–158
Social media VIP service	106–124
Content distribution on a real-estate/financial/business site	131
2.2K Facebook auto-likes/1K comments per month	150
Content distribution on an IT news/fashion/entertainment site	174
Content distribution on a healthcare site	189
Make a video trend for a certain search query service	222–266
Content distribution on dubious publication/a newspaper's classifieds section	266
10K WeChat followers	315
Content takedown service	394
Content takedown service on Tianya or mop.com	394–630
YouTube main page video appearance for 2 minutes	621
10K Facebook auto-likes per month	800
PR distribution to news outlets	802
1M YouTube views	999
10K votes/petition signatures	1,065

Tool/Service	Price (US$)
Public-opinion-monitoring service for 10 keywords	1,850
25K votes/petition signatures	2,664
Premium YouTube package (1M high-quality views and 50K likes)	3,150
Public-opinion-monitoring service for 20 keywords	4,175
Click farm service (1 server with remote control capacity for 30 phones	4,925
Content (with 4–6K characters) distribution on commercial news sites	5,328–9,768
Click farm service (1 server with remote control capacity for 50 phones)	7,815
YouTube main page appearance of 20 videos for 2–6 minutes	7,992
Click farm (1 server with remote control capacity for 100 phones)	14,524
Content (not marked as advertorial/paid) distribution	21,641
WeChat celebrity (with 10.7M followers) promotion	69,500
Weibo celebrity (with 78.3M followers) promotion	180,000

Note: The tools and services in the table above are arranged from cheapest to most expensive.

c. Hired Guns

75. **Dr Shashi Jayakumar**, Head of the Centre of Excellence for National Security at RSIS, noted that there are individual consultants and private sector entities specialising in hacking or interfering with elections with the aim of achieving a desired election result for their client. He said that the methods used include smears, hacking, spoofing webpages, and sending mass emails to influence outcomes. Dr Shashi gave the example of Andreas Sepulveda, a notorious "gun-for-hire" from Latin America who would rig elections for the highest bidder. According to Dr Shashi, in addition to using the methods described above, Sepulveda would organise real-world interventions as well, such as focus groups to understand ground sentiment.

76. Dr Shashi also noted that there appears to exist a growing shadow market for methods to influence target populations and outcomes in nations, using methods like those offered by Cambridge Analytica, which is reported to have profiled and micro-targeted the US electorate during the 2016 US Presidential Election. **Mr Dan Shefet** also noted how Cambridge Analytica sold information to the highest bidder.

77. **Ms Jennifer Yang** also said that there were "hired guns" in Indonesia, such as the Saracen Cyber Army. The Saracen Cyber Army is an online-based syndicate that created many social media accounts to spread hate speech for clients willing to pay for them. In fact, the Indonesian authorities have arrested organisers of the demonstration against Ahok on suspicion of paying Saracen to create and disseminate fake news. The Indonesian Centre for the Reporting and Analysis of Financial Transactions also reported that a number of undisclosed "high profile individuals" have been found to have transferred money to Saracen.

78. According to Ms Yang, the Saracen Cyber Army is not the only such player in Indonesia. As recently as February 2018, Indonesian authorities discovered a WhatsApp-based syndicate called the Muslim Cyber Army (which also operates through Facebook and Twitter), and indicated that there are many other such organisations that sought monetary gains in exchange for creating online fake news. In fact, the Indonesia Ministry of Communication and Information reported that as many as 800,000 websites have been found to have disseminated fake news, most of which was not reported to the Ministry.

ANNEX C: IMPACT OF ONLINE FALSEHOODS

(1) Short-Term and "Slow-Drip" Effect

1. **Dr Elmie Nekmat**, an Assistant Professor at the Department of Communications and New Media, National University of Singapore (NUS), said that the harmful effects of online misinformation and falsehoods can be either delayed or immediate. **Associate Professor Eugene Tan** similarly said that deleterious falsehoods could have an immediate and/or a "slow burn" effect. Disinformation expert **Mr Ben Nimmo** wrote about how online falsehoods can be used in short-term and long-term ways. **Dr Damien Cheong**, a Research Fellow at the National Security Studies Programme at the S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University (NTU), described how disinformation campaigns could take the form of a short or long game, where the short game created crises for the target, and the long game exacerbated existing crises and created more serious ones in the long run.

Immediate, one-off effects

2. **Mr Ben Nimmo** gave two examples of online falsehoods that had an immediate impact. The first was a false tweet in 2013 about a bomb attack on the White House, which was posted from a hacked Associated Press account. This falsehood triggered a short-term crash on the stock market. The second was a false claim in 2018 that South African President Jacob Zuma had resigned. This falsehood triggered a brief spike in the value of the rand.

3. **Mr Nimmo** also elaborated on the use of falsehoods in short-term ways. He described how short-term uses focused on a specific event, such as a vote, demonstration, natural disaster or security incident, to achieve an effect by the massive and sudden deployment of false stories or accounts. An example he gave of this was the use of leaked emails of the campaign of now-President Emmanuel Macron to suggest that Mr Macron had engaged in illegal activity, and the spread of conspiracy theories about the murder of Russian opposition leader in 2015 by thousands of bots, in order to drown out accurate information about the murder.

"*Slow burn" effects*

4. **Dr Elmie Nekmat** observed that the "drip-drip" effects of online falsehoods are likely to be cultivated over a period of time via constant exposure to a consistent set of information. He said such falsehoods can be made up of opinionated, biased information with strong extremist or partisan views, and that more often than not, the content of such information is not necessarily false but manipulated and twisted out of context.

5. **Dr Mathew Mathews**, a Senior Research Fellow at the Institute of Policy Studies (IPS), described the "slow-drip" effect as occurring when elements seek to exploit existing divisions and differences within a society for their own ends. He noted that in the Singaporean context, online falsehoods that can threaten social harmony can come in various forms and become an everyday experience. These can include reports that intentionally feature misinformation about particular ethnic, religious or immigrant groups and their loyalty to Singapore, their potential to commit anti-social acts or crimes, their lack of contribution to society, their overuse of state resources, or highlight and speculate about aspects of their culture which may not be well understood but deemed as at odds with majority culture.

6. Dr Mathews cautioned that such falsehoods, in combination and over an extended period of time, can have a corrosive effect, progressively chipping away at the harmony and cohesion that has been built up over time between different communities. This could move polarised communities further and further apart, and would leave Singapore more vulnerable to attempts to undermine its security and stability.

7. Dr Mathews gave two examples to illustrate his point. This first was regarding a school shooting in Florida on 14 February 2018. Dr Mathews said that the American community is also polarised on the gun control debate but that after the school shooting, bots were used to amplify certain points in the discussion to heighten emotions on the topic. The second was regarding *The Real Singapore* and the falsehood it spread about a Filipino family complaining about the playing of musical instruments during the Thaipusam procession. Dr Mathews noted that people quickly took to this story without questioning the veracity of the facts, and made comments maligning Filipinos. These examples illustrate the idea that there is an ongoing, low-level type deliberate online falsehood, which is then punctuated by high-level, high visibility events that would heighten tensions compared to if the ongoing erosion of trust had not taken place.

8. **Mr Ben Nimmo** explained that long-term uses of falsehoods typically focused on promoting or attacking a particular point of view. This could range from partisan and one-sided coverage, through hate speech, and into incitement to violence. An example was of British citizen Darren Osborne, who drove a van into a crowd outside a mosque in 2017. A UK judge found he had been exposed to racist and anti-Islam ideology over social media. Police investigations found he had been researching material from conspiracy theory and fake news websites in the weeks prior to the incident.[1]

9. **Dr Soon and Mr Goh** spoke about how, once an actor has started the process of – and achieved some success in – using disinformation to polarise a society, any future efforts to further polarise a society becomes increasingly easier, assuming that countermeasures are not done effectively.

[1] Kevin Rawlinson, "Finsbury Park-accused trawled far-right groups online, court told", *The Guardian* (23 January 2018).

10. **Mr Nicholas Fang** noted that if disinformation campaigns were waged at a lower intensity over a sustained period of time, the erosion of trust it engenders can lead to a society where every piece of news and information is doubted. He said that such an information crisis can have effects that are at least as damaging as a financial crisis, and create situations where leaders, institutions, and organisations are placed under suspicion as a matter of course, and where people are left without clear direction or confidence in their country and countrymen.

11. Several other representors made the point more generally. **Associate Professor Eugene Tan** from the Singapore Management University (SMU) School of Law noted that such "slow-burn" effects may have a greater impact than falsehoods with a one-off effect, because they are a lot more insidious and could operate in our societies without us even knowing it. **Mr Benjamin Ang**, a Senior Fellow and Coordinator of Cyber and Homeland Defence at the Centre of Excellence for National Security at RSIS, said information operations can work on slow burn issues that can be equally, if not more, pernicious. **Associate Professor Alton Chua** from the Wee Kim Wee School of Communication and Information, NTU agreed that it is reasonable to be concerned about the possibility of an insidious, low-level, "drip-feed" type of attack. **Dr Lim Sun Sun** agreed that the process of developing and spreading online falsehoods, because of the anonymity and deniability afforded by technology, allows an ongoing insidious process, laying the ground for further work.

(2) <u>Threats to National Security</u>

12. That online falsehoods could rise to be a national security threat was noted by several representors.

13. An RSIS research paper titled "Countering Fake News: A Survey of Recent Global Initiatives", attached by **Mr Benjamin Ang** to his written representation noted that fake news become a national security issue when it undermines the foundations (e.g. social cohesion, public institutions, peace and order) of the nation state. Similarly, **Ms Jennifer Yang** noted that foreign disinformation campaigns that undermine political figures or divide people on social, political, religious, or cultural lines are national security issues. She also agreed that domestic matters that touch on racial and religious issues can sometimes become a national security issue. **SCCA/SPC** and the **PAP Policy Forum** agreed that deliberate online falsehoods are a serious problem that poses a potential threat to social cohesion, peace and stability, and national security. Similarly, **Associate Professor Alton Chua** concurred that deliberate online falsehoods pose a potential risk to national security, racial or religious harmony, economic stability and cultural or mental dimensions of sovereignty.

14. Representors such as **the groups of students from SMU and NUS** also agreed, generally, that falsehoods may threaten a state's national security and sovereignty.

Dr Gillian Koh agreed that based on what is reported, disinformation campaigns can have real world consequences in terms of trying to sow discord, erode trust between groups and communities, exploiting fractures and fault lines and ultimately, undermining democratic institutions.

a. Undermining of social cohesion

15. Senior Research Fellow from the Institute of Policy Studies, **Dr Carol Soon**, and Research Assistant **Mr Shawn Goh** observed that deliberate online falsehoods often mirror the cracks and fissures that pervade each country. They noted that those who produce such falsehoods are astute in exploiting the pain points found in political systems and societies, and capitalising on people's anxieties, doubts, fears, and insecurities.

16. It was said that disinformation tactics tended to involve sowing division. **Dr Shashi Jayakumar** said that an aggressor could attempt to "peel off" one particular ethnic group or religion, using social media and disinformation to appeal to deeply ingrained historical and cultural issues and setting off one group against others or even against the government. **Associate Professor Alan Chong** from the Centre for Multilateralism Studies at RSIS said that information operations aim to weaken a potential adversary in peacetime by disseminating information that sows doubt and tension amongst a target population, and noted that there could be any number of ethnic, religious, or ideological features susceptible to such operations. **Dr Liew Kai Khiun** also noted that foreign influences seek to exploit and magnify existing social divisions. **Mr Nicholas Fang** said that disinformation campaigns typically feed on a society's area of vulnerabilities and fragilities, seek to amplify areas of doubt and unhappiness, and through the use of media and technology, perpetuate falsehoods voluminously and at great speed.

17. Similarly, political data scientists with the Technical University of Munich, **Dr Simon Hegelich** and **Mr Morteza Shahrezaye** said manipulative attempts using falsehoods can, in the long run, amplify existing tensions in society, resulting in polarisation. They said that from their own empirical work on the debates on Facebook and Twitter, it appears that there is already a measurable effect of polarisation, which is caused by the uneven distribution of information in these networks.

18. Likewise, **Associate Professor Eugene Tan** said that a deliberate falsehoods campaign works very well when there are existing social, political, and trust cleavages in society, which provide fertile terrain for foreign interference. He observed that the alleged foreign meddling in the US may not have made much headway had there not been deep internal rifts and political alienation among Americans.

19. **Dr Thio Li-ann** pointed out that falsehoods can damage the sense of solidarity and common identity and sharing of a range of common experiences by citizens.

Importantly, this sense of solidarity and common identity is exactly what is needed for a society's long-term health. Otherwise, society may devolve into 'tribes' championing single-issue agendas, without the ability to compromise and arrive at reasonable accommodations, or to uphold fundamental values crucial to the survivability of society.

20. **QSearch**, a social media analytics company, noted that deliberate online falsehoods exploit and exacerbate pre-existing social and racial tensions, whose causes are beyond the responsibility of the attacker. Lawyer **Mr Darius Lee** noted that the Internet can be used to accentuate individual biases, and exacerbate ideological fault-lines by polarising different segments of society deeper into their "echo chamber". The **Roman Catholic Archdiocese of Singapore** in its written representation referred to how the Pope, in his speech for the 2018 World Communications Day, warned that "fake news" damages the social fabric because it "exploits people's prejudices and weaknesses to generate fear and anger".[2]

21. *Political fault-lines.* **Mr Ben Nimmo** commented that disinformation campaigns tend to gradually inflame tensions and hollow out the political centre at the expense of the fringes. He said that disinformation campaigns inflame local tensions by focusing on divisive issues, including LGBT rights, gun control, race, and immigration. He gave an example of how troll-factory Facebook groups triggered a standoff between supporters and opponents of an Islamic centre in Texas.

22. Another example was given by **Dr Elmie Nekmat**, who highlighted how 9,097 posts related to energy policies and events posted between 2015 and 2017 were found to have manipulated Americans' opinions about pipelines, fossil fuels, fracking, and climate change via social media and stirred up tensions between conservatives and activist groups.

23. More generally, **Dr Carol Soon** agreed that deliberate online falsehoods try to polarise people and bring more and more people from the middle to the hard extremes. **Mr Nicholas Fang** agreed that deliberate online falsehoods can radicalise or push individuals to extreme points of view.

24. *Economic fault-lines.* **Ms Nataliia Popovych** and **Mr Oleksiy Makhuhin** from the Ukraine Crisis Media Centre noted that pensioners and people living in poverty in the Ukraine are vulnerable to foreign disinformation.

25. *Identity-based fault-lines.* **Dr Cherian George** observed that simple ideas can be used by political actors to activate tribal identities in a way that is very difficult to fight. He said that hate propaganda, which always involves disinformation, has been used (a) to facilitate crimes against humanity, such as genocides, ethnic cleansings, and brutal colonial conquests, and (b) as an instrument of identity

[2] Roman Catholic Archdiocese of Singapore, Appendix III: Written Representations, Paper No. 49, page B215.

politics, to mobilise supporters, intimidate opponents, and put pressure on authorities. He noted that even if they do not culminate in violence, such tactics worsen social division and discrimination, undermining national cohesion.

26. Dr George said that disinformation is used in hate propaganda on two levels. At the macro level, disinformation is used to emphasise the in-group's noble characteristics and portray the out-groups as inherently untrustworthy because of certain irredeemable cultural, religious, or ideological traits. This keeps the "us-versus-them" attitudes simmering on the backburner.

27. At the micro level, disinformation is used to create events to demonstrate how the out-group poses a clear and present danger to the in-group. Dr George noted that these events could be entirely fabricated, or involve half-truths about actual events. The stories may relate to attacks on the in-group by members of other communities; or government decisions said to disadvantage the in-group; or the appearance of cultural symbols (books, films, cultural practices, places of worship) deemed to be deeply offensive. According to Dr George, hate propagandists use these news stories to whip up indignation and outrage, thus instigating their followers to take desired actions.

28. Several examples from Singapore were cited. **Dr Liew Kai Khiun** from the Wee Kim Wee School of Communication and Information at NTU highlighted a recent incident involving comments posted by seemingly Myanmar-based user accounts on social media regarding the Rohingya issue. He noted that these comments were posted about articles on the Rohingya issue written by Singapore's mainstream media, and suggested that Singapore's mainstream media is a "Muslim media" and that Rohingyas do not exist in Myanmar. He said that the inflammatory nature of these comments, some of which have Islamophobic overtones, have created an online backlash from Singaporean Muslims, resulting in heightened tensions along religious and ethnic lines between users from the two countries.

29. **Dr Mathew Mathews** highlighted another example of a pre-mediated attempt to spread false rumours in Singapore. This was the false claim spread by The Real Singapore that a Filipino family had complained about some Singaporeans playing musical instruments during the annual Thaipusam procession in 2015, which led to a commotion between Hindu participants and the police. He shared how he personally witnessed how quickly netizens took to the story without questioning the facts. He warned that with the Internet, websites such as The Real Singapore could spread such articles at great speed and with grave consequences for public opinion and societal cohesion.

30. A **group from Nanyang Polytechnic** also cited the same incident as an example of how deliberate online falsehoods could cause rising racial tensions. They cautioned that although direct impact towards the citizens could not be seen by the single incident alone, the fact was that in the long run, feelings of hostility could be fuelled, causing more problems in the near future. They also noted that *The*

Real Singapore had also run several other articles to stir up ill-feelings towards certain racial groups in Singapore. In their view, it was evidence that such "small but impactful"[3] news was capable of causing hostility and anger towards a certain racial group, which was detrimental to Singapore's multi-racial society.

31. **Ms Chong Nyet Chin**, the Director of Food Safety and Quality at NTUC FairPrice, were of the view that falsehoods can impact social cohesion and religious cohesion in Singapore, citing as an example the fact that some people actually believed the hoax that NTUC FairPrice was selling halal pork.

32. **Ms Jennifer Yang** shared Indonesia's experience with online falsehoods. She said that in Indonesia, disinformation campaigns utilise sectarian and racist narratives that play on ethnic and religious sentiments, and that growing Islamism in Indonesian domestic politics has been accompanied by the rise of such campaigns. For example, she recounted how online hoax campaigns had polarised public opinion during the Jakarta gubernatorial elections in 2017. She described how a video that was edited to make a candidate, Basuki Tjahaja Purnamo (also known as Ahok) appear to have said something that he had not was used to accuse him of blasphemy, culminating in a protest movement and creating a divided "battleground."[4] Her views were corroborated by the representative from **MAFINDO**, an Indonesian hoax-busting organisation, who observed that falsehoods in Indonesia had targeted people of different ethnics, religion, political affiliations and other interests.

33. Ms Yang further described how during election campaign periods in 2012 and 2014, disinformation campaigns sought to put Jokowi's Javanese Muslim identity into question, casting him and members of his family as Chinese and Christians, labels that carry connotations of ethnic and religious minority statuses in Indonesia. In a country where the Chinese and Christian population have been prevented from holding the highest public office, disinformation campaigns could have an effect of dissuading some voters from voting for the targeted candidates.

34. Ms Yang also noted that disinformation campaigns can conflate long-standing domestic inter-ethnic issues with international affairs, creating tension both locally and abroad. For example, Indonesia's Islamists converge "the issue of China's economic and political rise with the position of ethnic Chinese in Indonesia, producing a toxic mash that threatens to undermine social stability in the country."[5] In December 2016, for instance, the Chinese embassy in Indonesia expressed concern over online anti-Chinese sentiment following media reports accusing China of deploying biological weapon against Indonesia, after four Chinese nationals were arrested for planting bacteria-contaminated chili seeds.

[3] Zheng Liren et al., Appendix III: Written Representations, Paper No. 60, page B343.
[4] Jennifer Yang, Appendix IV: Minutes of Evidence, page C308, para 2624.
[5] Jennifer Yang, Appendix III: Written Representations, Paper No. 82, page B507, para 11.

35. An example from the UK given by **Dr Mathews** involved a photograph that had circulated on the Internet in the immediate aftermath of the London Westminster Bridge attack in 2017. The photograph depicted a woman wearing a hijab and talking on the phone at the site of the attack. The BBC reported that thousands shared the picture which claimed the woman, as a Muslim, was indifferent to the suffering of victims around her, and that #BanIslam was one hashtag circulating with the image. It was later revealed that the image was shared by a Twitter user which was a fake account created by a foreign country used to influence UK and US politics.

36. Other representors acknowledged the impact of falsehoods on identity-based fault lines. **Associate Professor Eugene Tan** noted disinformation campaigns are usually centred on racial, religious, and other social fault-lines, and that race, religion, and language issues could be fertile terrain for some with malicious intent towards Singapore. **Dr Carol Soon** observed that any racial community anywhere in the world is capable of being targeted with modern technology and algorithms as part of a disinformation campaign. The written representation from **Roses of Peace**, a ground up, youth-driven initiative, described "fake news" as a "destructive force in the digital age"[6], and that "fake news" exploits social fault lines and undermines racial and religious cohesion. Mr Mohamed Sa'at Bin Abdul Rahman, **editor of Berita Harian of the Singapore Press Holdings**, voiced concern about falsehoods spreading through the mother tongue. He said the mother tongue is emotive and can be exploited to influence opinion. He cited the example of the word "jihad", which has been misused with serious consequences.

b. Incite public unrest and violence

37. **Mr Septiaji Eko Nugroho** recounted how in 2016, a violent mob damaged seven Buddhist monasteries in North Sumatra after disinformation spread through chat apps about a Chinese woman complaining about the morning prayer call. He also recalled how in 2017, a false digital flyer containing the police logo spread, advising people to be careful about child kidnapping. Many people believed this disinformation and as a result, a father who was bringing rice to his children was beaten to death as he was suspected of being a child kidnapper.

38. **Mr Ben Nimmo**, a **group comprising a lawyer and SMU law students**, **Mr Cheah Wenjie** and **Mr Chester Su** (students from the NUS Faculty of Law), **Mr Carlos Nicholas Fernandes** (a technology entrepreneur), and **Mr Benjamin Goh** all cited the "Pizzagate" example, where foreign disinformation agents spread a false story that then-presidential candidate Hillary Clinton was complicit in a paedophile ring managed from a pizzeria in Washington DC. The false story resulted in threats made against the pizzeria owner and groups that had performed at the pizzeria, and prompted a man to show up at the pizzeria armed with a rifle, and fire three shots.

[6] Roses of Peace, Appendix III: Written Representations, Paper No. 158, page B1384.

c. Instigate public disorder and instability

39. A **group from Nanyang Polytechnic** raised the example of panic-buying of salt in China due to the spread of misinformation that it would ward off radiation poisoning from the threat of Japan's nuclear emergency. This led to the 10-fold rise of salt prices and many stores running out of stock. They expressed the view that it was not unbelievable for Singaporeans to fall for a similar hoax, especially in times of crisis or vulnerability, when society would be more susceptible to such impact.

40. The same group also cited the hoax about former Prime Minister and Minister Mentor of Singapore, Mr Lee Kuan Yew's death in 2015. They emphasised the speed at which the falsehood spread, and how within the day, several international media outlets had already reported on his supposed death. Although their erroneous reports were subsequently hastily corrected, some panic had already occurred at a societal level.

41. A relevant example given by **Mr Ben Nimmo** was of the false tweet that the White House had been bombed, which led to a massive fall in the stock market, which was fortunately quickly reversed. In a similar vein, **a group of SMU law students** noted that the spread of false news about the credit issue of a bank could result in it closing down, and suffering real capital loss due to a large proportion of the population withdrawing their assets.

d. Threaten territorial sovereignty

42. **Mr Ruslan Deynychenko** said that foreign disinformation campaigns against Ukraine resulted in one part of Ukraine being illegally annexed by a foreign country, and troops and unidentified military men with machine guns and heavy weapons entering another part of Ukraine. He said that this led to the death of 10,000 people and forced millions to become refugees. Mr Deynychenko also shared how foreign disinformation about atrocities committed by Ukraine against Russian speaking citizens motivated Russian citizens to fight against Ukrainian government forces in Donbas. He cautioned that we should never ignore the existence of propaganda, citing the example of Ukraine, as one day there might be people killing each other because they were persuaded to hate each other. He said that disinformation campaigns are "a powerful weapon [that can] be pointed to any country at any time very, very quickly."[7]

[7] Ruslan Deynychenko, Appendix IV: Minutes of Evidence, page C163, para 1402.

(3) Harm to Public Institutions and Decision-Making

43. Considerable concerns were expressed by local representors about the impact of online falsehoods on democracy. For example, **Mr Benjamin Goh** said that prolonged exposure to false information amplifies the negative effects of misinformation, which further erodes the quality of discourse, a central pillar of democracy. Lawyer **Mr Darius Lee** said that widespread falsehoods can illegitimately skew public opinion, thereby undermining the proper functioning of democracy. More generally, **Associate Professor Eugene Tan** said that the threat of deliberate falsehoods strikes at the core of representative democracy. **Dr Gillian Koh** made the point that deliberate online falsehoods can end up corroding democracy and affecting healthy public debate.

44. **Ms Er Shengtian Rachel** and **Mr Joel Jaryn Yap Shen**, students from the NUS Faculty of Law, explained two ways in which deliberate online falsehoods hinder democracy. First, they undermine representative government as voters are unable to make informed choices between competing candidates and policies. Second, they undermine deliberative political debate, which destroys the feedback loop between the government and the governed.

a. Damaging society's shared public space and impeding informed participation in public discourse

45. *Impeding informed participation.* **Dr Thio Li-Ann**, a professor at the NUS Faculty of Law, observed that the propagation of deliberate online falsehoods can undermine deliberative democracy. This is because the working of a democratic society depends on its members being informed, not misinformed. Dr Thio also shared how falsehoods seek to undermine the process of allowing citizens to engage with a range of representative views of issues of common concern. This can weaken society because a range of representative views is required for understanding accurately where another citizen is coming from, for facilitating compromise and overlapping consensus, and to cultivate a commitment to pluralism. In the context of falsehoods during election campaigns, Dr Thio also explained how online falsehoods could divert attention from the real issues. While one may have the ability to articulate one's side of the truth, this would take effort and time.

46. **Mr Ben Nimmo** shared how the intent of disinformation campaigns is to make as many people as possible as angry as possible because people are easier to manipulate when they are angry, and less likely to have a sensible debate. **Dr Mathew Mathews** observed that when disinformation is amplified, people's emotions on the issue may become stronger.

47. *Disengagement from the public space.* **Dr Ullrich Ecker** said that being exposed to misinformation can cause people to stop believing in facts altogether, and decrease their engagement in public discourse. He said if trust in facts is eroded

such that facts no longer matter or are even portrayed as "unknowable", then objective evidence becomes irrelevant and policy making is no longer constrained by reality.

48. An example of this impact was given by **Mr Jakub Janda**, who shared that 53% of Czechs believed there is propaganda both for and against a foreign country in the public space, and that they cannot trust anything.

49. **Dr Thio Li-ann** observed how a deluge of fake information may cause people to give up being an engaged participant in civic life, as fake information crowds out reliable news, rendering it near impossible to judge the veracity of content, tell the truth from falsehood, wheat from tares. In the absence of reliable informational sources, the wearied person may retreat to the less taxing world of entertainment and the vapid titter-tattle of gossip. According to Dr Thio, if this takes place on a large scale, it would be a loss for the democratic process and culture.

50. Similarly, **Mr Nicholas Fang** noted that a population may become disinterested in news and information as a result of frustration or helplessness in terms of knowing who or what to trust and believe. He cautioned that this could lead to an "information crisis" where every source or platform of information is called into doubt, and which could lead to societal paralysis, dysfunction, conflict and chaos. **Mr Ben Nimmo** said that the spread of the concept of "fake news" may contribute to a further sense of alienation from all media, and a growing mistrust of all received values, which would seriously undermine democratic debate.

51. The influence of online falsehoods on people's belief in objective data was also noted by other representors. Lawyer **Mr Zhulkarnain Abdul Rahim** stated that deliberate online falsehoods devalue and delegitimise voices of expertise, authoritative institutions, and the concept of objective data. **Associate Professor Eugene Tan** said that it would be very harmful if people decide to just switch off and not believe anything that they read.

52. *Media.* **Mr Warren Fernandez**, the Editor-in-Chief of Singapore Press Holdings' English/Malay/Tamil Media Group, said that there is a constant drip feed online, attacking the mainstream media by questioning their credibility and pointing to delays in information. He explained that such delays are a result of needing to verify dubious information, but with the information spreading online, aspersions are cast against the mainstream media for not reporting on the information quicker. This view was shared by **Mr Walter Fernandez**, Editor-in-Chief at Mediacorp.

53. Similarly, **Mr Ben Nimmo** said that the distrust of mainstream media has been actively fostered by "alternative" news outlets from various political extremes, who have a shared interest in weakening the political centre and the credibility of established outlets. **Mr Warren Fernandez** shared that according to the Edelman Trust Barometer, trust in the media has declined in 22 out of 28 countries

surveyed, including Singapore, and that is because of the conflation of mainstream media and social media.

b. Obstructing public institutions in policy-making and the delivery of public services

54. *Trust in public institutions.* Several representors expressed concern generally about the impact of online falsehoods on trust in public institutions. **Dr Damien Cheong** cautioned that public institutions in Singapore may be targeted by disinformation operations. He said that disinformation actors can target the police, which is a highly trusted institution in Singapore, by creating incidents to generate distrust against the police. He said that undermining trust in the police will undermine trust in the state, and noted that such incidents have been generated.

55. **Mr Nicholas Fang** and **Dr Lim Sun Sun** noted that deliberate online falsehoods reduce social trust between people, institutional trust with the police, courts, and other organisations, trust in democracy and our process of politics and governance, and trust in the process of receiving and consuming news and information. The **editors of Singapore Press Holdings** stated that a major consequence of any spread of misinformation was a concurrent rise in mistrust, which could undermine Singapore's institutions, policies or values. Similarly, **Dr Gillian Koh**, agreed that deliberate online falsehoods can erode trust in key institutions, including the police.

56. Psychological research has shown how misinformation impacts trust in public institutions. **Dr Ullrich Ecker**, an Associate Professor at the School of Psychological Science at the University of Western Australia, said that mere exposure to conspiratorial discourse, even if the conspiracy claims are dismissed, makes people less likely to accept official information. He cited one study as showing that exposure to conspiracy claims adversely affected trust in government services and institutions, including those unconnected to the conspiracy claims.

57. Evidence was led on the experience in other countries. **Ms Jennifer Yang** noted that rumours and conspiracy theories clearly reduced trust between Indonesian citizens of different political, cultural, and religious affiliations, as well as between the government and its constituents. **Mr Ben Nimmo** shared how a foreign troll factory tried to widen the divide between the Black Lives Matter movement and the police by running Instagram accounts in favour of both the Black Lives Matter movement, as well as the police and the right to shoot Black Lives Matter activists. This troll factory even put out a fake video that purportedly showed an African-American woman being shot by a policeman in Atlanta, Georgia.

58. *Impeding policy-making*. Expert representors shared other countries' experiences on the matter. **Mr Jakub Janda**, the Head of the Kremlin Watch Program and Director of the European Values think-tank in the Czech Republic, stated that disinformation operations in Europe have resulted in (a) European countries

finding it impossible to craft constructive policies on issues such as migration, and (b) deteriorating trust in the European Union. He said that due to foreign disinformation campaigns, one-quarter to one-third of the Czech population believes that Ukraine is governed by a fascist government. This means that it is almost impossible for the Czech government to support Ukraine with, for example, humanitarian aid. He also said a quarter of Czechs believe disinformation, which results in figures such as four in ten Czechs blaming the US for the crisis in Ukraine. Mr Janda cautioned that if disinformation is not countered properly, it can result in the public losing trust in democratic institutions, in free media, and in democratic political parties.

59. **Dr Elmie Nekmat** observed that disinformation campaigns have involved attempts to influence public debates on domestic policies. For example, between 2015 and 2017, 9,097 posts relating to energy policies and events were found to manipulate Americans' opinions on pipelines, fossil fuels, fracking, and climate change.

60. **Dr Simon Hegelich** and **Mr Morteza Shahrazaye** said that from the data they analysed, they got the impression that the turn in the public debate about the refugee situation in Germany may have been affected by manipulative attempts. They observed that while many people in the real world were trying to help the refugees, social media platforms were flooded with negative comments. They said that people from the political right were using all kinds of online manipulation techniques to create this negative trend.

61. In conducting research into the processes by which people form their opinions and beliefs, **Dr Ullrich Ecker** and his colleagues explained why this was of public interest. They stated that if a majority believes in something that is factually incorrect, the misinformation may form the basis for political and societal decisions that run counter to a society's best interest.

62. More generally, **Mr Nicholas Fang** observed that a misinformed public is not good for a country, as it will result in individuals, social groups, communities, and organisations making decisions based on incorrect or inaccurate data. T**he groups of students from SMU and NUS** also concurred, that falsehoods may impair a government's ability to formulate policy.

c. Undermine of right to a representative government and representative politics

63. **Mr Ben Nimmo** gave an example of the impact of falsehoods on voting processes. This was a false claim from a Russian observer that the counting of the Scottish referendum did not meet international standard. This false claim fed calls for a revote (not merely a recount) counted by impartial international parties. The petition to this effect gathered over 100,000 signatures.

64. **Mr Jakub Janda** spoke on the impact of an alleged foreign State-sponsored disinformation campaign that sought to undermine the reputation of the Ukraine government. He referred to a referendum in the Netherlands on whether the EU should enter into a trade agreement with the EU, where 59% of Dutch people who voted against the trade agreement purportedly did so because they believed the Ukrainian government to be corrupt; 19% of them believed the unproven claim that Ukraine had shot down MH17, an event that killed 193 Dutch citizens.[8] If true, this would also be an example of a falsehood that may have impaired policy-making.

65. **Dr Elmie Nekmat** noted that disinformation campaigns tend to be strategically aimed at influencing election outcomes, by steering public discourse and altering public opinion within short, immediate time periods. **Dr Carol Soon** and **Mr Shawn Goh** also said that deliberate online falsehoods as part of a disinformation campaign have wreaked havoc on domestic politics and allegedly influenced referendum and election outcomes in other countries. They also spoke of how deliberate online falsehoods that disrupt democratic processes are a severe threat. **Ms Jennifer Yang** was of the view that disinformation surrounding Indonesian domestic politics could have had an effect on voters.

66. While **Dr Hegelich and Mr Shahrezaye** had expressed scepticism about whether falsehoods can actually influence people's voting behaviour, they acknowledged that, at the very least, anyone who is monitoring what is going on on social media might get a wrong impression and make bad decisions.

d. Waste of Public Resources

67. Several representors noted that dealing with online falsehoods wasted resources. The representative from **MAFINDO** described how there would be a big wave of falsehoods each time there were elections in Indonesia, and how some of the falsehoods, though "quite silly", could cause the government to have to commit time and resources to clarify the disinformation. **Mr Raja Mohan** characterised the matter as an "opportunity cost" that arises because of the wastage of time spent addressing the issues of online falsehoods, when the time could have been spent fixing actual problems.

(4) <u>Harm to Individuals</u>

a. Interference in individual decision-making

68. **Dr Ullrich Ecker** said it was well-established that misinformation often continues to influence people's memory, reasoning, and decision-making, even after people have received clear and credible corrections. This can be because people misremember a corrected "myth" as true, draw inappropriate inferences from the

[8] "Putin's Asymmetric Assault on Democracy in Russia and Europe: Implications for US National Security", *US Congress* (10 January 2018), pp 113-115.

information they have received, or make inadequate decisions based on misguided beliefs. He said that misinformation can make people feel more concerned or threatened than the evidence warrants. Dr Ecker and his colleagues have also noted that if individuals are misinformed, they may make decisions for themselves and their families that are not in their best interest and can have serious consequences. An example of this was how parents decided not to immunize their children, following unsubstantiated claims of a vaccination-autism link, which, they said "has had dire consequences for both individuals and societies."[9]

69. In a similar vein, **Dr Simon Hegelich** and **Mr Morteza Shahrazaye** noted that manipulative attempts with social bots, trolls, and hyperactive users can create the impression that a specific opinion is very popular or unpopular online. They said that anyone monitoring social media might end up getting the wrong impression and make bad decisions.

70. Another impact of falsehoods on individuals was on the authenticity of their political participation. **Dr Thio Li-Ann** said that falsehoods can thwart the liberty of individuals to effectively participate in the political process in an informed manner if people vote based on the misinformation.

71. Representors shared how it was getting harder to distinguish real news from fake news. **Dr Ullrich Ecker** noted that it has become increasingly difficult even for experienced and well-informed news consumers to reliably distinguish valid information from misinformation. Similarly, **Mr Nicholas Fang** commented that being able to identify and recognise fake news is not a skill-set possessed by the majority of people. In fact, given the level of sophistication involved, even trained professionals familiar with the news industry have a difficult time discerning what is real and what is not. He said that various state and non-state actors have refined and improved their capabilities in producing tools and products that are virtually indistinguishable from the real thing, and spread through such insidious channels as to fool even the semi-trained eye. The **editors of SPH** also acknowledged that people have difficulty distinguishing between fake and real news.

72. **Dr Carol Soon** shared that over 60% of respondents in the Edelman Trust Barometer Global Report 2018 agreed that the average person does not know how to tell good journalism from rumours or falsehoods. **Ms Myla Pilao**, Director for Technology Marketing at Trend Micro, shared that in a recent US survey,[10] as many as 20% of respondents did not feel confident about discerning what was fake news and what was real news. She also shared that even if readers could tell fake from real news, 88% of the respondents said they felt confused by disinformation.

[9] Lewandowsky et al., "Misinformation and its Correction: Continued Influence and Successful Debiasing", *Psychological Science in the Public Interest* 13(3) (2012) 106, p 107.
[10] Trend Micro Inc, Appendix III: Written Representations, Paper No. 86, pages B853-854.

b. Provocation of harassment and insults

73. Two representors shared stories of how they were victimised by deliberate online falsehoods. **Ms Gan Siok Bin** shared how she received lewd messages from strangers from foreign countries as a result of people spreading falsehoods about her in an online forum. **Mr Prakash Kumar Hetamsaria** shared how his photo was used in an article to suggest that he was a new citizen who was disappointed with Singapore and thinking of giving up his Singapore citizenship. This despite the fact that he has been a citizen since 1999 and was an active grassroots leader. The false association resulted in xenophobic and racist comments being directed at him, which caused distress to him, his wife and his young daughter.

74. **Dr Shashi Jayakumar** also gave the example of nineteen-year old Mr Anas Modamani, a Syrian refugee whose selfie with German Chancellor Angela Merkel was used by far-right groups to falsely claim that he was an ISIS member who carried out terrorist attacks. Since then, his photograph appeared in other false stories on social media linking him to terrorist attacks across Europe. In a news interview, Mr Modamani spoke of being too afraid to leave his house because of the false stories.[11]

75. More generally, other representors, including **Mr Jev Akshay**, **Mr Yeo Chee Hian**, **Mr Ngoh Wang Long**, **Mr Cheah Wenjie** and **Mr Chester Su**, and **Mr Benjamin Goh** pointed out that deliberate online falsehoods can adversely affect a person's reputation.

c. Harming of health

76. **Mr Septiaji Eko Nugroho**, founder of the Indonesian Anti-Hoax Community or MAFINDO, observed that there are a number of misleading health claims on social media. These claims sometimes mix the truth with falsehoods, and sometimes try to get people to buy products that could actually be dangerous. He said that according to some health institutions in Indonesia, patients stop their medical treatment so that they can follow the claims made online instead. This results in people dying from lack of proper medical attention.

77. Mr Nugroho also gave an example of the impact of false medical claims. He said a person suffered a stroke in the office and instead of taking him to the hospital, which was ten minutes away, his colleagues followed a false claim circulating on social media that suggested taking a needle and put it in the stroke patient's fingers and ears. This led to the person's death.

78. **A group from Nanyang Polytechnic** noted how vaccines, a vital public health tool, is under threat from growing public mistrust in immunisation and the rise of "fake news" drowning expert voices. They gave the example of parents in

[11] Stephanie Ott, "How a selfie with Merkel changed Syrian refugee's life", *Al Jazeera* (21 February 2017).

Indonesia refusing to let their children be vaccinated against infectious diseases after falsehoods were spread about the ingredients of the measles-rubella vaccine and that it was *haram*. They also shared how false information is spread in respect of beauty products. For example, weight loss pills touted as "100% natural" were found to contain sibutramine, which is an illegal substance that increases the risk of heart attacks and strokes, and causes other serious adverse effects.

d. Causing of financial and other harm

79. The representative from **SingTel** gave evidence of how customers of SingTel had been victims of scams in the past. These scammers made certain representations using the SingTel brand, or certain things that had been done by SingTel, in order to extract a commercial gain from SingTel's customers.

80. The representative from **NTUC FairPrice** suggested that some of the scams that had affected NTUC FairPrice in the past, such as the gift voucher scam, could have been part of a ploy to obtain customer's personal data.

81. NTU accountancy undergraduate **Mr Chua Jun Hao** cited the incident where a false claim posted on SGX's website about a capital acquisition by one listed company from another company. The claim had been quickly removed and debunked. Mr Chua highlighted that shareholders, investors and managers could have been misled by the misinformation to make wrong judgments. Investors could also have suffered financial losses had the stock price plunged.

(5) <u>Harm to Businesses</u>

82. The **Singapore Corporate Counsel Association** and the **Singapore Press Club** ("**SCCA/SPC**") said that corporations have been the target of online falsehoods, and that such falsehoods affect the corporation's reputation and image in the minds of consumers and can even impact upon public health and public safety concerns. **NTUC FairPrice** shared their own experience, noting that some of the consequences of online falsehoods against corporations include loss of business, reputational risk, and deterioration of customers' confidence, goodwill and trust.

a. Triggering of alarm over food product safety

83. The **SCCA/SPC** gave a few examples of falsehoods that affected corporations. One was a video circulating on WeChat, which claimed that Ayibo Food's seaweed was made of plastic. Chinese food safety officials had to intervene to counter the allegations in the video. Another was a report circulating through social media that Malaysia's Health Ministry issued a notice to Nestle Malaysia to withdraw all their instant noodles as the noodles contained lead. This falsehood was even aired on a local TV channel. Malaysia's Health Ministry had to issue a statement, saying that the noodles were safe and the report was untrue. Yet another was the report that Coca-Cola has recalled its Dasani water products after a clear

parasite was found in bottles across the US. Coca-Cola had to clarify that this was not true, and the US Food and Drug Administration indicated that they were not aware of any current recalls or disease outbreaks associated with Dasani water.

84. **Mr Jonas Kor**, the Director of Corporate Communications and Brands at NTUC Fairprice, described the "plastic rice" incident involving NTUC FairPrice. In 2017, a false story was spread that NTUC Fairprice's house-brand rice was made of plastic pellets and not rice. The person who spread the story claimed that his friend who was a pharmacist confirmed that the rice was made of plastic. This created a lot of fear and concern among NTUC Fairprice's customers and the public. NTUC Fairprice made a police report and worked with authorities like the AVA to assure the public that the story is false. The SCCA/SPC also noted that this online falsehood surfaced a public health and food safety concern on the sale of plastic rice in Singapore, even though the falsehood targeted NTUC FairPrice's house brand jasmine fragrant rice.

85. **NTUC FairPrice's** written representation also described the "Fake Chin Chow Incident", which involved videos circulating on social media which suggested that the "Tan Soon Mui Grass Jelly" product was made of plastic. Netizens queried this on NTUC FairPrice's social media page. Investigations later revealed that the allegation that the product was made of plastic was wholly false.

b. Straining of ties with customers

86. **Mr Hazrul Jamari** described a falsehood involving an elderly man who allegedly found pork cubes in a halal grocery store, which he had helped debunk. According to Mr Hazrul, he was concerned that the spreading of falsehoods would lead to trust between the local community and the store being affected.

87. Mr Kor from NTUC Fairprice described the incident involving a gift voucher scam, where a false story was spread that NTUC Fairprice would give people $400 to $1,000 in gift vouchers if they complete a survey as part of its anniversary celebrations. He said this would create a reputational risk for NTUC Fairprice because customers would think it is a legitimate survey and participate, expecting to be rewarded. He agreed that customers could turn up and become angry with the company for not giving them what they felt entitled to. He noted that the false story mixed the truth and falsehoods, as it used the fact that it was NTUC Fairprice's 45-year anniversary to perpetuate the false story. He also highlighted how it was being circulated with increasing frequency over the years, with three incidents occurring in the first three months of this year.

c. Smearing of business reputation

88. **Ms Myla Pilao** said that a company can spread false and negative comments about a competitor to rake in more business, nothing that this has happened in New

Zealand. She also said that business can be made to look bad using altercated audio and video files that render realistic-looking footage.

89. Representors shared examples of how business reputation can be affected by deliberate online falsehoods. **Mr Benjamin Goh** related how the "Pizzagate" story (a false story that claimed that then-presidential candidate Hillary Clinton was running a paedophilia ring from a pizzeria in Washington D.C.) spread explosively, fuelled by Twitter, Facebook, and Instagram, resulting in attacks against the pizzeria, and reputational losses to the owner of the pizzeria.

90. **Mr Jonas Kor** from NTUC FairPrice spoke about the "halal pork" incident. He said that the image of the Pasar Fresh Pork product with a halal sticker on it first surfaced in 2007. NTUC FairPrice made a police report and clarified via social media, their website and the mainstream media that this was a false image. They had to repeat these actions in 2011 and 2014, when the false image resurfaced.

91. Similarly, **Mr Sean Slattery**, the Vice-President of Regulatory and Interconnect at Singtel, shared how Singtel has in the past been the subject of commercial scams that affect its reputation. Such scams include promoting "get rich quick" online communications using the Singtel name and brand, or posting misleading information relating to a Singtel service to entice users to provide personal information or deposit cash or access to cash. **Mr Tim Goodchild**, the Head of Government and Strategic Affairs at **StarHub**, indicated that StarHub had similar concerns.

92. **SCCA/SPC** noted that legitimate advertising by corporations that appear next to online falsehoods or offensive content can destroy brands and their image.

d. Causing of financial losses

93. **Mr Jonas Kor** from NTUC FairPrice explained how companies may incur manpower costs and other losses when dealing with deliberate online falsehoods. Using the "plastic rice" incident as an example, he said that when false claims are made about food products, manpower is needed to investigate the claim, as well as update the public and assure them. He noted that sometimes, the corrective action taken is not sufficient as the falsehood has already caused unnecessary public alarm. **Ms Chong Nyet Chin** noted that if many such falsehoods are made against NTUC FairPrice, this will increase costs of manpower and resources, and these costs could eventually be passed down to the consumers, although that was an outcome the company aimed to avoid.

94. **Mr Zhulkarnian Abdul Rahim** identified the need to allocate ever-diminishing resources to debunking inaccurate information as one of the corollary harms of "fake news". He also related how in November 2016, there was a fake press release about a French building company, Vinci, which claimed that there had been a massive fraud and that the CFO had resigned, leaving a multi-billion euro

deficit. This caused Vinci's shares to dive immediately and by the time Vinci corrected the false press release, their share price had fallen, wiping billions of euros from their market value.

95. **Mr Benjamin Goh** said that the after the "Pizzagate" story broke and the shooting took place at the pizzeria, the owner had to spend $70,000 on security measures. He hired two guards to stand at the entrance during business hours, installed an alarm system and a network of cameras both inside and outside the pizzeria, and installed a panic button to alert the police in case of an emergency.

ANNEX D: DIFFICULTIES IN COMBATTING ONLINE FALSEHOODS

(1) Human cognitive tendencies

1. Psychological scientist **Dr Ullrich Ecker**, from the School of Psychology of the University of Western Australia, provided research-backed evidence on the psychology of why individuals believe misinformation. Dr Ecker is an expert on the psychology of misinformation processing and has conducted research into misinformation processing for around ten years. The findings he and his colleagues made included the following:

    a. *Misinformation may be more severe than ignorance.* There is a distinction between ignorance and belief in misinformation, and reliance on misinformation may be even more severe than ignorance. (Ignorance was defined as the absence of relevant knowledge.) Research had found that ignorance rarely led to strong support for a cause. However, false beliefs based on misinformation were often "held strongly and with (perhaps infectious) conviction."[1]

    b. *People's default is to accept information, including misinformation.* People usually cannot recognise that a piece of information is incorrect until they receive a correction or retraction. The deck is usually stacked in favour of accepting information rather than rejecting it, unless there are indicators that cast doubt on the motives of the source of the information. Breaking away from the default of acceptance requires more attention and mental resources. Hence, "[i]f the topic is not very important to you, or you have other things on your mind, misinformation will likely slip in."[2]

    c. *Compatibility with existing beliefs (also known as confirmation bias).* Numerous studies have shown that information is more likely to be accepted by people when it is consistent with their existing assumptions of what is true. The way we process information favours the acceptance of information that is compatible with one's pre-existing beliefs.[3]

    d. *Effect of repetition.* Misinformation has a stronger effect if it is repeated often. Repeated exposure to a statement is known to increase its acceptance as true. Repetition effects may create a perceived social consensus even when no consensus exists. [4]

[1] Stephan Lewandowsky et al, "Misinformation and its correction continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, p 108.
[2] Stephan Lewandowsky et al, "Misinformation and its correction continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, p 112.
[3] Stephan Lewandowsky et al, "Misinformation and its correction continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, p 112.
[4] Stephan Lewandowsky et al, "Misinformation and its correction continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, p 113.

2. Dr Ecker also shared his research on more subtle types of misinformation.[5] He and his colleagues noted that misinformation in the real world is often subtly misleading, For example, accurate numbers or trends can be communicated in a manner that makes them appear to have more, or less, significance than they in fact do, such as by cherry-picking data points. They stated that "[t]here can be little doubt that misleading headlines result in misconceptions in readers who do not read beyond the headlines." Their research further found that misleading headlines can lead to misconceptions and misinformed behavioural intentions in individuals.

3. The points made by Dr Ecker were corroborated by research findings from psychological studies reviewed by **Dr Carol Soon and Mr Shawn Goh**, from the Institute of Policy Studies in their paper "What Lies Beneath the Truth".[6] Their findings included the following points that showed the biases that tend to lead to people believing in falsehoods:

    a. Due to the deluge of information online, people rely on cognitive shortcuts to assess the information they encounter, and do not interpret information in a rational, neutral and objective manner.[7]

    Confirmation bias

    b. There exists the mental shortcut of confirmation bias, which is the tendency for people to accept information consistent with their pre-existing beliefs and reject information that contradicts them.[8] Research suggests that there may in fact be a neurological basis underlying confirmation bias: using confirmation bias to make decisions makes people feel good, in the same way as when they experience the positive effects of alcohol or opiate, eat chocolate, have sex or fall in love. As a result, people tend to focus on information that support their confirmation bias, and ignore information that contradicts their beliefs.[9]

    c. There have been several studies demonstrating the effect of confirmation bias. For example, one study found that individuals with higher prejudice towards homosexuals perceived fictitious scientific information that confirmed homosexual stereotypes as more convincing than individuals with lower prejudice. As for individuals with lesser prejudice towards

[5] Ullrich Ecker et al, "The Effects of Subtle Misinformation in News Headlines", *Journal of Experimental Psychology: Applied* (2014), Vol 20, No 4 323-335.
[6] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017).
[7] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 18.
[8] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 20.
[9] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 36.

homosexuals, they perceived fictitious scientific information that dis-confirmed homosexual stereotypes as more convincing than individuals with higher prejudice.[10]

d. Another study showed how confirmation bias also affects how people seek information.[11] Research subjects were made to listen to pre-recorded speeches on refuting arguments that linked smoking with lung cancer, and on the hypocrisy and wrongdoings of Christianity. The speeches were partially masked by static, and subjects were allowed to press a button that would reduce the static for a few seconds if they wanted to get a clearer listen. The study found that smokers pressed the button more than non-smokers when listening to the speech that debunked the relationship between smoking and cancer, and non-frequent churchgoers pressed the button more than frequent churchgoers when listening to the speech that attacked Christianity.

Familiarity bias and illusory truth effect

e. Research has established that repeated exposure to false information can influence people to believe that a falsehood is true.[12] This is also known as the "illusory truth effect".

f. Exposing people to false information will increase belief in the false information as people rely on familiarity as a heuristic in their cognitive processing. Repeated false information feels more familiar and truer even if it goes against what an individual already knows.

g. The illusory truth effect extends to not just to the perceived accuracy of plausible information, but highly implausible and partisan statements as well. In a study conducted on the effect of "fake news" on the 2016 US Presidential Election, researchers found that a single exposure to a fake news headline was sufficient to lead to an increased perception of accuracy. A second exposure led to an even greater perception of accuracy with the effect compounding over time. Furthermore, the increased perception of accuracy occurred despite the presence of explicit warning labels that indicated that the story was contested by fact checkers.

h. Research has also found that the illusory truth effect occurs even to those with knowledge about the topic that is the subject of the falsehood.

Social influence

[10] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 21.
[11] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 21.
[12] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 29.

i. There exist conformity cascades, where people tend to go along with the majority despite private doubts, in order to conform to the expectations of others and continue to be a member of the group. This effect is especially strong in tightly knit groups.

j. There also exist informational cascades, where people tend to believe in a rumour because others appear to believe it as well. This is because people tend to rely on the words of others as valid evidence of reality when they lack information of their own. People tend to think that something is probably true if they hear it from many others.

k. Ordinary people who are in the middle ground and not the extremes are susceptible to such influences.

4. **Dr Elmie Nekmat**, an Assistant Professor in communications and new media at NUS, also gave evidence on the psychological aspects of online falsehoods. Dr Nekmat has a PhD with an inter-disciplinary minor in educational and social psychology, and studies media psychology and effects, and digital media literacies. Dr Nekmat's evidence included the following points about the biases that tend to affect how people process information:

a. *Confirmation bias.* There exists confirmation bias, which involves assessing new information based on how compatible it is with pre-existing beliefs.

b. *Optimism bias.* There exists the tendency to believe that one is less susceptible to falsehoods than others. This tendency is found in people of all ages and different backgrounds.

5. At a more general level, the role that psychological biases play in the effectiveness of falsehoods was also recognised by several other representors. **Dr Lim Sun Sun**, Professor of Media and Communication at the Singapore University of Technology and Design, and **Mr Nicholas Fang** highlighted that online falsehoods leveraged on people's psychological pre-dispositions and emotions to take effect. **Mothership** agreed that heuristic tendencies like confirmation bias, illusory truth effect, and backfire effect add to the worry of the spread of deliberate online falsehoods. **Mr Dan Shefet** cautioned, in relation to deliberate falsehoods, that psychological research into the persuasion points of people carried the serious danger of being abused.

6. Confirmation bias was a phenomenon also recognised by representors such as **Mr Jakub Janda**, **Dr Damien Cheong**, **MAFINDO**, and in written representations from **Kwok Siang, a group from Nanyang Polytechnic**, and **Mr Gaurav Keerthi**.

7. In the same vein, lawyer **Dan Shefet** spoke about how human beings are prone to believing what is sensational and scandalous, and how people are more easily

manipulated into believing what is negative than what is positive. **A group of SMU students** highlighted as an area of concern how people rely a lot on their subjective emotions and views that appeal to their personal belief.

8. The familiarity and illusory truth effect was also referred to by other representors. **Dr Claire Wardle** referred to the "familiarity heuristic"[13] and how, if people hear the same information before, they are more likely to believe it. In his written representation, **Mr Benjamin Goh** wrote that people's brains may take "short-cut[s] to credibility"[14] by believing something to be true when people see multiple messages about the same topic. **Ms Myla Pilao** stated that "each time fake news is posted and reposted, readers of the same content grow familiar with it and take it as truth".[15] She also noted that the more the number of "likes" or reviews, the more one thinks that a piece of information is true.

9. Optimism bias was also referred to by **Mr Nicholas Fang**, who noted that it is a common response by the average person to believe that he or he will be able to recognise and resist any attempted fake news.

10. *Heuristic tendencies are greater online*. **Dr Carol Soon and Mr Shawn Goh** shared their research on the following ways in which the online environment tend to make people more susceptible to believing falsehoods:

    a. When online, people tend to not engage in the "deep processing" required for critical thinking, and tend to rely more on cognitive biases. "Shortcuts" that are used to assess credibility of information sources can cause individuals to be more susceptible to perceiving false information as accurate. Research has shown that instead of systematically processing the content of a website, users tend to rely on superficial aspects such as the overall visual appeal, layout, typography, font size and colour schemes to assess the site's credibility.

    b. People tend to be less sceptical about the information shared by their friends because they trust them. Research has found that the credibility of the most proximate source of the information, such as a Facebook friend, tends to exert the greatest influence on the assessment of the information's credibility.

    c. Algorithms are now being used to personalise information flows. Examples include websites like *Yahoo News* and a start-up funded by *The New York Times* to cater their headlines to audiences' interests and desires. In environments where people are provided with the news they prefer to read or hear, people are less sceptical of the information they receive.

[13] Claire Wardle, Appendix III: Written Representations, Paper No. 94, page B926.
[14] Benjamin Goh, Appendix III: Written Representations, Paper No. No 167, page B1434.
[15] Trend Micro Inc, Appendix III: Written Representations, Paper No. 86, page B587.

11. The speed at which online falsehoods gain "critical mass" in a short time can boost reliance on heuristic biases, according to **Dr Elmie Nekmat**. He explained that when falsehoods go viral, it is aggregation on social media, through "likes", shares and comments, that become compelling indicators of credibility of information that cannot be verified. Also, the tone of user messages and comments surrounding a story can influence how people think and feel about it, and can compel one to feel the same way, thereby "reinforcing inherent biases and attitudes when the tones are consistent with individual beliefs."[16] Dr Nekmat regarded these factors as exposing limitations of public education and media literacy efforts.

12. Similar views on the greater reliance on mental shortcuts online were expressed by several other representors. **Dr Thio Li-ann** noted that with the Internet today, the problem is no longer information deficits but a surfeit of information. The process of discerning good from bad arguments, truths from falsehoods is complicated where there is deliberate sowing of misinformation. She observed that the surfeit of information can overload the brain and hamper clear thinking, especially where falsehoods are mixed with the truth. **Dr Claire Wardle** made a similar point, noting that because of the overload of information today, mental shortcuts become more powerful. The **groups of students from SMU and NUS** also agreed that, given the deluge of information today, people tend to use mental shortcuts to process information.

(2) Weakness of truth compared with falsehoods

13. *Influence of falsehoods is difficult to reverse*. The influence of falsehoods is by its nature difficult to reverse, as shown by substantial psychological research.

14. This was shown by **Dr Ullrich Ecker,** whose research made the following points, among others:

    a. It is well-established that misinformation continues to influence people's memory, reasoning and decision making *even after* people have received clear and credible corrections.

    b. This is known as the "continued influence effect," which arises in part from failure of memory integration and memory retrieval.[17] As a result, the effect occurs even in cases where people do not have a vested interest or motivation to believe one thing over another.

[16] Elmie Nekmat, Appendix III: Written Representations, Paper No. 149, page B1304, para 9.

[17] Ullrich Ecker and Li Chang Ang, "Political Attitudes and the Processing of Misinformation Corrections", *University of Western Australia* (2017); Ullrich Ecker et al, "Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction", *Psychonomic Bulletin & Review* (2011) 18(3), 570; Briony Swire et al, "The Role of Familiarity in Correcting Inaccurate Information", *Journal of Experimental Psychology* (2017).

c. There is a "wealth of studies" showing that it is "extremely difficult to return the beliefs of people who have been exposed to misinformation to a baseline similar to those of people who were never exposed to it."[18]

d. The effect of corrections can wear off relatively quickly over time. Subsequently, people can return to accepting false claims as true simply because the false "myths" are familiar.

e. Repeating corrections does not entirely offset the influence of misinformation that is repeated often, especially by different sources.

15. The same effect was discussed by **Dr Carol Soon and Mr Shawn Goh,** who referred to it as "belief perseverance", a phenomenon where individuals retain newly created beliefs even after being informed that the initial information on which the beliefs were based was incorrect. According to them, belief perseverance suggests that impressions, once formed, are difficult to change. Thus, once a piece of false information is out in the open, it may be too late to blunt its influence. As summarised by Dr Soon and Mr Goh, research suggests that exposure to false information may have long term effects, while corrections may unfortunately be short-lived.

16. *Role of emotions.* Falsehoods that trigger negative emotions are generally harder to correct. **Dr Carol Soon and Mr Shawn Goh** cited research that has found evidence of a negativity bias, where information that evoke negative emotions is processed more thoroughly, leaves a stronger impression, and more resistant to correction than falsehoods evoking positive emotions.[19]

17. *Motivated reasoning - role of ideological world views and identities*. Motivated reasoning is the tendency to find justifications for existing wrong conclusions, despite conflicting facts. This phenomenon helps explain why people continue to be influenced by falsehoods despite the issuance of corrections.[20]

18. **Dr Carol Soon and Mr Shawn Goh** explained the psychological basis for why individuals engage in motivated reasoning. They made the following points from their research:

a. One of the reasons why individuals engage in motivated reasoning is to preserve their self-identity and group identity. People are motivated to defend their beliefs in the face of counter-evidence because if they do not, they risk losing their identity and membership in the group that they are in. Therefore, the sense of belonging people may have to the group is very powerful: it

[18] Stephan Lewandowsky et al, "Misinformation and Its Correction: Continued Influence and Successful Debiasing", *Psychological Science in the Public Interest* 13(3) (2012) 106, p 114.
[19] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies*, p 34.
[20] Stephan Lewandowsky et al, "Misinformation and Its Correction: Continued Influence and Successful Debiasing", *Psychological Science in the Public Interest* 13(3) (2012) 106, p 118.

allows one to more easily accept false information and dismiss the truth. For this reason, Dr Soon and Mr Goh agreed that an aggressor could attempt to "peel off" one particular ethnic group or religion, by using disinformation to appeal to deeply ingrained historical, cultural issues, in order to set off one group against another.

b. Research has established that an individual's political beliefs and identity contribute to motivated reasoning and can increase one's susceptibility to believing false information. Motivated reasoning and one's political identity also play a role in an individual's rejection of the validity of a scientific source. According to Dr Soon and Mr Goh, the research shows that politically motivated reasoning can cause one to reject even a properly researched, independent, objective and scientific point.

19. The role of an individual's worldview in making one resistant to corrections was also addressed in **Dr Ullrich Ecker's** research, which included the following points:

a. Where a piece of information consistent with existing knowledge or beliefs is accepted, it is highly resistant to change.[21]

b. A key reason why falsehoods persist in influencing people despite corrections is their worldview, or personal ideology. Corrections that contradict one's worldview are more difficult to process, less familiar, and less supported in one's social network.[22]

c. Dr Ecker and his colleagues had concluded that "personal beliefs can facilitate the acquisition of attitude-consonant misinformation, increase reliance on misinformation, and inoculate against the correction of false beliefs."[23]

20. *Corrections can back-fire*. The impact of cognitive biases on how people process corrections was underscored by research showing that corrections can back-fire, by increasing people's belief in the falsehood.

21. **Dr Ullrich Ecker** referred to several studies where backfire effects were observed in attempts to correct misinformation, such that people became even more committed to the misinformation.[24] For example, corrections about misinformation that President Bush's tax cuts in the early 2000s had increased

[21] Stephan Lewandowsky et al, "Misinformation and Its Correction: Continued Influence and Successful Debiasing", *Psychological Science in the Public Interest* 13(3) (2012) 106, p 112.
[22] Stephan Lewandowsky et al, "Misinformation and Its Correction: Continued Influence and Successful Debiasing", *Psychological Science in the Public Interest* 13(3) (2012) 106, p 118.
[23] Stephan Lewandowsky et al, "Misinformation and Its Correction: Continued Influence and Successful Debiasing", *Psychological Science in the Public Interest* 13(3) (2012) 106, p 120.
[24] Stephan Lewandowsky et al, "Misinformation and Its Correction: Continued Influence and Successful Debiasing", *Psychological Science in the Public Interest* 13(3) (2012) 106, pp 119-120.

revenues, or that there had been weapons of mass destruction in Iraq, led to a backfire effect among Republican participants. A similar effect was reported in a study relating to climate change. Messages highlighting the adverse effects on health caused by climate change led to a decline in support among Republicans for climate mitigation policies.

22. His evidence was corroborated by that of **Dr Carol Soon and Mr Shawn Goh**, who explained the following about the back-fire effect:

    a. The "worldview backfire effect" is particularly strong when it comes to corrections inconsistent with an individual's sense of cultural identity and their fundamental beliefs about how society should operate.

    b. This backfire effect was support by a study that used neuroimaging to investigate the neural systems involved in maintaining political beliefs in the face of counter-evidence. The study found that when the subjects were challenged on their strongly held political beliefs, there was more activation in areas of the brain that correspond with self-identity and negative emotions. This study suggested that humans may in fact be neurologically "hardwired" to hold on to pre-existing beliefs in the face of counter-evidence.[25]

    c. There was also the "familiarity backfire effect", which is based on the idea that familiarity towards a piece of information increases its chances of being accepted as true. As a result, the act of debunking false information may reinforce the information in people's minds. Dr Soon and Mr Goh cited a study which found that identifying medical claims as false helped people remember it as false in the short-term, but paradoxically increased its chances of being remembered as true after a three-day delay.[26]

23. *Biases are facilitated by conditions online.* **Dr Claire Wardle**, an expert engaged by the Council of Europe to provide a study on "fake news", made the point that social media algorithms are designed to encourage the predisposition of individuals to seek out, consume and engage with information that supports their world view. This is further discussed below, in relation to online echo chambers.

(3) Falsehoods travel faster and wider than the truth

24. *Corrections lag behind the falsehood.* Representors acknowledged the difficulty that corrections faced in keeping up with falsehoods. **Dr Soon and Mr Goh** spoke about how corrections may not reach a wide enough audience. During the hearing, Dr Soon drew attention to a recent study by MIT[27] that looked at a large number

[25] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), pp 36-37.
[26] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 31.
[27] Soroush Vosoughi et al, "The spread of true and false news online", *Science* 359, 1146-1151 (2018).

of tweets on Twitter over a ten-year period. The study found that falsehoods were 70% more likely to be re-tweeted than the truth.

25. **Mothership** agreed that it was sometimes very hard to correct misimpressions, once a falsehood is out there. They elaborated that an article debunking a falsehood will be read by more people if it is published with speed, but will not reach as many readers if the falsehood is circulating for a day or longer. They agreed that falsehoods travel faster than the truth.

26. Underscoring how powerful online falsehoods can be, **Professor Hany Farid** was of the view that technology to prevent the upload of false images (as opposed to just deleting the images) was "incredibly important"[28], because on the Internet, two hours is an eternity and things go viral in a matter of minutes or hours.

27. Several examples of how falsehoods spread further and faster than corrections were given by representors.

28. **Mr Ben Nimmo** gave an example from the 2017 Catalan independence referendum, where a photo of police pushing back against demonstrators under a Catalan flag was uploaded by a Twitter user. Within an hour and a half, a Spanish fact-checking organisation tweeted the truth: that the image was a fake, with the flag included using Photoshop. The tweet containing the truth was retweeted over 3,700 times, while the fake was retweeted over 12,600 times.

29. Another example from Mr Nimmo was that of a forged letter, purporting to expose connections between Britain's GCHQ intelligence agency and the Obama administration. At one stage, this letter was reposted by a Twitter user named "Babushka", whose post was retweeted 500 times. "Babushka" subsequently posted another tweet, indicating that the letter might be fake, but this correction was only tweeted a dozen times.

30. **Mr Prakash Hetamsaria** gave evidence on how his photograph was posted on the *All Singapore Stuff* website and how he was falsely identified as a new citizen disappointed with Singapore and considering giving up his citizenship. The article was shared over 44,000 times. Mr Hetamsaria's Facebook clarification, on the other hand, was shared only a handful of times.

31. Making the point more generally was a **group of SMU students and a lawyer,** who observed that the ease of dissemination afforded through the Internet and social media platforms leads to a "crowding-out" of truth and fact. They cited the quote that "*a lie can travel halfway around the world before the truth can get its shoes on*".[29]

[28] Hany Farid, Appendix IV: Minutes of Evidence, page C625, para 5313.
[29] Sui Yi Siong et al., Appendix III: Written Representations, Paper No. 130, page B1137, para 21; Appendix IV: Minutes of Evidence, page C999.

32. Also relevant are findings that show how false news can be more attractive than mainstream news. A recent study cited by representatives of the **Ukraine Crisis Media Centre**, Ms Nataliia Popovych and Mr Oleksiy Makhuhin, found that the level of Facebook interaction (i.e. comments, shares, and reactions) generated by a small number of false news outlets matched or exceeded that produced by the most popular news brands.

33. *Reasons for lag are difficult to overcome.* There are psychological reasons for why falsehoods are spread more than corrections.

34. **Dr Ullrich Ecker's** research showed that people seem to mainly pass on information that will evoke an emotional response, regardless of its truth or believability. Emotional arousal generally increases people's willingness to pass on information. Hence, stories likely to evoke disgust, fear or happiness are spread more readily and widely through social media than neutral stories.[30]

35. **Dr Carol Soon** highlighted that the recent MIT study on the spread of false news on Twitter, mentioned above at [24], had found that emotion played a fairly important role in why falsehoods tended to spread further and deeper than the truth.

36. Anecdotal evidence of the role of emotions in the spread of online falsehoods was given by disinformation expert **Mr Ben Nimmo,** with reference to the incident of the falsehoods spread during the 2017 Catalan independence referendum mentioned at [28] above. In his view, the event showed the difficulty of fake-busting in a heated and viral information environment, particularly where the falsehood plays into one's emotions.

37. He explained that a difficult problem in dealing with online falsehoods was what to do with the "willing audience", namely, those emotionally invested in believing that the fake is true, and are therefore willing to share it. There were also those who would knowingly share the false story in the belief that doing so served a higher purpose. In that regard, he referred to indicative evidence that many of those sharing the falsehoods attacking Mrs Hillary Clinton during the 2016 US Presidential Election did not believe in what they were sharing but hoped others would. Accordingly, there was the emotional investment both in wanting to believe the story, and wanting to spread it.

38. *Speed of online falsehoods a significant concern.* More generally, significant concerns were expressed about the speed at which falsehoods spread online. **Dr Elmie Nekmat** pointed out that the effects of deliberate online falsehoods in social media can occur rapidly and impact broad segments of society within a short period of time. Student **Zubin Jain** observed that while in the past, falsehoods

[30] Stephan Lewandowsky et al, "Misinformation and its correction continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, p 108.

could take hours to spread, social media has now removed the "grace period" and a message can be spread to the entire population in a mere couple of hours.

39. When asked by the Committee about why they, as young Singaporeans, were concerned about deliberate online falsehoods, a **group of SMU students and a lawyer** explained that their concerns stemmed from the speed of dissemination of online falsehoods and how damaging falsehoods could be in that very short period of time. Another **group of SMU students** echoed the same concern, noting how deliberate online falsehoods are shared at a very rapid speed across various forms of communication channels.

(4) Social transformations caused by the digital revolution

a. Online echo chambers

40. *Online echo chambers*. Several expert representors, including a psychological scientist, political data scientists and a computer scientist, gave evidence on the existence of online echo chambers, and their role in facilitating the influence of falsehoods online.

41. Online echo chambers have been described as the "fractionation of the information landscape",[31] according to research provided by **Dr Ullrich Ecker**. His research made the following points:

    a. In this phenomenon, blogs, which tended to be political, linked primarily to other blogs of similar persuasion and not to those with opposing viewpoints. There was research showing that half of blog readers sought out blogs that supported their views, while only 22% sought out blogs that espoused opposing views, creating so-called "cyber-ghettos."[32]

    b. The repetition of misinformation in social media echo chambers is particularly influential, because it can give rise to the wrong belief that there is high social consensus that the misinformation is true.[33]

42. Drawing on their research on the impact of online echo chambers created by algorithms, **Dr Carol Soon and Mr Shawn Goh** made the following points about echo chambers and how they reinforce confirmation biases:

    a. Algorithms used by social media platforms predict what people like based on what they consume and personalise their information exposure, thereby

[31] Stephan Lewandowsky et al, "Misinformation and its correction continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, p 108.
[32] Stephan Lewandowsky et al, "Misinformation and its correction continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, p 111.
[33] Stephan Lewandowsky et al, "Misinformation and its correction continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, p 113.

reinforcing filter bubbles and echo chambers in which they are exposed to information and opinions that are consistent with their pre-existing beliefs.

b. Findings of a study by data scientists at Facebook demonstrated that the filter bubble / echo chamber effect was real, even if smaller than expected, and that Facebook's algorithm increased people's chances of encountering information, including false information, that reinforced their world-view.[34]

c. This makes people over-confident in their mental frameworks, and dramatically amplifies people's confirmation biases.

d. Another study on the spread of misinformation on Facebook found that the homogeneity of echo chambers was the primary driver of misinformation online.[35] They may also increase group polarisation, where deliberation among like-minded people entrenches false information.

43. To elaborate on the Facebook study[36] mentioned at [42.b] above, that study examined how 10.1 million US Facebook users navigated the site over a six-month period. The study found that an average of 29% of the news stories displayed by Facebook's news feed presented views that conflicted with the user's ideology. It also found that individuals' choices of what information to consume had a stronger effect than Facebook's filtering algorithm. The results of the study were criticised, as noted by Dr Carol Soon and Mr Shawn Goh. One of these criticisms questioned the methodology of the study: as the 10.1 million users surveyed had self-identified as liberal or conservatives in their profiles, the results of the study could not be generalised to all Facebook users.[37] This is because people who self-identify their politics are likely to behave differently from those who do not. As Dr Soon and Mr Goh pointed out, the findings of the study at the very least demonstrated that Facebook's algorithm did result in a filtering effect.

44. The existence of online echo chambers was corroborated by evidence from expert representors Dr Simon Hegelich and Mr Morteza Shahrezaye, Dr Hany Farid, and Dr Kevin Limonier, all of whom had conducted relevant empirical research.

45. There is a "measurable effect of polarisation caused by uneven distribution of information"[38] in social networks, according to the empirical research of **Dr Simon Hegelich and Mr Morteza Shahrezaye** on Facebook and Twitter debates. Dr Hegelich and Mr Morteza Shahrezaye are political data scientists from the

[34] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 45.
[35] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017), p 45.
[36] Eytan Bakshy et al, "Exposure to ideologically diverse news and opinion on Facebook", *Science* (5 June 2015).
[37] Carol Soon and Shawn Goh, "What Lies Beneath the Truth: A Literature Review on Fake News, False Information and More", *Institute of Policy Studies* (30 June 2017) citing Farhad Manjoo, "Facebook Use Polarizing? Site Begs to Differ", *New York Times* (7 May 2015).
[38] Simon Hegelich and Morteza Shahrezaye, Appendix III: Written Representations, Paper No. 74, page B443.

Technical University of Munich, and Dr Hegelich was an expert invited by German Chancellor Angela Merkel to brief her political party on the phenomenon of social media manipulation. They highlighted that the reality of social networks was more complex than "simple explanations"[39] such as filter bubbles or echo chambers.

46. The role of social media algorithms in driving engagement with falsehoods was highlighted by **Professor Hany Farid.** Dr Farid described how algorithms are deliberately designed by humans to make decisions to engage users more, referring to it as the "algorithmic optimisation of engagement".[40] News that is sensational, presented as a conspiracy theory or fake, which people tend to be more engaged in, are being driven by such algorithms. Dr Farid agreed that if platforms are simply maximising for engagement, then platforms are vulnerable to manipulation. He cited some "startling analyses"[41] which showed that one can start with a video about the moon landing and within a few clicks, end up with a conspiracy theory about how the moon landing was faked. Another example cited was how one could start with a video about a moderate Muslim cleric and end up watching Al-Qaeda videos within five clicks.

47. **Dr Kevin Limonier** referred to the same phenomenon using the term "algorithm jail".[42] He described how he had conducted an experiment where he created fake profiles on Facebook to test Facebook's algorithms. The experiment involved "liking" pro-Russian media outlets and content to see what would happen to one's news feed. It found that a week later, only information of that nature appeared on the news feed. He concluded that, assuming Facebook was one's only source of information, the world would quickly become a "digital jail."[43] Users would only see the news and contents that the algorithms on social media networks decide for its users, such that users would only have one point of view. He caveated that Facebook may have changed their algorithms since, but it was difficult to tell as Facebook did not disclose how their algorithms worked.

48. Providing anecdotal evidence was the representative from Indonesian hoax-busting organisation **MAFINDO**, who described how groups of people with the same ideas are inclined to group together, leading to reduced levels of tolerance and amplification of polarisation.

49. A considerable number of other representors acknowledged the impact of online echo chambers.

50. Technology entrepreneur **Carlos Nicholas Fernandes** illustrated the impact of online echo chambers with a hypothetical: he wrote that if one were leaning

[39] Simon Hegelich and Morteza Shahrezaye, Appendix III: Written Representations, Paper No. 74, page B443.
[40] Hany Farid, Appendix IV: Minutes of Evidence, page C630, para 5354.
[41] Hany Farid, Appendix IV: Minutes of Evidence, page C630, para 5354.
[42] Kevin Limonier, Appendix IV: Minutes of Evidence, page C198, para 1693.
[43] Kevin Limonier, Appendix IV: Minutes of Evidence, page C198, para 1693.

slightly toward Donald Trump during the presidential campaign, and were provided with more positive views of Donald Trump with damning information on Hillary Clinton, one would automatically move further away from the center and closer to the most extreme supporters of Donald Trump. Thus, he concluded that while people may not have extreme positions, social media can seed our minds with ideas and beliefs, and also amplify existing gaps.

51. **Mr Zhulkarnain** observed that there is a risk that people see more content that reinforces their own thinking if they end up frequently interacting with posts and videos that reflect the similar views of their friends or family; thus creating echo chambers which create divisions of ideologies within a society. **Mr Gaurav Keerthi**, in the context of discussing a website that he had created to help break filter bubbles and echo chambers described filter bubbles as "basically the algorithms that Google designs or search engines design to provide you tailored information, but… end up feeding your predisposed biases".[44] As for echo chambers, he described them as "basically your social networks, they share stories that you already agree with so you don't get exposed to the other side".[45] A **group from Nanyang Polytechnic** explained how the echo chamber effect perpetuates group polarisation and implicit biases.

52. A **group of SMU students** noted how people may filter and receive information from certain preferred sources only, precluding an engagement with competing views that may provide the truth. They described this as the "echo chamber" effect, where participants in online communities constantly have their own opinions echoed back to them, which reinforces their original (potentially false) beliefs. They then remarked that given the way social media algorithms work, it may be impossible or unlikely for consumers to be provided with alternative information. For example, Facebook's algorithms are designed to populate users' news feeds with content similar to material previously "liked".

53. Other representors acknowledged the phenomenon more generally. **Associate Professor Eugene Tan** referred to the "personalisation algorithm"[46] and explained that it is responsible for the way people experience websites they visit, or when they receive targeted advertisements on social media. He noted that in today's age, different people are exposed to different realities because of their news feeds. **Mothership** explained that these computing codes fail to factor in human emotional complexity effectively and accurately, giving rise to "echo chambers".

b. Disruptions to the news ecosystem

54. *Lowering of barriers to entry for anyone to publish*. The barriers for non-professional sources of news to enter the news ecosystem, regardless of their

[44] Gaurav Keerthi, Appendix IV: Minutes of Evidence, page C459, para 3994.
[45] Gaurav Keerthi, Appendix IV: Minutes of Evidence, page C459, para 3994.
[46] Eugene Tan, Appendix III: Written Representations, Paper No 150, page B1310, para 10.

quality, have been lowered. Anyone can publish news on the Internet. People are increasingly relying on social media as a source of their news.

55. **Dr Ullrich Ecker** explained how the Internet had facilitated the spread of misinformation as it had side-lined the use of conventional "gate-keeping" mechanisms, such as professional editors.[47] He highlighted the lack of editorial gate-keeping and commitment to journalism ethics and standards. He further observed that Internet users have moved from being passive consumers of information to actively creating content on social media and blogs.[48]

56. The rise of "citizen journalists" was described by **Dr Thio Li-ann.** She stated that anyone can be a "citizen-journalist" if one has access to the Internet. She mentioned that such a "citizen-journalist" is not subject to the rigors of checking mechanisms and editorial oversight in ensuring the veracity of information. Dr Thio also observed that there are few ethical guidelines or constraints on those who play informational roles via social media, in contrast to the ethos of professional journalists. She further cautioned that the anonymity of the Internet may lead to publication in a reckless or negligent fashion.

57. Other representors acknowledged the point more generally. **Mr Ben Nimmo** highlighted that social media has allowed peer-to-peer interactions on an unprecedented scale by allowing malicious actors to bypass traditional editorial verification and spread their falsehoods unchecked. A **group from Nanyang Polytechnic** also made a similar point, explaining that advancements in technology have made it simple for individuals to post their views on social media platforms.

58. The growing reliance of many on social media as their main source of news was highlighted by a considerable number of representors. **Dr Claire Wardle** noted the fact that social feeds, rather than news websites, are often people's direct connection to news. **Accountancy student Mr Chua Jun Hao** cited a 2017 Reuters article which showed that the majority of people obtain their news online, via social media. **Ms Nataliia Popovych** and **Mr Oleksiy Makhuhin** wrote that audiences worldwide rely on the Internet and social media as primary sources of news and information. **Mr Calvin Cheng** stated that Google, Twitter, Facebook and Wikipedia have become the go-to sources for information globally. **Dr Ullrich Ecker** noted that many even regard blogs and social media posts as "trustworthy sources of information".[49] **Ms Jennifer Yang** noted that many Indonesians increasingly prefer the views and opinions from personal networks, seeing communication from the government and mainstream media as less

[47] Stephan Lewandowsky et al, "Misinformation and its correction continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, p 110.
[48] Stephan Lewandowsky et al, "Misinformation and its correction continued influence and successful debiasing", *Psychological Science in the Public Interest* (2012) 13(3) 106, p 110.
[49] Ullrich Ecker, Appendix III: Written Representations, Paper No. 44, page B183, para 15.

trustworthy. A **group of SMU students** was of the view that Facebook is possibly where most young people get their access to news nowadays.

59. In the context of Singapore, **Mr Zhulkarnain** cited the Reuters Institute Digital News Report 2017 which found that, in terms of news consumption, 61% of Singaporeans obtained their news from social media, with Facebook and WhatsApp being the preferred social media and messaging apps.

60. *Expectation of real time news.* The impact on the news industry of the instantaneous nature of the spread of information online was highlighted by **Professor Gerald Steinberg from NGO Monitor**. He noted that to keep up with social media, journalists were finding themselves under pressure to report things that they may not have had the time to fully verify. By taking the time to be careful to verify the details of an incident, such as a terrorist attack, mainstream media was losing power to other actors who did not feel so constrained.

61. *Disruption of business model of newspapers*: The business model of newspapers has been disrupted. **Editor of The Straits Times, Mr Warren Fernandez**, explained that the business model for media had been fundamentally disrupted. This business model depended largely on print or digital advertising. However, advertising revenue was being channelled away to only a few key players. In his view, the result of this was challenges to the ability to respond to news developments, including "fake news." He emphasised that sustaining his newspaper's newsroom required "tremendous resources",[50] and the ability of newspapers to verify facts was heavily resource intensive. Without the ability to sustain their news operations, the newspaper would not be able to continue playing their role. He felt it important to consider the business models for quality journalism.

62. Former news editor and journalist **Mr Nicholas Fang** highlighted that traditional media platforms were facing financial pressures due to rising competition from digital media, and an "almost infinite"[51] number of other sources of news online, which made the model of charging consumers for access to news increasingly unrealistic. In his view, these financial pressures diverted attention from the media companies' role of delivering quality journalism. He referred to the "reality of pressure to attract more eyeballs",[52] as consumers turned to social media and other channels for their news and information. The need to grow revenues had increased pressures to cater to the demands of the majority of consumers, often resulting in a rush to the "lowest common denominator of popular demand"[53] and the rise of "clickbait" to draw advertising dollars.

[50] Warren Fernandez, Appendix IV: Minutes of Evidence, page C497, para 4290.
[51] Nicholas Fang, Appendix III: Written Representations, Paper No. 144, page B1276.
[52] Nicholas Fang, Appendix III: Written Representations, Paper No. 144, page B1276.
[53] Nicholas Fang, Appendix III: Written Representations, Paper No. 144, page B1276.

63. Other representors alluded to how these digital transformations had led to shifts in the way the mainstream media reported the news. During the hearing, Dr Carol Soon agreed that international news media had started catering their headlines to audiences' interests and desires; in other words, they were providing people with the news that they want to read, or want to hear.

64. The impact of the digital revolution on consumer preferences was also raised by **the representative from Channel NewsAsia, Mr Walter Fernandez**. He stated that echo chambers built by algorithms have created a system where people want news that resonates with their own personal view or the view of their friends.

65. The general impact of the digital revolution on the mainstream news industry was referred to by **Dr Ullrich Ecker**. In a joint research article on understanding the "post-truth" era, he elaborated on the "rapid transformation of the media landscape."[54] He observed that with the "plethora"[55] of voices online today, the number of journalists working for daily papers in the US had dropped from around 55,000 in 2000 to 33,000 in 2015.

c. *Transformation of political discourse*

66. Political data scientists **Dr Simon Hegelich and Mr Morteza Shahrezaye** expounded on how the digital revolution had transformed political discourse. In their view, "[n]ever before has the political communication of so many people changed in such a short time."[56] Their analysis made the following points:

a. In democracies, there is a private sphere and a public sphere. In the private sphere, citizens follow their private interests and motivations. In the public sphere, the focus is on the general welfare or public good. The distinction between the public and the private sphere was a conceptual one, but necessary for a democracy.

b. What is wrong from a public point of view may be right from the private point of view, and vice versa. The public sphere requires the integration of contradicting private interests. (Dr Hegelich and Mr Shahrezaye also take the view that there should be no public regulation of falsehoods because this would involve determining what is right and wrong for a private person. The incorrect conflation of a false statement of fact with moral notions of right and wrong is dealt with in Part II.A.4.)

c. A change in the technical means of communication is necessarily a fundamental change for democracy. Historically, the invention of the printing

[54] Stephan Lewandowsky et al, "Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era", *Journal of Applied Research in Memory and Cognition 6* (2017) 353, p 359.
[55] Stephan Lewandowsky et al, "Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era", *Journal of Applied Research in Memory and Cognition 6* (2017) 353, p 359.
[56] Simon Hegelich and Morteza Shahrezaye, Appendix III: Written Representations, Paper No. 74, page B441.

press, advertising, daily newspapers, radio and television had changed democratic discourse.

d. The digital revolution today had made the amount of information in the public sphere "explode",[57] and also made social media the new channel of private communication. Also, decisions about what should be public were today made increasingly by algorithms. The institutions that used to safeguard the distinction between the public and private spheres, such as the media, were losing influence.

e. Social media was not designed for political communication. It was intended to connect private persons to increase their outreach. Communication on social media is guided by private affinity and emotions. In contrast, political discourse should not be so convenient, but should be the result of debates and compromise between legitimate interests.

f. Political communication on social media has been vulnerable to manipulation and polarisation.

g. However, the importance of social media for political communication is likely to grow. There is a need to learn how to use these platforms for political communication.

67. In a similar vein, how deliberate online falsehoods undermine the democratic concept of the "marketplace of ideas" was discussed by **Dr Thio Li-ann**. Professor Thio made the following points in her analysis:

a. The mainstream media operates as a public forum, exposing people to a wide range of speakers, unanticipated topics and viewpoints, and exposing viewpoints to a diverse public. This would allow citizens to engage with a range of representative views of issues, in order to understand where other citizens are coming from, and for facilitating compromise and overlapping consensus where possible.

b. However, with people now choosing to go online to obtain their news, they are denied this exposure to differing viewpoints. Technology today allows people to filter the kind of news we want to hear. By customising the news one receives, this is harmful to a well-functioning democracy insofar as it is important to be exposed to and engaged with viewpoints and topics through unanticipated encounters one cannot control (*e.g.*, the reader cannot control the type of articles a paper publishes).

c. There is a difference between a physical town hall and the Internet as a space for discussion. In a physical town hall, everyone could see one another's

[57] Simon Hegelich and Morteza Shahrezaye, Appendix III: Written Representations, Paper No. 74, page B441.

facial features, reactions and non-verbal speech. These are all not present on the Internet. Without these kinds of filters and self-restraint, the very worst impulses can come out of people. In online interactions, there is an absence of the human factor to moderate how people communicate.

68. The impact of online echo chambers on political discourse was recognised by several representors.

69. **Mr Nicholas Fang**, wrote about how technology-enabled filters automatically feed users with information they show a prior preference for. According to Mr Fang, this then creates a society where people only see parts of issues and not the broader picture, which then impedes the formation of viable solutions, and any coherent debate or discussion. The **National Council of Churches of Singapore** similarly described how social media algorithms decide what content is shown to users, causing groups of users to consume the same information and not exposing them to alternative information or opinion. Eventually, this leads to serious distortions of public debate.

70. Lawyers, **Mr Dan Shefet** and **Mr Darius Lee** had similar views. Mr Shefet said that in his view, the real threat to democracy came from the fact that everyone gets different news because of the filter bubble effect. Mr Lee wrote about how the use of filters to selectively feed stories to users based on their preferences has been shown to promote greater balkanisation and polarisation of society into ideological echo chambers.

71. The impact of anonymity on political discourse was also addressed. The **group from Nanyang Polytechnic** also highlighted the lack of accountability on individuals on the Internet, noting how perpetrators are able to hide behind anonymity or fake identities. Lawyer **Mr Darius Lee** also observed how the internet enables users to hide behind the anonymity of cyberspace, thus reducing the need for accountability in delivering one's ideas. The representative from **MAFINDO** similarly observed that the information ecosystem exploits the anonymity allowed on the Internet and social media.

72. More generally, how social media is not well-suited for political communication was acknowledged. **Mr Gaurav Keerthi** remarked that social media is not optimised or designed for robust discussion and debates of policy issues, but instead designed for social interactions, social networking and connecting with friends. According to the representative from **MAFINDO**, many people still have the false impression that they are free to speak anything on social media without consequences.

ANNEX E: DISINFORMATION OPERATIONS ALLEGEDLY CONDUCTED BY RUSSIA

1. This annex sets out the evidence received by the Committee of disinformation operations allegedly conducted by Russia, together with their alleged impact. As the Committee has consistently clarified, the Committee is not in a position to draw any conclusions on whether any country is indeed responsible for the alleged actions or intentions attributed to them by others. It is also not within the Committee's remit to assess whether these alleged actions were conducted for geopolitical or other reasons. Statements set out below should be regarded as statements made by representors. These statements do not reflect the Committee's views.

a. Motivations and Strategies

2. According to the US Senate Committee on Foreign Relations, Russia has sought to exacerbate divisions in Western democracies, weaken their democratic systems, and amplify their perceived weakness and problems.[1] This is apparently to prove that the Western democratic model is not worth pursuing, thereby increasing its own relative power.[2] In Mr Deynychenko's analysis, Russia's disinformation operations also seek to "reduce [the State's] ability to resist Russian aggression, change its foreign policy and create conditions for its inclusion in [Russia's] sphere of influence",[3] especially in Eastern Europe. This is apparently to fulfil Russia's long-term goals of: (a) creating or re-establishing the Russian empire in accordance with the borders of the former USSR; and (b) re-establishing its influence with all Russian-speaking people, not only in Russia, but also abroad.

3. *Russia's "perpetual state of war"*. UCMC gave evidence that Russia views their information operations as perpetual regardless of their relations with any government. The essence of this "hybrid", or "non-linear", war is to be able to wage war without officially announcing it.[4] The main battlefield, in this form of warfare, is "the mind of the enemy",[5] and information operations become of strategic importance. According to Dr Raska, the goal is to manipulate the adversary's perceptions, shape its decision-making process, and strategic choices, while minimizing the scale of kinetic force needed.

4. *The "Gerasimov Doctrine"*. The "Gerasimov Doctrine" – named after the current Chief of the General Staff of the Armed Forces of Russia, Valery Vasilyevich Gerasimov – sets out what many claim to describe how and why Russia uses disinformation operations. Dr Shashi explained that the doctrine emphasizes the

[1] "Putin's Asymmetric Assault on Democracy in Russia and Europe: Implications for US National Security", *US Congress* (10 January 2018), p 99.
[2] Edward Lucas and Peter Pomeranzev, "Winning the Information War", *Center for European Policy Analysis* (2016), p 2.
[3] Ruslan Deynychenko, Appendix III: Written Representations, Paper No. 78, page B469.
[4] Peter Pomerantsev and Michael Weiss, "The Menace of University: How the Kremlin Weaponises Information, Culture and Money", *Institute of Modern Russia* (September 2013), p 29.
[5] Michael Raska, Appendix III: Written Representations, Paper No. 97, page B950.

uses of propaganda and subversion as a military tool to achieve the aims of an aggressor State. The doctrine recognizes that "the information space opens wide asymmetrical possibilities for reducing the fighting potential of the enemy";[6] and that modern warfare is now conducted by a rough 4:1 ratio of non-military to military measures. Properly effected, disinformation operations can transform "a perfectly thriving state ... in a matter of months and even days ... into an arena of forced armed conflict, become a victim of foreign intervention and sink into a web of chaos, humanitarian chaos and civil war."[7]

5. *One tool amongst many*. According to Mr Janda, disinformation operations form only one part of Russia's complex toolkit of instruments used to undermine the sovereignty and security of a target State. Other non-military measures include economic pressure, disruption of diplomatic ties, and supporting radicals and extremist groups in the target countries. [8]

b. Actors and platforms relied on

6. In his written representation, Mr Janda describes how Russia uses multiple platforms and actors to conduct disinformation operations. These various modalities are set out in detail below.

7. *State-sponsored media outlets*. It is often claimed that Russian media outlets like *Russia Today (RT)* and *Sputnik* act as megaphones for Russia in spreading disinformation. These sponsored outlets are allegedly highly effective and difficult to combat for the following reasons:

   a. They target the popular medium in which majority of the population receive news. In Ukraine, television is the dominant news medium (followed by the Internet). Almost all Ukrainians (96.8%) watch TV for news at least weekly compared to just 48.3% going online for news.[9] According to Mr Deynychenko, Russian media outlets have sought to capitalize on this and attempted to spread their television signals into Ukrainian territory by improving their technical capabilities (*e.g.* installing more powerful transmitters).

   b. They cut across various language barriers: Dr Limonier observed that *RT* now broadcasts its content in at least 6 languages, including on TV cables in 4 countries. *Sputnik*, on the other hand, has a radio and Internet news service broadcast in 34 languages including common ones (English, French, Chinese and Spanish) as well as those which are rarer (Georgian, Latvian, Dari).

[6] Shashi Jayakumar, Appendix III: Written Representations, Paper No. 59, page B329.
[7] Shashi Jayakumar, Appendix IV: Minutes of Evidence, page C123, para 1088.
[8] Kremlin Watch Program of the European Values Think-Tank, "Kremlin Hostile Disinformation Operations. Situational report on Czech Republic and Central European context", *European Values Think-Tank* (18 October 2016), p 4.
[9] "Contemporary Media Use in Ukraine", *Gallup, Broadcasting Board of Governors* (2014), p 1.

c. They have creatively modified their means of engagement. Compared to Cold War-era propaganda, modern Russian propaganda today is claimed to be "enjoyable" and "emotionally engaging".[10] Mr Deynychenko observed that daily news on these media outlets have been substituted with engaging political talk shows. "Guests" who are introduced as "experts" are invited on these talk shows to spread the Russian narrative.[11] To augment the perceived credibility and blur the line between trustworthy broadcasting and disinformation, these outlets have allegedly also recruited well-known media and journalism personalities from US and Europe to front the shows.[12]

d. They lay claim to traditional liberal-democratic ideals of free speech, critical journalistic inquiry and independent thought. These outlets allegedly exploit the ideals of freedom of information and expression to inject disinformation in target societies. For example, RT's conspiratorial ethos is encompassed by its slogan of "*Question More*". While it appears to advocate media literacy, encourage people to think critically and maintain a healthy scepticism about media content, the underlying message attempts to suggest that any mainstream narrative in the news cannot be trusted.[13] The alleged goal is to systematically influence populations to become less trusting of mainstream, established news networks, and to "choose the side of the freethinkers and support Russia, portrayed as the ideal country".[14]

8. *Social media.* Various analysts have alleged that Russia has effectively exploited the anonymity, ambiguity, ubiquity and flexibility of the Internet, in particular social media, which was unavailable and unimaginable during Soviet times.[15] Social media acts as a cheap distribution channel or gateway to Russian media outlets. Because they are designed to hijack users' attention, it makes them excellent conduits for the dissemination of falsehoods.[16] Dr Limonier observed that an important share of visits on the websites of RT and Sputnik comes from redirections from social networks. These Russian media outlets allegedly attract their audience by publishing "quirky" articles with catchy titles, and sensational or emotional content (usually having little to do with their editorial line) on social media networks. The main intention, according to Dr Limonier, is to get users re-directed to their own websites.

---

[10] Edward Lucas and Peter Pomeranzev, "Winning the Information War", *Center for European Policy Analysis* (2016), p 9.
[11] Ruslan Deynychenko, Appendix III: Written Representations, Paper No. 78, page B472.
[12] Monika Richter, "The Kremlin's Platform for 'Useful Idiots' in the West: An Overview of RT's Editorial Strategy and Evidence of Impact", *European Values Think-Tank* (18 September 2017), p 24.
[13] Monika Richter, "The Kremlin's Platform for 'Useful Idiots' in the West: An Overview of RT's Editorial Strategy and Evidence of Impact", *European Values Think-Tank* (18 September 2017), pp 13, 14.
[14] Kevin Limonier, Appendix III: Written Representations, Paper No. 73, pages B430-431.
[15] Edward Lucas and Peter Pomeranzev, "Winning the Information War", *Center for European Policy Analysis* (2016), p 10.
[16] "The Fake News Machine", *The Economist (published in The Straits Times)* (4 March 2018).

9. *Bots and Trolls*. Evidence was received by the Committee on how Russia allegedly uses bots and trolls prolifically to spread and amplify falsehoods:

    a. Use of bots: Mr Deynychenko observed that Russian propaganda has significantly increased its activity in the social media networks through bots. Dr Limonier has used data obtained from Twitter to identify thousands of "French" accounts which had relayed Russian propaganda from Russian-linked platforms, many of which exhibited behaviours similar to what one would expect of a bot.

    b. Use of trolls: The US Senate Committee on Foreign Relations has accused Russia of employing individuals who would set up thousands of fake social media accounts to derail online debates and amplify pro-Russian narratives.[17] According to a New York Times investigation, in 2015, hundreds of young Russians were employed at a "troll farm" in St Petersburg known as the Internet Research Agency ("IRA") where many worked 12-hour shifts in departments focused on different social media platforms. These "trolls" earned between $800 to $1,000 a month, an attractive wage for recent graduates new to the work force. They were trained to provoke unrest and discontent amongst Americans on social media, by leveraging on hot-button issues or policies in the US.[18] Further details on the activities of the IRA are set out below at [25]-[28]. Many of these Russian "troll farms" are reported to have spread pro-Kremlin messages on the web, attacked Russia's opponents and drowned out constructive debate online.[19]

10. *Use of local actors to amplify content.* According to Mr Janda, the disinformation produced by the Russian state media would not have had the same significant effect if not for the ecosystem of local actors in the target country; whose interests converge with that of Russia. These local actors – whether knowingly or not – allegedly assist in the penetration of information space by the Russian state media, through their circulation of content. In this regard, Dr Limonier presented evidence on the "galaxy" of Twitter users who allegedly took part in the propagation of discourse produced by Russian platforms. Dr Limonier found that the "Russosphere" was not homogenous, either based on the individuals' profile or their political orientation, such that a large part of the discourse could be said to exist without any action from Russia. What links these users and discourse together to form a coherent whole were several "central" accounts – the Russian media outlets and the accounts of political personalities.

[17] "Putin's Asymmetric Assault on Democracy in Russia and Europe: Implications for US National Security", *US Congress* (10 January 2018), pp 43-44.
[18] "Putin's Asymmetric Assault on Democracy in Russia and Europe: Implications for US National Security", *US Congress* (10 January 2018), pp 44-45.
[19] Edward Lucas and Peter Pomeranzev, "Winning the Information War", *Center for European Policy Analysis* (2016), p 10.

11. The Committee will now set out the experience and impact of four countries which were allegedly targeted by Russian disinformation operations. They are: (a) Ukraine, (b) the Czech Republic; (c) the United States of America; and (d) France.

c. Ukraine: Experience and Impact

12. *Overview.* According to UCMC, Russia sees a sovereign and independent Ukraine as an affront to its nationalistic aspirations. It has been claimed that the objective of Russian disinformation operations in Ukraine is to destabilize it psychologically and to advance a conviction that it is a failed State. This is allegedly to destroy both domestic and international support for a Ukraine that is independent from Russia;[20] ultimately weakening the country's resistance to Russian influence and aggression. Being a neighbouring State, with a huge proportion of Russian-speaking people, Russian disinformation operations in Ukraine are said to have achieved considerable success.

13. The disinformation tactics allegedly used include targeting groups vulnerable to Russian influence, using falsehoods to support overarching and emotive narratives to confuse and demoralise the Ukrainian population, including its armed forces, and leveraging "useful idiots", *i.e.* opinion leaders among local academia, think thanks, politicians, community leaders, to advance the narratives and make these narratives appear as belonging to the locals.

14. *Impact.* The Committee received evidence of the following impact experienced by Ukraine, as a result of disinformation operations allegedly conducted by Russia.

15. *Fuelled existing tensions between different communities*. Mr Deynychenko gave evidence of how Russian disinformation operations have targeted and fuelled existing tensions between different groups of people, by focusing on historical examples of conflict between them. This was apparently the result of a sustained campaign of lies, rumours and disinformation being spread on how, for example, the Russian-speaking minority in Ukraine are the subject of persecution.

16. *Discredited Ukraine's standing in other EU countries*. Disinformation was also allegedly disseminated not just in Ukraine but in other countries, particularly neighbouring European countries, to discredit Ukraine's standing in the EU:

    a. Disinformation affecting relationship with Sweden: A forged official letter from Sweden's Ministry of Justice was circulated online to suggest that Ukraine had sought to improperly influence a case involving war crimes before the Swedish courts. This letter was allegedly disseminated by Russian media and had reached the Swedish public,[21] to undermine the support among the Swedish public for Ukraine.

[20] Edward Lucas and Peter Pomeranzev, "Winning the Information War", *Center for European Policy Analysis* (2016), p 15.
[21] "Fake Swedish letter in Russian media", *StopFake* (15 September 2015).

b. Disinformation affecting relationship with Poland: According to Mr Deynychenko, the Russian media had at one point deliberately played up the historical relations and conflicts between Ukraine and Poland. This led to radicals in both countries burning the flags of the other country, desecrating monuments and military cemeteries, with active coverage of these events by the press in Poland and Ukraine.

c. Disinformation affecting relationship with Netherlands: Russian media outlets had allegedly spread the falsehood that the Ukrainian military had shot down Flight MH17,[22] which led to the death of 193 Dutch citizens. When Netherlands held a referendum in April 2016 to approve a trade agreement between EU and Ukraine, the referendum saw a relatively low turnout (just 32% of Dutch population), with about two-thirds voting against the agreement.[23] According to a poll cited by a Ukrainian foreign ministry official, 59% of Dutch who voted against the trade agreement did so as they believed the Ukrainian government to be corrupt; and 19% of them believed the unproven claim that Ukraine had shot down MH17.[24]

17. *Loss of territorial sovereignty and lives*. According to Mr Deynychenko, the disinformation operations in Ukraine ultimately culminated in the loss of territorial sovereignty and Ukrainian lives – *i.e.* the annexation of Crimea, and the armed conflicts in other parts of Eastern Ukraine which claimed thousands of lives. Many of the Russian-linked fighters who fought in Ukrainian soil were reported to have disclosed that they were motivated to fight because of the Russian television coverage of supposed Ukrainian "atrocities" against Russian-speaking citizens. Mr Deynychenko also provided the following account of how disinformation operations in Ukraine were a prelude to armed conflict:

a. In March 2014, there was a large number of false news articles coming from Kremlin-controlled news sources about the presence of Ukrainian refugees at the Ukrainian-Russian border, using fake photos and videos of long lines of refugees which were taken elsewhere.
b. At the same time, Russia had prepared rooms for thousands of Ukrainian refugees.

c. Armed operations by Russian-backed forces commenced in Eastern Ukraine a month and a half later. This led to many victims being forced to leave their homes, and actually seeking asylum in Russia.

[22] Monika Richter, "The Kremlin's Platform for 'Useful Idiots' in the West: An Overview of RT's Editorial Strategy and Evidence of Impact", *European Values Think-Tank* (18 September 2017), p 21.
[23] Monika Richter, "The Kremlin's Platform for 'Useful Idiots' in the West: An Overview of RT's Editorial Strategy and Evidence of Impact", *European Values Think-Tank* (18 September 2017), p 35.
[24] "Putin's Asymmetric Assault on Democracy in Russia and Europe: Implications for US National Security", *US Congress* (10 January 2018), pp 113-115.

d. To Mr Deynychenko and his team at StopFake, this showed how Russia had all along plotted the organisation of an armed conflict in Eastern Ukraine, and had even prepared for the appearance of the refugees long before they existed.

d. Czech Republic: Experience and Impact

18. *Overview*. Although the Czech Republic has no historic Russian minorities, it is reported that pro-Kremlin disinformation still finds its way into the country through local voices in their native tongues.[25] It has been suggested that the goal of disinformation operations in the Czech Republic is to shift public opinion against the West,[26] and displace the influence of the EU and NATO in the country. According to Mr Janda, Russia has influenced extremists and fringe politicians in the Czech Republic to share and spread pro-Russian propaganda and disinformation. Disinformation has also been allegedly spread through pro-Russian websites, informal groups and communities of social media – and these networks online have managed to shift fringe views to the mainstream, thereby reaching and influencing a significant number of Czech citizens.[27]

19. *Impact*. The Committee received evidence of the following impact experienced by the Czech Republic, as a result of disinformation operations allegedly conducted by Russia.

20. *Undermined trust within the population*. According to Mr Janda, Russian disinformation operations have undermined the level of trust within the Czech population towards the Czech government, allied organisations and states, democratic political parties, and the mainstream media. As a result, 53% of Czechs believed that there was both pro-Russian and anti-Russian propaganda in the Czech public space and they could not trust anything.

21. *Influenced governance and policy options*. Falsehoods portraying the Ukrainian government as fascist have allegedly impeded the Czech government's ability to render humanitarian aid to Ukraine. According to Mr Janda's research, a quarter to a third of Czechs believe that the Ukrainian government is fascist. In Mr Janda's view, the falsehoods spread about Ukraine in the Czech Republic have impacted the policy options of the Czech Government. Mr Janda also warned that if the threat of disinformation and influence by Russia continues to be underestimated in the Czech Republic, one can expect Czech politicians to become more submissive to pro-Kremlin narratives; and public institutions to be penetrated by the Kremlin's influence.

[25] Edward Lucas and Peter Pomeranzev, "Winning the Information War", *Center for European Policy Analysis* (2016), p 33.
[26] Edward Lucas and Peter Pomeranzev, "Winning the Information War", *Center for European Policy Analysis* (2016), p 33.
[27] Edward Lucas and Peter Pomeranzev, "Winning the Information War", *Center for European Policy Analysis* (2016), pp 33-34.

e. United States of America: Experience and Impact

22. *Overview*. According to national security reporters from the *Washington Post*, after the Cold War, senior policymakers in the US wrongly assumed Russia would be a partner and largely pulled the US out of information warfare. In contrast, Russia – whilst weakened by the breakup of the USSR – had allegedly seized on influence campaigns and cyberwarfare as equalizers as both were cheap and easy to deploy, and hard for an open and networked society such as the US to defend against.[28] The complacency of the US left it unprepared to deal with Russian disinformation operations adequately. Even when the US was alerted to the Russian threat in 2014, senior US officials were reported not to have been "particularly alarmed by the threat, reflecting a widely held belief inside the US Government that its democratic institutions and society weren't ... as vulnerable".[29]

23. *Aims*. According to US intelligence agencies, Russia's strategic goal was to undermine the US-led liberal democratic order.[30] Disinformation operations were allegedly launched to undermine public faith in the US democratic process (*i.e.* the 2016 US Presidential elections), denigrate and harm Hillary Clinton's electability,[31] and sow discord and discontent in US society generally.[32]

24. *Key strategies*. Some of the key strategies of how Russian disinformation operations were allegedly conducted in the US are set out below.

25. *Use of covert, long-term, infiltration of local social media communities to gain influence*. One of the key strategies of Russian agents was allegedly to infiltrate US communities on social media by first ingratiating themselves with genuine members of the community, then using the approval of those members to take a stance as a representative member of the community. According to the Indictment by US Special Counsel Robert Mueller ("Mueller Indictment"), these activities began as early as 2014 by the IRA.[33]. IRA created false US personas, and operated social media pages and groups, which were designed to attract US audiences. Over time, they managed to reach significant numbers of Americans.[34] For example,

[28] Adam Entous et al, "Kremlin trolls burned across the Internet as Washington debated options", *The Washington Post* (25 December 2017).
[29] Adam Entous et al, "Kremlin trolls burned across the Internet as Washington debated options", *The Washington Post* (25 December 2017).
[30] "Assessing Russian Activities and Intentions in Recent US elections", *Intelligence Community Assessment* (6 January 2017), p ii.
[31] "Assessing Russian Activities and Intentions in Recent US elections", *Intelligence Community Assessment* (6 January 2017), p ii.
[32] Ben Nimmo, "Understanding the Role of Russian Propaganda in the US Election", *New Atlanticist, Atlantic Council* (17 August 2016); see also Robert Mueller, *Indictment by the United States Office of Special Counsel* (16 February 2018), para 6.
[33] Robert Mueller, *Indictment by the United States Office of Special Counsel* (16 February 2018), para 3.
[34] Robert Mueller, *Indictment by the United States Office of Special Counsel* (16 February 2018), para 4.

according to Twitter, the IRA managed at least 3,814 troll accounts; and 1.4 million American users are known to have interacted with these accounts.[35]

26. Two prominent examples of false US personas created by IRA for the purposes of disinformation operations are as follow:

    a. "Jenna Abrams" Twitter Account: This was a fake account of a non-existent person created by the IRA, using the image of a young American woman. At one point, "Jenna Abrams" had over 70,000 followers and was quoted by dozens of high-profile media outlets. Once the account had attracted a following, it started pushing divisive views on immigration, segregation, and Donald Trump, especially as the 2016 US Presidential election loomed.

    b. Fake Tennessee Republican Party ("TRP") Twitter Account: IRA had also impersonated the Tennessee Republican Party on Twitter, to repeatedly send out inflammatory falsehoods before it was finally shut down by Twitter. The fake TRP account gained 152,099 followers and posted a total 10,985 Tweets and Retweets, of which 9,852 were original Tweets (2,092 were posted during the 2016 US Presidential election time period). Original Tweets from this account received more than 67 million impressions within the first seven days after posting. In comparison, the *authentic* TRP account had only 13,800 followers and had Tweeted or Retweeted 8,768 times as of November 18, 2017. Of those, 200 were original Tweets that received 240,000 impressions within the first seven days after posting.[36] The Tweets of the fake TRP account were even amplified, inadvertently, by Retweets from the likes of Kellyanne Conway and Donald Trump Jr.[37]

27. IRA also created fake social media groups with the use of bots and artificial intelligence which over time, were populated by authentic supporters of the causes these groups championed. Two prominent examples were the "United Muslim of America" and the anti-Islamic "Heart of Texas" Facebook groups, which posted inflammatory posts that allegedly led to an actual public protest, the details of which are described below at [34] below.

28. *Production and purchase of political advertisements online to influence elections*. According to the Mueller Indictment, IRA and their co-conspirators had produced, purchased and posted advertisements on US social media and other online sites expressly advocating for the election of Donald Trump or expressly opposing

[35] "Update on Twitter's review of the 2016 US election", *Twitter Blog* (19 January 2018).

[36] "Sean Edgett's Answers to Questions for the Record", *Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism Hearing on Extremist Content and Russian Disinformation Online: Working to Find Solutions, October 31, 2017* (19 January 2018), pp 16-17.

[37] "Sean Edgett's Answers to Questions for the Record", *Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism Hearing on Extremist Content and Russian Disinformation Online: Working to Find Solutions, October 31, 2017* (19 January 2018), p 32.

Hillary Clinton.[38] Similar to many of the Tweets and posts of the fake social media accounts, these advertisements were able to reach a wide number of people. According to Facebook, between June 2015 and August 2017, fake accounts associated with the IRA spent approximately $100,000 on more than 3,000 Facebook and Instagram ads. An estimated 11.4 million people in the US saw at least one of those ads during the relevant period.[39]

29. *Impact.* The Committee received evidence of the following impact experienced by the United States, as a result of disinformation operations allegedly conducted by Russia.

30. *Inflamed social divides.* As mentioned earlier, online falsehoods spread by sources linked to Russia allegedly targeted already divisive issues in the US, such as race, LGBT rights, gun control, and immigration. Dr Shashi pointed out that these Russian-linked sources often targeted, and promoted, all sides of the political spectrum on controversial issues, for the purpose of simply turning different groups or communities against each other.

31. One example was how Russian trolls allegedly widened the divide between the "Black Lives Matter" supporters and the police in the US. It was reported that at least 29 known Russian trolls had tweeted about Black Lives Matter and police shootings, spreading divisive content widely over a nine-month period.[40] Some of the divisive content disseminated by Russian trolls include a message stating that activists working on the Black Lives Matter movement who disrespected the American flag should "*be immediately shot*"; while another suggested that "*Black people have to [practise] an eye for an eye. The law enforcement officers keep harassing and killing us without consequences*".[41] The fake TRP Twitter account also posted inflammatory materials which included anti-Muslim messages, and claimed that unarmed black men killed by police officers deserved their fate.[42]

32. *Undermined democratic process.* It is widely claimed that a key goal of Russian disinformation operations in the US was to attack Hillary Clinton and weaken her candidacy.[43] This has led to a perception that the US 2016 Presidential Elections had been interfered by foreign agents.

---

[38] Robert Mueller, *Indictment by the United States Office of Special Counsel* (16 February 2018), para 48.
[39] Testimony of Colin Stretch, *Hearing before the United States Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism Hearing on Extremist Content and Russian Disinformation Online: Working to Find Solutions*, October 31, 2017, pp 5-6.
[40] Denise Clifton, "Russian trolls stoked anger over Black Lives Matter more than was previously known", *Mother Jones* (30 January 2018); Kanyakrit Vongkiatkajorn, "How Russia exploited Black Lives Matter, Sean Hannity, and mass shootings", *Mother Jones* (17 February 2018).
[41] Curt Devine, "'Kill them all' – Russian-linked Facebook accounts called for violence", *CNN* (31 October 2017).
[42] Kevin Collier, "Twitter was warned repeatedly about this fake account run by a Russian troll farm and refused to take it down", *BuzzFeed* (18 October 2017).
[43] "Assessing Russian Activities and Intentions in Recent US elections", *Intelligence Community Assessment* (6 January 2017), p ii.

33. Some of the messages that were spread by the IRA during the election period included:

a. Allegations of voter fraud by the Democratic Party, spread through the fictitious US personas and groups on social media.[44]

b. A conspiracy theory that a mysterious explosion in Washington, DC, killed an employee of the Democratic National Committee, and that the death was linked to Hillary Clinton. This was shared by the fake TRP Twitter account.[45]

c. Anti-vote messages, containing false information such as "*American Muslims are boycotting elections today, most of the American Muslim voters refuse to vote for Hillary Clinton because she wants to continue the war on Muslims in the Middle East and voted yes for invading Iraq*".[46] This was shared by the "United Muslims of America" social media account.

34. <u>*Incited public protest*</u>. IRA also initiated the creation of opposing Facebook groups which allegedly triggered an actual standoff on the streets between supporters and opponents of an Islamic centre in Texas. Through the Facebook pages controlled by IRA – "*Heart of Texas*" and "*United Muslims of America*" – a protest and a counter-protest were organised on May 21, 2016, in Houston, Texas. Participants were urged to battle on the streets and to bring their firearms to the protest. The total cost to the IRA for this entire enterprise, which led to a clear security threat, was reportedly only US$200.[47]

35. <u>*Aftermath*</u>. According to US intelligence agencies, Russia will apply lessons learned from its disinformation operations aimed at the 2016 US Presidential election to influence efforts in the US and worldwide in the future. This is because Russia would have seen the 2016 US Presidential election influence campaign as at least a qualified success.[48] Despite this clear and present threat, it has been reported that the US is still struggling to find a coherent and effective response against Russian disinformation operations, due to its domestic politics and legal constraints in imposing effective countermeasures.[49]

---

[44] Robert Mueller, *Indictment by the United States Office of Special Counsel* (16 February 2018), para 47.

[45] Andrew Prokop, "23 tweets from @TEN_GOP, one Russian-run Twitter account mentioned in Mueller's new indictment", *Vox.com* (16 February 2018).

[46] Robert Mueller, *Indictment by the United States Office of Special Counsel* (16 February 2018), para 46(c).

[47] Natasha Bertrand, "Russia organized 2 sides of a Texas protest and encouraged 'both sides to battle in the streets'", *Business Insider* (1 November 2017).

[48] "Assessing Russian Activities and Intentions in Recent US elections", *Intelligence Community Assessment* (6 January 2017), p 5.

[49] Adam Entous et al, "Kremlin trolls burned across the Internet as Washington debated options", *The Washington Post* (25 December 2017).

f. France: Experience and Impact

36. <u>*Overview.*</u> According to Dr Kevin Limonier, tools of Russian State propaganda, comprising media outlets RT and Sputnik, were a "distinctive feature"[50] of the 2017 French Presidential Election. He shared with the Committee his research on the use of social media by RT and Sputnik to grow their influence in France.

37. <u>*Methods.*</u> Four methods used by RT and Sputnik were identified by Dr Limonier, as follows:

    a. *Grand narratives supported by selective editorial content*. The outlets allegedly sought to promote a narrative that cast the Western world as hegemonic, and Russia as a champion of free-thinking and multi-polarity. To do so, they allegedly published articles that exploited any information that could be used to discredit the US, EU or NATO, using catchy titles, and selectively omitting salient facts.

    b. *Use of social networks to reach people with different views.* An "important" share of visits on the websites of RT and Sputnik reportedly comes from re-directions from social networks. As mentioned at [10] above, Dr Limonier had carried out a preliminary mapping of the "galaxy" of Twitter users who relayed content from RT and Sputnik. The mapping showed that their content was able to reach a politically varied audience.

    c. *Manipulate social media to gain visibility.* As mentioned at [8] above, Sputnik and RT reportedly took advantage of the algorithms of social media by using "click-bait" to drive user engagement, and to promote the visibility of these outlets. The "click-bait" comprised sensational or emotive articles.

    d. *Bots.* Dr Limonier's research also found that bots (and trolls) were also a prominent feature of the sphere of Russian influence online. He found bots which typically engaged in abusive behaviour regularly relaying or interacting with the online platforms of Russian media outlets.

38. <u>*Impact*</u>. Dr Limonier's observation was that Russia had gained a "prime position"[51] in the geopolitics of cyberspace, and it had been increasingly successful. The narratives put out by RT and Sputnik had allegedly gained an "undeniable following"[52] in France, and were enjoying a growing audience in the West generally.

[50] Kevin Limonier, Appendix III: Written Representations, Paper No. 73, page B426.
[51] Kevin Limonier, Appendix III: Written Representations, Paper No. 73, page B438.
[52] Kevin Limonier, Appendix III: Written Representations, Paper No. 73, page B426.

ANNEX F: MEASURES TAKEN BY TECHNOLOGY COMPANIES

1. This Annex sets out a non-exhaustive list of the measures relevant to online falsehoods that Facebook, Google, and Twitter have said they are taking.

Facebook

Facebook's social media platform

2. According to Facebook, the following measures are being taken on its Facebook social media platform that are relevant to online falsehoods:

    a. Prohibiting inauthentic accounts, and requiring users to use their authentic names.

    b. Using algorithms to "down rank" content in News Feed that is inauthentic, including hoaxes and misinformation, and "click bait". It has also made updates to reduce the presence in News Feed of content from low-quality websites, such as those that produce "click bait", sensationalism or spam.

    c. Beginning tests in the US to prioritise news from publications rated by the community as trustworthy. Facebook has also made updates to reduce posts and ads in News Feed that are from low-quality websites, such as "click bait", sensationalism or spam.

    d. Testing a button that will allow people to easily access additional contextual information to articles shared in News Feed.

    e. Removing content that impersonates others. However, Facebook will not remove content on the basis that it is false.

Facebook Advertising

3. According to Facebook, the following measures relevant to online falsehoods are being taken on its advertising platform:

    a. Ensuring that spammers who make money by posting "click bait" cannot run advertisements carrying such "click bait" on Facebook. It will also ban repeat offenders from advertising on Facebook. Facebook Pages that contain mostly hoaxes and false news, and have "a large number of shocking, or malicious ads" may not be eligible to run ads, and their posts will show up lower in the News Feed.

    b. Making its advertising service more transparent, by enabling the public to view all the advertisements that a Facebook Page is running.

WhatsApp

4. WhatsApp is testing a "forwarded message" tag warning users when a message has been forward multiple times, indicating that it is spam.[1]

5. In July 2018, WhatsApp began testing a new feature globally, which limited the forwarding of messages, photos and videos to 20 chats at a time, whether among individuals or groups.[2] In India – where false information circulating on WhatsApp has led to a spate of violent incidents – a lower limit of 5 chats was set.[3] WhatsApp also removed the quick forward button next to media messages for its users in India.[4]

Google

Google Search and Google News

6. According to Google, it is taking the following measures relevant to online falsehoods on its search engine and news aggregator:

    a. Taking steps to prevent its Google Search algorithm from being exploited to amplify "poor quality or misleading" information, by "working to make improvements" to surface more high quality and credible results in response to their users' queries. However, Google will not remove content on the basis it is false, unless pursuant to a legally valid request.

    b. Introducing a "fact check label" in Google News and Google Search, which flags when a claim has been fact-checked by a publisher or fact-checker, and links to the fact check. A labelled article will also be shown next to a related article whenever possible.

    c. Introducing in the US "publisher knowledge panels", which informs users on topics covered by a publication and the awards it has received. Google aims to refine this feature and make it available globally.

Google Advertising

7. According to Google, it is taking the following measures relevant to online falsehoods on its advertising platforms:

    a. Not allowing misleading, inappropriate or harmful ads on Google Ads.

---

[1] "WhatsApp starts labelling forwarded messages, feature live on Android beta", *Indian Express* (9 June 2018).
[2] Alex Hern, "WhatsApp to restrict message forwarding after India mob lynchings", *The Guardian* (20 July 2018).
[3] "More changes to forwarding", *WhatsApp Blog* (19 July 2018).
[4] "More changes to forwarding", *WhatsApp Blog* (19 July 2018).

b. Google AdSense and DoubleClick prohibit website owners who misrepresent who they are and deceive users with their content from running advertisements.

YouTube

8. According to Google, improving its YouTube algorithms, so that in "breaking news" situations, they would prioritise authoritative sources over freshness and relevance.

Twitter

Twitter's social media platform

9. According to Twitter, it is taking the following measures relevant to online falsehoods on its social media platform:

    a. Implementing and continuing to develop technology to prohibit malicious automation, such as botnets, as well as accounts that display spam behaviour, or coordinated and abusive behaviour.

    b. Improving how it detects when accounts may have been hacked or compromised.

Twitter Advertising

10. According to Twitter, it is placing all advertisements run on its platform in a Transparency Centre.[5]

[5] Alex Kantrowitz, "Twitter will end dark ads and establish a 'transparency center'", *BuzzFeed News* (24 October 2017).

ANNEX G: MCCY'S RESPONSE TO THE SELECT COMMITTEE ON DELIBERATE ONLINE FALSEHOODS ON RECOMMENDATIONS ON GOVERNANCE AND STRENGTHENING PUBLIC TRUST

1. We received from the Select Committee a summary of recommendations regarding how we could strengthen trust between the people and the government. These recommendations revolved around the principles of communication, accountability, transparency and participation in the Government's policy and decision-making processes.

2. The Government agrees that these are important values that underpin good governance, which in turn allows for greater trust to be built between the government and the people. To engender trust, there needs to be commitment among all the parties involved – and not just government - to engage each other on the same principles of open communication, accountability, and integrity, as well as the sincere desire to serve the broader public interest taking into account the geo-political, social, and economic developments in and outside of Singapore.

3. Today, the Government provides mechanisms and platforms, and builds capability across the people, private and public sector so that there can be broader involvement among Singaporeans and organisations to partner the government and each other, to build the Singapore we want to see. These efforts speak to the recommendations received by the Select Committee, and the Government is heartened that we are on the right track. However, we acknowledge that there is always room for improvement and we will strive to do so, as a collective effort with Singaporeans. What follows provides information on the Government's current efforts to strengthen public trust, the institutional mechanisms and other ongoing initiatives that seek to ensure communication, accountability, transparency and participation.

<u>Institutional Mechanisms</u>

4. Embedded in Singapore's governance institutions are mechanisms for ensuring accountability, transparency and participation. For example:

a. The government of the day is elected through free and fair elections;
b. The Courts have the power to exercise judicial oversight over the Executive, to ensure that Executive action is in accordance with Singapore's Constitution and laws;
c. The Auditor-General's Office, an independent organ of state, enhances public accountability in the management and use of public funds and resources; and

d. Ministers explain the rationale of policies in Parliament while Parliamentarians, on behalf of their constituencies, can seek information from public institutions through Parliamentary Questions.

5. Beyond institutional mechanisms, there are also efforts to strengthen engagement and partnerships with citizens. Examples of these efforts are set out below.

Strengthening Engagement and Partnership with Citizens

6. The government has always taken the view that engaging and partnering our citizenry in developing and implementing policies and programmes foster stronger society. During the early days of nation-building, the government established the People's Association to engage the citizenry on the rationale behind various national policies. Several large-scale engagements provided opportunities for Singaporeans to contribute their ideas and partner with us. For example, in 2002, the Remaking Singapore Committee was formed to reshape the political, social and cultural norms of Singapore, looking beyond economics to understand the changing aspirations and expectations of Singaporeans. In 2006, the government restructured the Feedback Unit to form REACH, or "Reaching Everyone for Active Citizenry @ Home", to lead the government's efforts in engaging and connecting with citizens.

7. In recent years, we have refined these approaches but have been led by the same goals. We stepped up efforts to engage wider segments of the population, and involve them in diverse areas of policy and building our future Singapore. In 2012, Our Singapore Conversations was launched to engage Singaporeans on their hope and aspirations for Singapore; over 47,000 people participated. This was followed by SGfuture as part of the SG50 celebrations, where Singaporeans shared their ideas for a better Singapore, and came together to turn their ideas into action. In 2016, the Committee on the Future Economy saw over 9,000 businesses, members of the workforce and Singaporeans participate in shaping our economic future ahead. More recently, the Government has embarked on a discussion series to engage Singaporeans from all walks of life in charting the way ahead for Singapore.

8. The government also conducts extensive policy communication and consultations on a regular basis with stakeholders such as businesses, interest groups, and religious organisations. Before new Bills are introduced, consultations are also conducted with stakeholders directly affected by the changes, and a consultation document is put up on the REACH website to obtain public feedback. Beyond consulting citizens on policy design, we partner citizens to implement policies and develop programmes. More details are below

Policy Communication, Consultation and Co-Creation

9. To communicate and consult the public on policies, the government taps on both digital and offline platforms, including:

a. Gov.sg and associated platforms. The Gov.sg website and its social media platforms inform citizens of government-related news, initiatives and policies. Factually on Gov.sg also helps to clarify widespread or common misperceptions of government policies, so that citizens are better informed on issues that concern them. There are similar platforms and efforts in various Ministries. The government also engages partners and stakeholders to support these platforms and messages.

b. REACH. REACH engages the public, businesses and professional groups to understand their sentiments towards issues so as to enable government agencies to formulate better policies. Agencies consider the information collected from REACH platforms seriously in their decision-making process. In 2017 and 2018, REACH conducted over 50 public consultations on Bills and policies that have significant public interest. These include:

- Amendments to the Criminal Procedure Code and Evidence Act
- Approaches to Managing Personal Data in the Digital Economy
- Employment Act Review
- Healthcare Services (HCS) Bill
- Regulatory Framework for the Use of Private Residential Properties as Short-Term Accommodation Self-Employed Persons' Top Concerns in the Future Economy
- Tobacco Control Measures

Various stakeholder groups including the legal community, civil society, businesses, and members of the public contributed feedback in these consultations.

c. REACH also conducts Listening Points and dialogues in the heartlands and other key nodes to gather feedback from Singaporeans from all walks of life. From January 2017 to May 2018, over 150 Listening Points and dialogues were conducted to engage Singaporeans on a range of issues including transport, cost of living, jobs and economy, terrorism, cyber security, fake news, elected presidency and the President’s Address. These are complemented by online engagement via REACH’s Discussion Forum and

social media platforms, where REACH works with government agencies to address questions that Singaporeans have about various policies and issues.

d. PA's "Ask Kopi Kaki" (AKK) initiative and Community Kopi Talks. PA has been strengthening its engagement with citizens through its volunteer network and community partners to better understand citizens' needs and aspirations, as well as to share government policies with them. For example, the Ask Kopi Kaki (AKK) initiative provides a simple and accessible way for citizens to learn about government policies that are relevant to their needs, based on their life stage. AKK kiosks are available at community centres, and grassroots volunteers are also trained to help residents navigate government schemes so that they can get the support they need. PA staff and volunteers also gather and surface feedback to government agencies to improve policies and processes. For specific policies where there is a high-level of interest among the public, PA organises regular PA Kopi Talks at the national and community level, where policymakers share in greater depth their policy rationale, and listen to citizens' feedback.

e. National Steering Committee on Racial and Religious Harmony. MCCY works closely with partners to nurture community advocates who are able to rally different community segments in times of crisis, and show solidarity and unity. MCCY regularly engages apex religious leaders through the National Steering Committee on Racial and Religious Harmony, and religious and community leaders at the local level through the Inter-Racial & Religious Confidence Circles (IRCCs). Through workshops and exercises, MCCY has been partnering the IRCCs and religious organisations to build capability within their organisations and in the community to grow the skills and knowledge to strengthen trust among different communities, and community resilience.

10. Government agencies and advisory councils also regularly consult our stakeholders on an ongoing basis, as part of their policy reviews and implementation plans. For example:

a. MTI and its Statutory Boards engage trade associations and chambers, businesses (SMEs/MNCs) and students/youths through avenues such as dialogues and forums to gather feedback and sentiments on policies/announcements, such as the ratification of Free Trade Agreements and introduction of grants/programmes;

b. MOM engages NGOs and community groups on issues related to foreign workers and foreign domestic workers;

c. The Public Transport Council conducts focus-group discussions and regular surveys with commuters to understand their experience and gather feedback; and

d. HDB engages residents and community stakeholders through focus group discussions in the early stages of each phase of its Remaking Our Heartland (ROH) Programme. For its latest batch of ROH towns at Woodlands, Toa Payoh and Pasir Ris, HDB started its engagement with residents and stakeholders before its plans were formulated. Their views were then incorporated when developing the ROH plans for these towns. These plans were presented through a series of public exhibitions to take in feedback for further refinement.

11. In recent years, we also created avenues for citizens to be more deeply involved in designing policies or innovating new solutions to foster public trust and facilitate citizens' understanding of the policy rationale and trade-offs. There have been several successful efforts, and we will continue to introduce more opportunities. For example:

a. In 2016, MSF supported the Social Development Network (SDN) Council to conduct a Community Panel involving about 70 participants. The participants brainstormed ideas to address the challenges of singlehood and how to facilitate a dating-friendly environment. Experts also shared their insights on the dating and social landscape with the participants. Some of the ideas were incorporated into the bi-annual Spark Connections campaign organised by SDN;

b. In 2017/8, MOH's Citizens' Jury saw 76 citizens from diverse backgrounds deliberating on the issue of diabetes prevention and management. Over two months, they developed community-based solutions, before submitting their joint recommendations to MOH. A few participants in MOH's Citizens' Jury are going further to implement their own recommendations in their community; and

c. The government's Ideas! Portal crowdsources ideas and solutions from the community via an online platform. Several agencies have posted challenges for citizens to contribute ideas, including NYC's "Open Lab", which called for ideas from youths on how to create social good; the URA-REDAS Spark Challenge, a competition that called for ground-up innovations to raise the quality of the urban environment and also provided a platform for shortlisted projects to be tested by the public in commercial developments; and the Cool Ideas for Better HDB Living Initiative, where residents are encouraged to co-create innovative solutions to improve the HDB living environment and to foster stronger community involvement.

Other Partnerships with Citizens

12. Other than policy formulation, the government partners with citizens and stakeholders to better communicate its policies, deliver services, and improve solutions for the community. For example:

a. Many public agencies have volunteer programmes to involve citizens in programme delivery. For example, the Silver Generation Office equips and supports its volunteers - the Silver Generation Ambassadors - to deliver personalised, last-mile communication and outreach for the Pioneer Generation Package and other relevant government schemes. Silver Generation Ambassadors are trained to explain policies and schemes that are in place to support our seniors, and through the Community Networks for Seniors programme, to bring together various services from our community organisations, voluntary welfare organisations and government agencies to serve our seniors better;

b. NParks' Friends of the Parks scheme is a ground-led initiative that enables local communities including residents, recreational groups such as hikers and bikers, tenants, nature groups, researchers and regular park users to collaborate and lead initiatives to promote the active and responsible use of the parks. These initiatives include developing educational and awareness programmes and activities, and conducting habitat enhancements, guided walks, biodiversity surveys, among others;

c. URA has consulted the community extensively to develop the master plan for the 24 km-long Rail Corridor. Following a series of community exhibitions and workshops to collect public feedback to refine the plans, works have commenced on a 4 km signature stretch between the conserved Bukit Timah Railway Station and the Hillview area. This is the first step towards the Corridor becoming an exceptional and inclusive community space for people of all ages and abilities;

d. The SG Cares movement rallies corporates, the community and public agencies to champion causes and work together, promoting and facilitating active volunteerism and philanthropy, supporting ground-up initiatives through funding and other resources, and fostering partnerships among local stakeholders at the town level to create greater social impact.

Independent indicators of public trust

13. Current international assessments of public trust in Singapore are encouraging. For example:

a. According to the 2018 Edelman Trust Barometer, the Singapore general population's trust in public institutions is around 58%, which is 10 percentage points ahead of the global average, and even further ahead of countries such as the US (43%), Australia (40%) and the UK (39%);

b. The World Economic Forum's 2017-2018 Global Competitiveness Report ranked Singapore 2nd out of 137 countries in transparency in government policy-making;

c. Singapore continued to be perceived as having the lowest levels of corruption among 14 countries in Asia, as well as Australia and the United States, according to the 2018 Asian Intelligence Report of the Political & Economic Risk Consultancy.

14. Although public trust in Singapore is still high, it can easily be eroded by deliberate online falsehoods. To sustain and strengthen trust and partnerships with citizens, the government is continuing to grow its engagement capabilities and engage citizens as an integral part of their work. Citizen engagement training has been stepped up for public officers. To encourage innovation in citizen engagement approaches, the Citizen Engagement Seed Fund was set up in 2016 and has supported 16 new engagement projects by various agencies so far.

15. The government is committed to forge stronger partnerships and engagement with citizens, and create the best environment for citizens to build deep relationships and collaboration among themselves and with others, for greater impact. This shared responsibility to work for a common good is needed to build the best future we can for Singapore and Singaporeans.

Appendix I

MINUTES OF PROCEEDINGS

1st Meeting

Tuesday, 16 January 2018

2.00 pm

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Dr Janil Puthucheary
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

ABSENT

Ms Chia Yong Yong
Mr Desmond Lee

---

1. The Committee deliberated.

2. *Agreed* –

   (a) that the general public be invited to submit written representations to the Committee;

   (b) that the closing date for the submission of written representations be Wednesday, 28 February 2018;

   (c) that the invitation be advertised in the four local vernacular newspapers and published on the Parliament website;

   (d) that a press statement on the invitation for written representations to the Committee be issued; and

   (e) that Ministry officials be admitted to subsequent meetings of the Committee.

Adjourned till 9.30 am on
Monday, 5 March 2018

2nd Meeting

Monday, 5 March 2018

9.00 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Desmond Lee
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Ms Sun Xueling
Mr Edwin Tong Chun Fai

ABSENT

Mr K Shanmugam

In Attendance:

Ministry of Law:
Ms Lim Hui Min, Delphia, Senior Assistant Director, International Legal Division

Ministry of Communications and Information:
Mr Wong Zhilong, Assistant Director, Information Policy Division

1. The Committee deliberated.

2. Written representations received were considered.

3. *Agreed* –

   (a) that Papers 1 to 23, 25 to 102 and 104 to 107 be published.

   (b) that the Committee do meet to hear oral evidence on the following dates and times:

      (i) Wednesday 14 March, 11 am to 5.30 pm

      (ii) Thursday 15 March, 10 am to 5.30 pm

      (iii) Friday 16 March, 10 am to 5.30 pm

      (iv) Thursday 22 March, 10 am to 5.30 pm

(v) Friday 23 March, 10 am to 5.30 pm

(vi) Tuesday 27 March, 10 am to 5.30 pm

(vii) Wednesday 28 March, 10 am to 5.30 pm

(viii) Thursday 29 March, 10 am to 5.30 pm

Reserve date: Saturday 24 March, 10 am to 5.30 pm

(c) that if the Chairman is unable to be present for the Select Committee meetings on 14, 15, 16, 22, 23, 24, 27, 28 and 29 March 2018, Mr Seah Kian Peng be elected to act as Chairman on those dates.

(d) that the following representors be invited to give oral evidence:

(1) Mr Howard Lee (*Paper 12*)

(2) Mr Hazrul A. Jamari (*Paper 13*)

(3) Ms Han Hui Hui (*Paper 15*)

(4) Prof Hany Farid (*Paper 17*)

(5) Mr Prakash Kumar Hetamsaria (*Paper 18*)

(6) Mr Zubin Jain (*Paper 22*)

(7) Ms Bertha Henson (*Paper 26*)

(8) Mr Shriniwas Rai (*Paper 27*)

(9) Ms Gaurav Keerthi (*Paper 28*)

(10) Mr Teymoor Nabili (*Paper 31*)

(11) Mr Darius Lee (*Paper 32*)

(12) NTUC FairPrice Co-operative Ltd (*Paper 33*)

(13) European Values Think-Tank (*Paper 34*)

(14) Mr Ben Nimmo (*Paper 36*)

(15) Assoc Prof Alton Chua (*Paper 38*)

(16) Channel NewsAsia (*Paper 39*)

(17) National Library Board (*Paper 40*)

(18) Assoc Prof Ullrich Ecker (*Paper 44*)

(19) Assoc Prof Liew Kai Khiun (*Paper 46*)

(20) Prof Cherian George (*Paper 47*)

(21) Ms Kirsten Han (*Paper 48*)

(22) Titular Roman Catholic Archbishop of Singapore (*Paper 49*)

(23) Ms Rachel Er Shengtian and Joel Jaryn Yap Shen (*Paper 51*)

(24) Ukraine Crisis Media Centre (*Paper 54*)

(25) Dr Thio Li-ann (*Paper 55*)

(26) Ms Gulizar Haciyakupoglu (*Paper 56*)

(27) Mr Shashi Jayakumar (*Paper 59*)

(28) Ms Danielle Chee, Mr Darren Kang, Ms Felicia Chu, Ms Noor Syazana Bte Rafeeq Ahamed, Ms Jacelyn Loh, Ms Jelisa Tan, and Mr Zheng Liren (*Paper 60*)

(29) Masyarakat Anti-Fitnah Indonesia (Mafindo) (*Paper 61*)

(30) Dr Carol Soon Wan Ting and Mr Shawn Goh Ze Song (*Paper 62*)

(31) Mr Norman Vasu (*Paper 63*)

(32) Community Action Network (*Paper 72*)

(33) Castex Chair of CyberStrategy (*Paper 73*)

(34) Prof Simon Hegelich and Mr Morteza Shahrezaye (*Paper 74*)

(35) Mr Dan Shefet (*Paper 75*)

(36) Dr Janis Berzins (*Paper 77*)

(37) StopFake.org (*Paper 78*)

(38) Mr Zhulkarnain Abdul Rahim (*Paper 80*)

(39) Ms Jennifer Yang Hui (*Paper 82*)

(40) Dr Thum Ping Tjin (*Paper 83*)

(41) Mr Benjamin Joshua Ong (*Paper 84*)

(42) Trend Micro Inc (*Paper 86*)

(43) Mr Rajesh Sreenivasan (*Paper 87*)

(44) Ms Claire Wardle (*Paper 94*)

(45) Prof Kalina Bontcheva (*Paper 96*)

(46) Asst Prof Michael Raska (*Paper 97*)

(47) Mr Raja Mohan M K (*Paper 98*)

(48) Mr Mathew Mathews (*Paper 100*)

(49) Ms Simran Kaur Sandhu, Ms Gloria Chan Hui En, Mr Daryl Gan and Ms Cheah You Yuan (*Paper 101*)

(50) Mr Damien Cheong (*Paper 103*)

(51) Facebook (*Paper 104*)

(52) PAP Policy Forum (*Paper 107*)

(e) that the following persons or organisation be invited to submit a written representation by the closing date of 7 March 2018 and to give oral evidence:

(1) Mr Andrew Loh

(2) Mr Terry Xu

(3) Human Rights Watch

(f) that accredited local and foreign media be admitted to public hearings for the purposes of recording, broadcasting and reporting the proceedings;

(g) that members of the public be admitted to observe public hearings; and

(h) that a press statement be issued.

Adjourned till 10.00 am on
Friday, 9 March 2018

3rd Meeting

Friday, 9 March 2018

10.00 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Dr Janil Puthucheary
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

ABSENT

Ms Chia Yong Yong
Mr Desmond Lee

In Attendance:

Ministry of Law:
Ms Lim Hui Min, Delphia, Senior Assistant Director, International Legal Division

Ministry of Communications and Information:
Ms Angela Tan, Assistant Director, Information Policy Division

1. The Committee deliberated.

2. Written representations received were considered.

3. *Agreed* –

   (a) that the five written representations received late be accepted for consideration;

   (b) that Papers 103A and 108 to 164 be published.

   (c) that the following representors be invited to give oral evidence:

      (1) Prof Lim Sun Sun (*Paper 101*)

      (2) MARUAH (*Paper 112*)

      (3) NGO Monitor (*Paper 117*)

      (4) Asia Internet Coalition (*Paper 119*)

(5) Singtel (*Paper 121*)

(6) Mr Thiruprakassh S/O Suppiah (*Paper 122*)

(7) National Council of Churches of Singapore (*Paper 124*)

(8) Mediacorp Pte Ltd (*Paper 125*)

(9) StarHub Ltd (*Paper 126*)

(10) Dr Goh Yihan (*Paper 129*)

(11) Mr Sui Yi Siong, Mr Choo Hao Ren Lyndon, Ms Chen Lixin and Mr Aaron Yoong Joon Wei (*Paper 130*)

(12) Mr Benjamin Ang *(Paper 135)*

(13) Internet Society Singapore Chapter (*Paper 136*)

(14) Mr Andrew Loh (*Paper 137*)

(15) Google *(Paper 138)*

(16) Mr Nicholas Fang (*Paper 144*)

(17) Singapore Press Holdings (*Paper 148*)

(18) Asst Prof Elmie Nekmat (*Paper 149*)

(19) Assoc Prof Eugene Tan *(Paper 150)*

(20) Dr Gillian Koh *(Paper 152)*

(21) Twitter Inc *(Paper 153)*

(22) The Online Citizen *(Paper 154)*

(23) Singapore Press Club and Singapore Corporate Counsel Association *(Paper 155)*

(24) Roses of Peace (*Paper 158)*

(25) Mothership.sg (*Paper 159)*

(26) Dr Kweh Soon Han (*Paper 160*); and

(d) that a press statement be issued.

4. It was further agreed that a security agency be heard in private at a date to be fixed.

Adjourned till 11.00 am on
Wednesday, 14 March 2018

4th Meeting

Wednesday, 14 March 2018

11.00 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

ABSENT

Mr Desmond Lee

1. The Committee deliberated.

2. *Agreed* that Ms Gulizar Haciyakupoglu (*Paper 56*) and Mr Damian Cheong (*Papers 103 and 103A*) be heard in private.

3. The following witnesses were examined under oath or affirmation:

   (a) Dr Carol Soon Wan Ting and Mr Shawn Goh Ze Song *(Paper 62)*;

   (b) Mr Mathew Mathews *(Paper 100);*

   (c) Asst Prof Michael Raska *(Paper 97);*

   (d) Mr André Ahchak of the Roman Catholic Archdiocese *(Paper 49)*, Rev Dr Ngoei Foong Nghian and Dr Roland Chia Cheng Kim of the National Council of Churches of Singapore (*Paper124*), and Dr Kweh Soon Han of the Singapore Buddhist Federation *(Paper 160)*;

   (e) Mr Shriniwas Rai (*Paper 27*); and

   (f) Dr Goh Yihan (*Paper 129*).

Adjourned till 10.00 am on
Thursday, 15 March 2018

5th Meeting

Thursday, 15 March 2018

10.00 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

ABSENT

Dr Janil Puthucheary
Mr Desmond Lee

1. The Committee deliberated.

2. The following witnesses were examined under oath or affirmation:

   (a) Dr Shashi Jayakumar (*Paper 59*);

   (b) Mr Ruslan Deynychenko of StopFake.org (*Paper 78*);

   (c) Mr Jakub Janda of European Values Think-Tank (*Paper 34*) via video-conference;

   (d) Dr Janis Berzins (*Paper 77*) via video-conference;

   (e) Ms Nataliia Popovych and Mr Oleksiy Makhuhin of Ukraine Crisis Media Center (*Paper 54*) via video-conference;

   (f) Assoc Prof Kevin Limonier of Castex Chair of CyberStrategy (*Paper 73*) via video-conference; and

   (g) Mr Ben Nimmo (*Paper 36*) via video-conference.

Adjourned till 10.00 am on
Friday, 16 March 2018

6th Meeting

Friday, 16 March 2018

10.00 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

ABSENT

Dr Janil Puthucheary
Mr Desmond Lee

---

1. The Committee deliberated.

2. The following witnesses were examined under oath or affirmation:

   (a) Ms Gulizar Haciyakupoglu (*Paper 56*) in private;

   (b) Mr Damien Chong (*Papers 103 and 103A*) in private;

   (c) Mr Septiaji Eko Nugroho of Masyarakat Anti-Fitnah Indonesia (Mafindo) (*Paper 61*)

   (d) Asst Prof Elmie Nekmat (*Paper 149*);

   (e) Ms Myla V. Pilao of Trend Micro Inc. (*Paper 86*); and

   (f) Mr Morteza Shahrezaye (*Paper 74*)

Adjourned till 8.30 am on
Tuesday, 20 March 2018

7th Meeting

Tuesday, 20 March 2018

8.30 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Desmond Lee
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Ms Sun Xueling
Mr Edwin Tong Chun Fai

ABSENT

Mr K Shanmugam

---

1. The Committee deliberated.

2. *Agreed* -

   (a) that Reporters Without Borders (also known as Reporters Sans Frontieres or RSF) be invited to give oral evidence;

   (b) that the witness list be revised; and

   (c) that the written representations of Ms Gulizar Haciyakupoglu (*Paper 56*) and Mr Damien Cheong (*Papers 103 and 103A*) who were heard in private be not published.

3. The Committee heard evidence in private from a security agency.

Adjourned till 10.00 am on
Thursday, 22 March 2018

8th Meeting

Thursday, 22 March 2018

11.00 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Desmond Lee
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

ABSENT

Mr Pritam Singh

1. The Committee deliberated.

2. The following witnesses were examined under oath or affirmation:

   (a) Ms Jennifer Yang Hui (*Paper 82*);

   (b) Mr Zubin Jain (*Paper 22*);

   (c) Mr Simon Milner and Mr Alvin Tan of Facebook (*Paper 104*), Mr Jeff Paine of Asia Internet Coalition (*Paper 119*), Ms Irene Jay Liu of Google (*Paper 138*) and Ms Kathleen Mary Helen Reen and Mr Philip Chua Jin Wen of Twitter Inc (*Paper 153*); and

   (d) Mr Yuen Kuan Moon and Mr Slattery Sean Patrick of Singtel *(Paper 121)* and Mr Tim Goodman of StarHub Ltd *(Paper 126)*.

Adjourned till 10 am on
Friday, 23 March 2018

9th Meeting

Friday, 23 March 2018

9.30 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Desmond Lee
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

ABSENT

Mr Pritam Singh

---

1. The Committee deliberated.

2. *Agreed* -

   (a) that the six written representations received late be accepted for consideration;

   (b) that Papers 165 to 170 be published; and

   (c) that a press statement on the Select Committee's correspondence with Human Rights Watch be issued.

3. The following witnesses were examined under oath or affirmation:

   (a) Mr Gaurav Keerthi (*Paper 28*);

   (b) Dr Thio Li-ann (*Paper 55*);

   (c) Mr Walter Fernandez and Mr Jaime Ho of Channel NewsAsia (*Paper 39*), and Mr Warren Fernandez, Mr Goh Sin Teck and Mr Mohamed Sa'at bin Abdul Rahman of Singapore Press Holdings (*Paper 148*);

   (d) Mr Lien We King and Mr Martino Tan of Mothership.sg (*Paper 159*);

   (e) Mr Vikram Nair, Mr Benjamin Tay Yong Guan, Mr Jude Tan Kim Chooi and Mr Sujatha Selvakumar of the PAP Policy Forum (*Paper 107*);

   (f) Prof Gerald M Steinburg of NGO Monitor (*Paper 117*);

(g) Mr Poh Leong Sim, Mr Jonas Kor and Ms Chong Nyet Chin of NTUC Fairprice Co-operative (*Paper 33*);

(h) Mr Wong Taur-Jiun and Ms Angeline Lee of the Singapore Corporate Counsel Association; Mr Patrick Daniel, Mr Zakir Hussain and Ms Lau Joon Nei of Singapore Press Club and Dr Stanley Lai of Allen & Gledhill LLP (*Paper 155*); and

(i) Dr Gillian Koh (*Paper 152*).

Adjourned till 10.00 am on
Tuesday, 27 March 2018

10th Meeting

Tuesday, 27 March 2018

9.50 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Desmond Lee
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

---

1. The Committee deliberated.

2. *Agreed* -

   (a) that a press statement on the scheduling of witnesses be issued;

   (b) that a press statement on the Select Committee's correspondence with Reporters Without Borders (also known as Reporters San Frontieres or RSF) be issued; and

   (c) that the witness list be further revised.

3. The following witnesses were examined on oath or affirmation:

   (a) Prof Hany Farid (*Paper 17*) via video-conference;

   (b) Mr Benjamin Ang (*Paper 135*);

   (c) Mr Hazrul A Jamari (*Paper 13*), Mr Zulkarnain Abdul Rahim (*Paper 34*) and Mr Abbas Ali Mohamed Irshad, Mr Jonathan Tan Bingxian and Mr Nadim Kapadia of Roses of Peace (*Paper 158*);

   (d) Ms Ng Wai Yin, Mr Chow Wun Han, and Ms Sara Pek Leng Leng of National Library Board (*Paper 40*);

   (e) Prof Cherian George (*Paper 47*);

   (f) Mr Howard Lee (*Paper 12 and 12A*) via video-conference, Ms Kirstan Han (*Paper 48*), Mr Ngiam Shih Tung of MARUAH (*Paper 112*), and Mr Terry Xu of The Online Citizen (*Paper 154*); and

(g) Mr Jolovan Wham of Community Action Network (*Paper 72*).

4. The Committee further deliberated.

5. *Agreed* that a press statement reiterating the invitation to Human Rights Watch to give evidence before the Committee be issued.

Adjourned till 10.00 am on
Wednesday, 28 March 2018

11th Meeting

Wednesday, 28 March 2018

10.00 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Desmond Lee
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

---

1. The Committee deliberated.

2. The following witnesses were examined under oath or affirmation:

   (a) Assoc Prof Alton Chua (*Paper 38*), and Asst Prof Liew Kai Khiun (*Paper 46*);

   (b) Mr Prakash Kumar Hetamsaria (Paper 18) and Mr Raja Mohan M K (Paper 98);

   (c) Mr Dan Shefet (*Paper 75*);

   (d) Assoc Prof Eugene Tan (*Paper 150*);

   (e) Dr Norman Vasu (*Paper 63*); and

   (f) Mr Andrew Loh (*Paper 137*).

Adjourned till 10.00 am on
Thursday, 29 March 2018

12th Meeting

Thursday, 29 March 2018

10.00 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Desmond Lee
Mr Pritam Singh
Ms Rahayu Mahzam
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

ABSENT

Mr Seah Kian Peng

1. The Committee deliberated.

2. *Agreed* –

   (a) that two late representations received by the Committee be not considered; and

   (b) that a press release on the completion of public hearings be issued.

3. The following witnesses were examined under oath or affirmation:

   (a) Ms Rachel Er Shengtian and Mr Joel Jaryn Yap Shen (*Paper 51*), Ms Simran Kaur Sandhu and Ms Gloria Chan Hui En (*Paper 101*), and Ms Sui Yi Siong and Mr Chen Lixin (*Paper 130)*;

   (b) Dr Thum Ping Tjin (*Paper 83*);

   (c) Mr Nicholas Fang (*Paper 144*); and

   (d) Prof Lim Sun Sun (*Paper 110*).

4. The Committee further deliberated.

5. *Agreed* that a further press statement on the Select Committee's invitation to Human Rights Watch be issued.

Adjourned *sine die*

13th Meeting

Friday, 17 August 2018

4.30 pm

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Desmond Lee
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr K Shanmugam
Ms Sun Xueling
Mr Edwin Tong Chun Fai

---

1. The Committee deliberated.

2. *Agreed* –

   (a) that Dr Thum Ping Tjin's additional representation of 4 May 2018 be published in the Select Committee's Report;

   (b) that, in view of the full videos of all public hearings made available on the Parliament website, the summary of oral evidence for 29 March 2018 be not published on the Parliament website; and

   (c) that the summaries of evidence of Dr Gulizar Haciyakupoglu and Dr Damien Cheong be published in the Select Committee's Report.

3. The Committee further deliberated.

Adjourned *sine die*

14th Meeting

Monday, 3 September 2018

10.30 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Desmond Lee
Mr Pritam Singh
Mr K Shanmugam
Ms Sun Xueling

ABSENT

Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr Edwin Tong Chun Fai

1. The Committee deliberated.

Adjourned till 10.00 am on
Tuesday, 11 September 2018

15th Meeting

Tuesday, 11 September 2018

10.00 am

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Desmond Lee
Mr Pritam Singh
Ms Rahayu Mahzam
Mr K Shanmugam
Mr Seah Kian Peng
Ms Sun Xueling

ABSENT

Mr Edwin Tong Chun Fai

1. The Committee deliberated.

Report

2. The Chairman's report brought up and read the first time.

3. Resolved, "That the Chairman's report be read a second time paragraph by paragraph.".

   Paragraphs 1 to 585 inclusive read and agreed to.

   Annexes A to G inclusive of the Chairman's report read and agreed to.

4. Resolved, "That this report be the Report of the Committee to Parliament.".

5. Agreed that the Chairman do present the Report to Parliament on Thursday 20 September 2018.

Adjourned *sine die*

16th Meeting

Tuesday, 18 September 2018

5.30 pm

PRESENT

Deputy Speaker Charles Chong *(in the Chair)*
Ms Chia Yong Yong
Dr Janil Puthucheary
Mr Pritam Singh
Ms Rahayu Mahzam
Mr Seah Kian Peng
Mr Edwin Tong Chun Fai

ABSENT

Mr Desmond Lee
Mr K Shanmugam
Ms Sun Xueling

1. Agreed that the Chairman do present the Select Committee's Report to Parliament on Wednesday 19 September 2018.

Adjourned *sine die*

Appendix II

LIST OF INDIVIDUALS AND ORGANISATIONS FROM WHOM WRITTEN REPRESENTATIONS WERE RECEIVED BY THE SELECT COMMITTEE

Paper No.	Representor
1	Ong Junkai (Self-employed)
2	Yu Qinxu (Management Consultant)
3	Gan Siok Bin (Retired accountant)
4	Ler Han Qiang (Engineer)
5	Rongxiang Lin (Self-employed Computer Engineer)
6	COL (Ret) K. Kuharajahsingam (Counsellor)
7	Erwin
8	Dr Rex Yeap (Lecturer)
9	Dr Lee Hock Seng (Private Family Physician)
10	Chandra Das (Property Agent)
11	Yvonne Wong (Unemployed)
12 12A	Howard Lee (PhD Student)
13	Hazrul A. Jamari (Entrepreneur)
14	Ang Chin Chye (Lawyer)
15	Han Hui Hui (Blogger)
16	Yeo Boon Eng (Tutor)
17	Hany Farid (Professor and Chair, Computer Science, Dartmouth College)
18	Prakash Kumar Hetamsaria (CFO)
19	Toh Hwee Boon (Freelance Counsellor)
20	Anonymous

Paper No.	Representor
21	Raymond Khng Guan Gek (Unemployed)
22	Zubin Jain (Student)
23	Edwin Ho (Self-trading in financial markets)
24	Alex Tan
25	Nga Thio Ping (Retiree)
26	Bertha Henson (Adjunct Professor and part-time blogger)
27	Shiriniwas Rai (Lawyer)
28	Gaurav Keerthi (Founder of dialectic.sg and confirm.sg)
29	Kevin Seah (Private tutor, freelance editor/writer)
30	Wilson Na (Software Engineer)
31	Teymoor Nabili (Freelance Journalist, Host of the "Perspectives" current affairs show on Channel NewsAsia)
32	Darius Lee (Advocate and Solicitor)
33	NTUC FairPrice Co-operative Ltd
34	Jakub Janda (Head, Krelim Watch Program; Director, European Values Think-Tank)
35	Edmund Chow (Postdoc Research Fellow) Mohamad Abdillah Zamzuri (Director (Arts & Education)) Nadine Yap (Chief Customer Success Officer) Osman Sulaiman (Business Owner) Ravi Chandran Philemon (Executive Director) Wendy Koh Lai May (Educationist)

Paper No.	Representor
36	Ben Nimmo (Senior Fellow, Information Defense Digital Forensic Research Lab)
37	Matthew Soo Yee (Awaiting matriculation)
38	Assoc Prof Alton Chua (Associate Professor, and Associate Chair (Research), Wee Kim Wee School of Communication and Information, Nanyang Technological University)
39	Senior Editors of Channel NewsAsia
40	National Library Board
41	The Independent
42	Nicolas Arpagian (Director of Strategy, Orange Cyberdefense. Scientific Director; Cybersecurity Program, National Institute for Security & Judicial Studies (INHESJ – French Prime Minister Office))
43	Singapore Philosophy Group
44	Associate Professor Ullrich Ecker (Associate Professor Director, Community and Engagement, School of Psychological Science, University of Western Australia)
45	Anthony Chia (Business Consultant)
46	Assistant Professor Liew Kai Khiun (Assistant Professor, Wee Kim Wee School of Communication and Information, Nanyang Technological University)
47	Professor Cherian George (Professor of Media Studies, School of Communication, Hong Kong Baptist University)
48	Kirsten Han (Journalist and Writer)
49	Roman Catholic Archdiocese
50	Jev Akshay s/o Jeevan (Student and Writer)
51	Er Shengtian, Rachel Joel Jaryn Yap Shen (Law undergraduates, National University of Singapore)
52 52A	Calvin Cheng Ern Lee (Entrepreneur)
53	Chong Ja Ian (Teacher)

Paper No.	Representor
54	Nataliia Popovych (Co-Founder, Board Member, Ukraine Crisis Media Center) Oleksiy Makhuhin (Head of Hybrid Warfare Analytical Group of Ukraine Crisis Media Center)
55	Professor Thio Li-ann (Professor of Law, National University of Singapore)
56	Gulizar Haciyakupoglu (Research Fellow, Centre of Excellence for National Security (CENS))
57	Function 8
58	AWARE
59	Shashi Jayakumar (Head, Centre of Excellence for National Security and Executive Coordinator, Future Issues and Technology, S. Rajaratnam School of International Studies, Nanyang Technological University, Singapore)
60	Chee Muk Onn Danielle Kang Darren Chu Jian Ren Felicia Noor Syazana Bte Rafeeq Ahamed Jacelyn Loh Liang Nee Tan Ler Min, Jelisa (Students) Zheng Liren (Lecturer)
61	Septiaji Eko Nugroho (Founder, Mafindo/Indonesian Anti Hoax Community)
62	Soon Wan Ting, Carol (Senior Research Fellow (Institute of Policy Studies, Lee Kuan Yew School of Public Policy, National University of Singapore)) Shawn Goh Ze Song (Research Assistant (Institute of Policy Studies, Lee Kuan Yew School of Public Policy, National University of Singapore))
63	Norman Vasu (Senior Fellow, Centre of Excellence for National Security, S. Rajaratnam School of International Studies)
64	Seah Ming Yan Bertrand (Student)
65	Cedric Choo (Undergraduate Student, Yale-NUS College)
66	Ronald Chan (Semi-retired)

Paper No.	Representor
67	Ng Kok Hua (Retired)
68	Julian Sng Wei Meng (Tertiary Student)
69	Alan Soon (On behalf of roundtable discussion between Singapore Press Holdings, the National University of Singapore, the Singapore Management University, the Media Literacy Council, Mothership, the Asia Internet Coalition, Twitter, Facebook, Rajah & Tann and others)
70	Valerie
71	Datos Concepción
72	Community Action Network (CAN)
73	Kevin Limonier (Associate Researcher, Castex Chair of Cyberstrategy)
74	Simon Hegelich (Professor for political data science at the Bavarian School of Public Policy, Technical University of Munich) Morteza Shahrezaye (Researcher at the professorship for political data science at the Bavarian School of Public Policy, Technical University of Munich)
75	Dan Shefet (Individual Specialist to UNESCO and French lawyer specialized in European Law and IT Law)
76	Koh Jee Leong (Singapore Unbound)
77	Dr. Jānis Bērziņš (Director, Center for Security and Strategic Studies, The National Defense Academy of Latvia)
78	StopFake.org
79	Liew Siow Gian Patrick (Business Owner)
80	Zhulkarnain Abdul Rahim (Lawyer)
81	Isaac Neo Yi Chong (Student)
82	Jennifer Yang Hui (Associate Research Fellow, Centre of Excellence for National Security)
83	Dr Thum Ping Tjin (Historian and also the founder, Managing Director, and Research Director of New Naratif; Research Fellow in History and Coordinator of Project Southeast Asia at the University of Oxford (2014-present))

Paper No.	Representor
84	Benjamin Joshua Ong (Lecturer of Law (FDS) School of Law, Singapore Management University)
85	Chui Jian Wei (Civil Servant)
86	Trend Micro Inc
87	Rajesh Sreenivasan (Partner of Rajah & Tann Singapore LLP and Head of its Technology, Media & Telecommunications practice)
88	Lao Yuen Seong (Sales)
89	Liu Ching Man (Manager)
90	Chua Jiawen (Strategy Manager)
91	Associate Professor Alan Chong (Associate Professor Centre for Multilateralism Studies, Institute of Defence and Strategic Studies)
92	Jiang Haolie (University student)
93	Cheng Zai Hui
94	Claire Wardle (Research Fellow at the Shorenstein Center for Media, Politics and Public Policy, Harvard Kennedy School and Executive Director of First Draft)
95	Sin Kin Kok (Retired)
96	Professor Kalina Bontcheva (Professor of Text Analytics, University of Sheffield)
97	Assistant Professor Michael Raska (Assistant Professor, S. Rajaratnam School of International Studies, Nanyang Technological University)
98	Raja Mohan M K (Chief Programme Officer)
99	Sabaratnam Ratnakumar (Retiree)
100	Mathew Mathews (Senior Research Fellow, Institute of Policy Studies, Lee Kuan Yew School of Public Policy, National University of Singapore)
101	Simran Kaur Sandhu Gloria Chan Hui En Daryl Gan Cheah You Yuan (Singapore Management University (2nd Year LLB students))
102	Kriel.Agency

Paper No.	Representor
103 103A	Damien D. Cheong (Research Fellow, National Security Studies Programme, S. Rajaratnam School of International Studies, Nanyang Technological University)
104	Facebook
105	Benjamin Chen (Student of Nanyang Technological University)
106	Chua Jun Hao (Nanyang Technological University, Year 3 Accountancy Student)
107	PAP Policy Forum (PPF)
108	Kwek Suat Yee (Doctor)
109	Siew Yaw Hoong (Engineer)
110	Professor Lim Sun Sun (Professor of Media and Communications, and Head of Humanities, Arts and Social Sciences at the Singapore University of Technology and Design)
111	Anonymous
112	MARUAH (Working Group for an ASEAN Human Rights Mechanism, Singapore)
113	Yam Yi Jie
114	Qsearch
115	Tisane Labs Pte Ltd
116	Yeo Chee Hian (Engineer)
117	Prof Gerald M Steinberg (President, NGO Monitor)
118	Stephen Lim
119	Asia Internet Coalition (AIC)
120	Sudhir Thomas Vadaketh (Writer)
121	Singtel
122	Thiruprakassh s/o Suppiah (Manufacturing Manager)
123	Adrian Kwek (Senior Lecturer, Singapore University of Social Sciences (SUSS))
124	The National Council of Churches of Singapore (NCCS)

Paper No.	Representor
125	Mediacorp Pte Ltd
126	StarHub Ltd
127	Lim Puay Kuan (PA Trainer)
128	Ngoh Wang Long (Application Support)
129	Goh Yihan (Dean of the School of Law, Singapore Management University)
130	Sui Yi Siong (Lawyer) Choo Hao Ren, Lyndon Chen Lixin Aaron Yoong Joon Wei (Undergraduates)
131	Kwok Siang (Guo Xiang) (Student)
132	Cheah Wenjie Chester Su Yong Meng (3rd Year Undergraduates Full-Time, National University of Singapore)
133	Timothy Tan
134	Lim Sheng Kang Shaun (Fourth-year student, Faculty of Law, National University of Singapore)
135	Benjamin Ang (Senior Fellow / Coordinator Cyber and Homeland Defence, Centre of Excellence for National Security (CENS), S. Rajaratnam School of International Studies (RSIS), Nanyang Technological University)
136	Internet Society Singapore Chapter (ISOCSG)
137	Andrew Loh Hong Puey (Self-investor)
138	Google
139	Andrew Fung (Career Coach)
140	Teo Geok Choo (Housewife)
141	Chong Huat Kwong (Jeffery) (Senior Operations Manager)
142	Chen Sicong Tay Wei Jie, Joel (Undergraduate law students, Singapore Management University)
143	Ang Peng Hwa (Lecturer, Wee Kim Wee School of Communication and Information, Nanyang Technological University)

Paper No.	Representor
144	Nicholas Fang (Managing Director, Black Dot Pte Ltd)
145	Alan Ting Yee Chong (Director)
146	Chong Zi Liang
147	Carlos Nicholas Fernandes (Technology Entrepreneur and retired member of the MTI's Pro-Enterprise Panel)
148	Singapore Press Holdings (SPH)
149	Assistant Professor Mohamed Elmie Bin Nekmat (Assistant Professor, Communications and New Media, National University of Singapore)
150	Associate Professor Eugene Tan (Associate Professor of Law, School of Law, Singapore Management University)
151	Transient Workers Count Too (TWC2) and Humanitarian Organization for Migrant Economics (HOME)
152	Dr Gillian Koh (Deputy Director (Research) Institute of Policy Studies, Lee Kuan Yew School of Public Policy, National University of Singapore)
153	Twitter Inc
154	The Online Citizen (TOC)
155	Singapore Press Club and Singapore Corporate Counsel Association
156	Lim Shi Mei Benjamin Yiwen Smith (Postgraduate Students)
157	Dr Shobha Avadhani (Instructor, Centre for English Language Communication, National University of Singapore)
158	Roses of Peace
159	Mothership.sg
160	Singapore Buddhist Federation

Paper No.	Representor
161	Joses Ho (Research Fellow) Tan Jian Xiong David (Research Assistant) Ervin Tan (Lawyer) Gwyneth Teo (Journalist)
162	Jonathan Lim (Lawyer)
163	Roy Fung (Managing Director)
164	Tan Keng Sooi (Retiree)
165	Lim Boon Tiong Terence (Craftsman)
166	Media Literacy Council
167	Benjamin Goh
168	Anonymous
169	Embassy of the Russian Federation in Singapore
170	Iwan Rahabok (IT Architect)