What Lies Beneath the Truth:

A Literature Review on Fake News, False Information and More

Carol Soon Wan Ting, Senior Research Fellow

Shawn Goh Ze Song, Research Assistant

30 June 2017

iPS Institute of Policy Studies

CONTENTS

1

EXECUTIVE SUMMARY

The rise of the Internet and social media has resulted in a rapidly evolving information ecosystem, where information exchange occurs in unprecedented scale and speed. While these new dynamics offer exciting opportunities, they also introduce new challenges to governments and societies. Recent political events have thrown "fake news" into the spotlight and have created the impression that fake news is a new phenomenon and problem. However, fake news is but one type of false information in our information ecology. Other types of false information such as rumours, conspiracy theories, propaganda and disinformation have existed since the pre-Internet days, and have always been used by different groups and individuals to influence public opinion.

Section 2: Defining Fake News begins with addressing the potential difficulties when *fake news* is used as an umbrella term to refer to many different types of false information. Without a concise and consistent definition of the term, it is difficult to prescribe appropriate counter-measures to effectively address the problem. Experts generally agree that the term *fake news* refers to a very specific type of false information – deliberately fabricated false content that assumes the appearance of an authoritative information source, created with the intention to deceive, and often for political and(or) economic gains. This specific definition of *fake news* distinguishes it from other types of false information. Nonetheless, existing research on the various types of false information offer insights as to why fake news is a growing problem today.

With the abundance of information in our society today, people tend to rely on heuristics and social cues to evaluate the information that they encounter. However, this often leads them to succumb to their cognitive biases. *Section 3: The Persistence of False Information (Human Factors)* looks at the individual-cognitive and socio-

psychological factors that may increase people's likelihood of believing false information. For example, *confirmation bias* and *motivated reasoning* explain why people tend to embrace information compatible with their preexisting beliefs and reject information that contradict them, regardless of the accuracy of the information. Furthermore, correcting people's false beliefs may even backfire and cause them to hold onto those false beliefs even more strongly, especially when the correction contradicts their worldview. Other factors like negative emotions and social pressure also contribute to people's susceptibility to believing false information.

While the problem of false information is not a new one, the emergence of the Internet and social media has certainly presented novel challenges in terms of dealing with it. *Section 4: The Persistence of False Information (Media Factors)* looks at how the unique characteristics of the Internet and social media exacerbate the negative effects of people's cognitive biases. For example, filter bubbles and echo chambers created by the algorithms used by social media platforms tend to reinforce people's worldviews and entrench their existing biases. Social bots can also amplify the propagation of false information, thus increasing people's exposure to it.

Section 5: Consumers and Propagators of False Information looks at whether certain groups of people are particularly susceptible to believing false information. For example, research has shown that people who hold extreme political beliefs, and those who perceive power loss or social exclusion may be more prone to believing false information. This section also looks at examples of how governments, organisations and individuals with vested interests, and Internet subculture groups contribute to the production and propagation of false information.

It is evident that the problem of false information is a complex one. Thus, corresponding solutions to combat false information have to be multi-pronged. In *Section 6: Solutions and Policies*, we look at various existing and potential solutions, as well as specific government responses to the problem of false information. Fact checking serves as an essential near-term measure to debunk false information.

However, that alone is usually insufficient as fact checking efforts are often performed after a piece of false information has been spread widely, when the damage is already done. Furthermore, attempts at correcting false beliefs often backfire as previously mentioned. Fortunately, research shows promising results in successfully "vaccinating" or "inoculating" people psychologically, thus conferring people with a resistance against false information even prior to exposure to it. Other more medium- and long-term solutions that may address the problem of false information include leveraging the help of technology and machines, re-building quality and trust in journalism, and improving news literacy among the wider public.

Many governments around the world have looked into the possibility of regulating social media platforms to curb the spread of false information. However, experts generally agree that regulatory measures may be difficult to implement and may also be limited in their effectiveness. Furthermore, regulation may infringe upon freedom of expression, resulting in censorship and a chilling effect on public discussion. Instead, non-regulatory measures such as improving news literacy and trust in quality journalism are viable alternative solutions.

We conclude our literature review in *Section 7: Conclusion* by proposing potential research areas that may be relevant and useful to Singapore's context.

2

DEFINING *FAKE NEWS*

❖ The term *fake news* is currently used as an umbrella term to refer to many different types of false information. Without a concise and consistent definition, it will be difficult to prescribe appropriate counter-measures.

❖ *Misinformation* is the unintentional sharing of false information, whereas *disinformation* is the deliberate creation and non-accidental sharing of false information.

❖ Industry players and academics agree that the term *fake news* refers to a very specific type of false information – 1) driven by intention to deceive, 2) politically and/or economically motivated, 3) deliberate fabrication of false content, and 4) assumes the form and appearance of authoritative news sources.

Currently, the term *fake news* is used to refer to a wide variety of misleading or false information online. They range from rumours circulated via Instant Messaging (e.g. salt shortage in India), posts and updates on social media (e.g. free HPV vaccine programme as part of a Chinese conspiracy in Indonesia), inaccurate reports and commentaries on websites (e.g. a commentary published on *The Real Singapore* on a Filipino family's involvement in a commotion between Thaipusam participants and the police), sensational and partisan news stories (e.g. those published by *USConservativeToday.com* and *LibertyNewsWriters.com*), to fake news websites (e.g. *CNN-Alive.com*, *Metro-UK.com* and those by the Macedonian youths).

There is much ambiguity concerning the precise distinctions between *fake news* on the one hand, and ideologically slanted news, disinformation, misinformation, and propaganda on the other (Mele *et al.*, 2017). *Fake news* has become a catch-all phrase to refer to everything from news articles that are factually incorrect to rumours, Internet hoaxes, conspiracy theories, and political parodies. In some cases, overly broad usage of the term has bordered on abuse and exploitation, particularly when it is used to label information that is true but disliked by an individual or a group. One example would be US President Donald Trump who labels news reports published by mainstream media which he perceives to be biased against him as fake.

The broad use of the term *fake news* is problematic on three levels. First, without a clear definition, it is difficult to prescribe the most appropriate counter-measure — what types of false information to target, what and whom to monitor, whom to act against, what counter-measures to adopt, and how to communicate why action is needed.

Second, misuse of the term can have a negative impact on traditional institutions such as journalism. In the long term, it erodes the public trust in journalism, which in turn has implications on it, and on how and where people access reliable information needed for informed decision-making. Third, from the perspective of policy communication, contestation of the term may fuel suspicions on whether criticisms and measures taken against fake news may be politically-motivated (Marwick & Lewis, 2017).

This section will present the information ecology, of which fake news is part of. Following which, the section will focus on how *fake news* is defined by the industry and academia.

2.1

Misinformation and Disinformation

❖ *Misinformation* is the unintentional sharing of false or inaccurate information.

❖ *Disinformation* refers to the deliberate creation and sharing of information known to be false, typically driven by intentions to arouse passions, attract viewership, or deceive.

Misinformation is the unintentional sharing of false or inaccurate information, while *disinformation* refers to the "deliberate creation and sharing of information known to be false" (Mele *et al.*, 2017). While malicious intent is absent in misinformation, disinformation is typically driven by intentions "to arouse passions, attract viewership, or deceive". Perpetrators of disinformation often use subtle methods, such as feeding inaccurate quotes or stories to innocent intermediaries, or knowingly amplifying biased or misleading information (Weedon, Nuland & Stamos, 2017).

Fallis (2015) emphasised two characteristics of disinformation. First, it is misleading, meaning that the information is likely to create false beliefs, thereby subjecting people to harm, epistemic or otherwise. Second, it is "non-accidentally" misleading:

"It is this feature [non-accidental] that distinguishes disinformation from more innocuous forms of misleading information, such as honest mistakes and overly subtle satire. Focusing on non-accidentally misleading information puts us in a better position to detect this specific type of misleading information; for instance, much like lie-detection techniques, we can look for the intention to mislead instead of just looking for errors."

The information ecology encompassing both misinformation and disinformation is represented in a typology developed by First Draft, a non-profit coalition formed in 2015 to address challenges relating to trust and truth in the digital age. The typology classifies misinformation and disinformation based on: 1) the content that is created and shared, 2) the motivations (intent) of the content creators, and 3) the ways the content is disseminated (see *Figure 1*). These seven types occupy a scale which measures the intent to deceive (Wardle, 2017).

The motivations behind the seven types of misinformation and disinformation are:

1.   Poor journalism

2.   Parody

3.   To provoke or 'punk'

4.   Passion and partisanship

5.   Profit

6.   Political influence or power
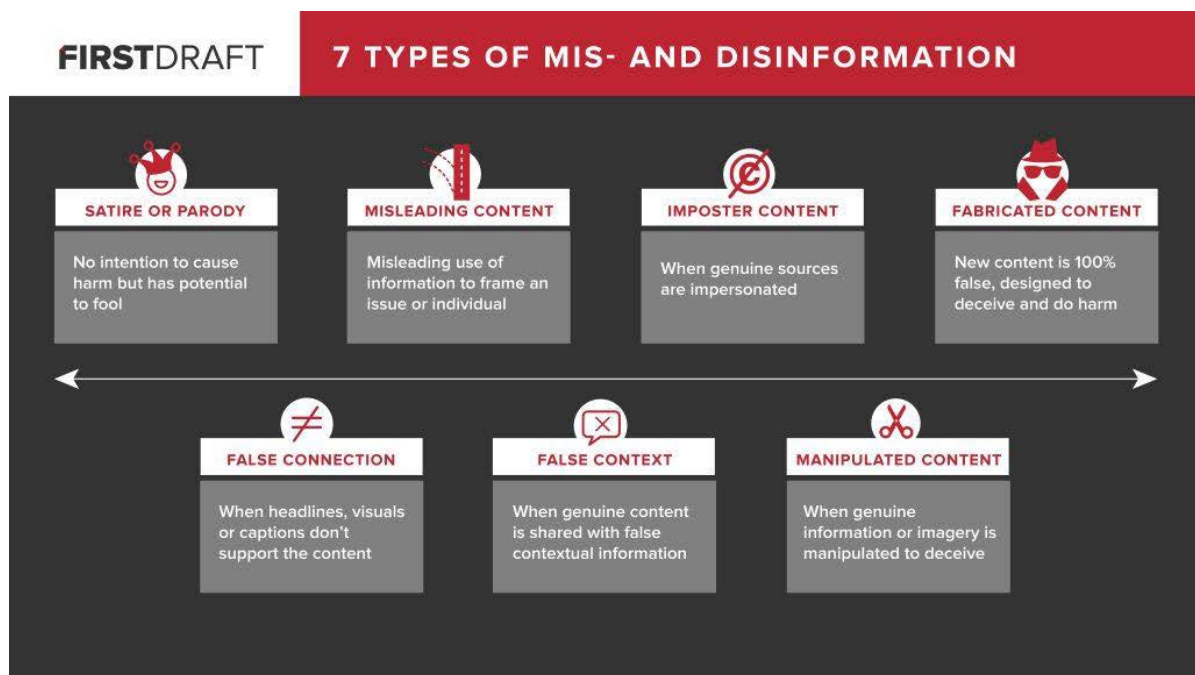
7.   Propaganda

Figure 1. Seven Types of Misinformation and Disinformation. Source: Wardle, 2017, https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79.

Figure 2 below illustrates the diverse content types which are outcomes of different motivations. For instance, false context (when genuine content is shared with false contextual information) can be a result of poor journalism or propaganda.

Figure 2. Misinformation Matrix (By Motivations). Source: Wardle, 2017, https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79.

Another example is the case of rumours. A rumour is information that has not been shown to be true and is passed from one individual to another (Allport & Postman, 1947; Rojecki & Meraz, 2016). Rumours have credibility not because they are supported by evidence, but because other people seem to believe them (Sunstein, 2014b). What is critical to note is that rumours, statements which are unverified at the point of reception, may turn out to be true, or partly or entirely false (DiFonzo & Bordia, 2007; Zubiagia *et al*., 2015). Thus, rumours can be either misinformation or disinformation. Another variable that determines whether a rumour is misinformation or disinfomation is the intent. According to Knapp (1944), there are three types of rumours – 1) those that lead to wishful thinking ('pipe-

dream' rumours), 2) those that increase anxiety or fear ('bogy' rumours), and 3) those that generate hatred ('wedge-driving' rumours). (Knapp, 1944; Zubiagia *et al.*, 2015).

2.2

Fake News

❖ The *term fake news* refers to a very specific type of false information, focusing on the 1) intent to deceive, 2) motivation (for economic gain or political influence), 3) deliberate fabrication (false statements of fact), and 4) form (assuming the appearance of an authoritative news source).

❖ The above characteristics distinguish fake news from other types of false information such as rumours, parodies, hoaxes and conspiracy theories.

<u>Definitions by Industry</u>

Fake news has attracted much attention and concern, largely due to the political events that took place in 2016. In the report *Information Operations and Facebook*, Weedon, Nuland and Stamos (2017) used the term *false news* to refer to "*news articles that purport to be factual, but which contain intentional misstatements of fact with the intention to arouse passions, attract viewership, or deceive*". The word "false", used instead of "fake", communicates the purpose and intent of communicators to mislead and deceive.

The two features — intent and designed to deceive — are echoed by (Chadwick, 2017) who defined fake news as "*fictions deliberately fabricated and presented as non-fiction with the intent to mislead recipients into treating fiction as fact or into doubting verifiable fact.*"

The following breaks down each component found in the definition:

- "Fictions" — distinguishes fake news from items which have a kernel of truth but are exaggerated or "sensationalised"

- "Fabricated" — emphasises the made-up and manufactured aspect of fake news

- "Deliberately" and "intent" — emphasises the reasons and motivations that drive the production of fake news, and differentiate fake news from flawed journalism that sometimes result from haste, carelessness, and partiality

- "Presented as non-fiction" — focuses on fake news producers' premeditation and calculation, distinct from people who simply spread it unthinkingly

- "Mislead" — indicates seriousness of purpose and distinguishes fake news from entertainment, pranks or satire.

- "Treating fiction as fact" and "doubting verifiable fact" — draws attention to the consequences of fake news.

A similarly narrow definition of *fake news* is adopted by *The New York Times.* The publication uses the term to refer to "*a made-up story with an intention to deceive, often geared toward getting clicks'* (Tavernise, 2016) — highlighting both the intention to deceive readers about the legitimacy of the content, and the implied economic imperative behind the production of fake news that typically follows a clickbait model.

The form which fake news assumes is highlighted by Mele *et al.*, 2017 who said that fake news is "*misinformation that has the trappings of traditional news media, with the presumed associated editorial processes*".

Such "trappings" are illustrated through Consultancy Buchanan Public Relations' (Keohane, 2017) comparison of fake news with "real news". *Figure 3* below presents the features of fake news and how they are different from those of real news.

Fake News	Real News
❖ Has a questionable domain or URL (may end with something like .com.co)	❖ Cites multiple sources from differing perspectives
❖ Cites only one source, or none at all	❖ Language is straightforward
❖ Stories are formulated around speculation	❖ Facts are proven with reports or statistics
❖ Stories are formulated around a single opinion, rather than all sides of the story	❖ Does not cite "alternative facts"
❖ Language is sensational and/or embellished	❖ Headline is truthful and reflects the story
❖ States conspiracy theories	❖ Contains all elements of a story – not just selective components
❖ Heavy partisan bias	❖ Minimal partisan bias
❖ Not reported on by any other news outlets	❖ Journalist has a reputable history for fair stories
❖ Notes an older publish date and time, or none at all	❖ Published date and time is recent
❖ Title does not match the content – clickbait	

Figure 3. Differences between Fake News and Real News. Source: Keohane, 2017, https://www.buchananpr.com/2017/03/what-is-the-new-definition-of-fake-news.

<u>Definitions by Academia</u>

Academics who study information seeking and the Internet have also examined the phenomenon of fake news — this section presents how they operationalised fake news.

In their study of the impact of fake news and social media during the 2016 US Presidential Election, Allcott and Gentzkow (2017) conceptualised *fake news* to be "*news articles that are intentionally and verifiably false, and could mislead readers*". Their definition provided the basis for their sample selection which included intentionally fabricated news articles (e.g. the now-defunct website *denverguardian.com* which was responsible for a widely shared article with the headline, "FBI agent suspected in Hillary email leaks found dead in apparent murder-suicide"), and articles that originated on

satirical websites (e.g. the now-defunct website *wtoe5news.com* which reported that Pope Francis had endorsed Donald Trump's presidential candidacy) but could be misunderstood as factual, especially when viewed in isolation on Twitter or Facebook feeds.

Their definition excluded what they called "close cousins of fake news", such as unintentional reporting mistakes, rumors that did not originate from a particular news article, conspiracy theories, satire that was unlikely to be misconstrued as factual, false statements by politicians, and reports that were slanted or misleading but not outright false.

A similar approach in defining fake news was adopted by Pennycook, Cannon and Rand (2017) who studied the cognitive effect of familiarity on the believability of fake news. They defined fake news as "*news stories that were fabricated and promoted on social media in order to deceive the public for ideological and/or financial gain*".

The emphasis on disinformation being disguised as news articles is also common in the field of machine learning and natural language processing. Ruchansky, Seo and Liu (2017) proposed three characteristics in the identification of fake news in machine learning — the text (e.g. whether the headline matches the body of the article, and the consistency and quality of the language), the response generated (e.g. emotions such as anger elicited by opinionated and inflammatory language), and the source of the article (e.g. structure of the URL, credibility of the media source, and profile of the journalist who authored it).

In their work which predicts and detects deception, Rubin, Chen, and Conroy (2015) focused on three types of fake news:

- *Serious fabrications* such as fraudulent reporting, yellow press, and tabloids which present unverified information and leverage eye-catching headlines ("clickbaits"), exaggerations, scandal-mongering, or sensationalism, to increase traffic or profits.

- *Hoaxing,* another type of deliberate fabrication or falsification in the mainstream or social media, where attempts to deceive audiences may be picked up and mistakenly validated and amplified by traditional news outlets (e.g. the #Columbian Chemical plant hoax).

- *Humorous fakes* (e.g. news satire) which are distinguished from fabricated news, where readers are made aware of the humorous intent and are unlikely to take the information at face value.

The preceding sections present the variations of different types of communications in the information ecology, of which fake news is a part of. While the term *fake news* has been used to refer to a wide gamut of content, its application by both industry and academia is more stringent, focusing on the intent (to deceive), motivation (for economic gain or political influence), deliberate fabrication (false statements of fact), and form (assuming the appearance of an authoritative news source). These characteristics, when combined, distinguish fake news from other types of false information such as rumours, parodies, hoaxes and conspiracy theories (with the exception of rumours which could turn out true).

3

THE PERSISTENCE OF FALSE INFORMATION:

HUMAN FACTORS

❖ Individuals tend to embrace information compatible with their preexisting beliefs and reject information that contradict them, regardless of its accuracy.

❖ Increasing familiarity to false information through repetition can increased its perceived accuracy.

❖ Correcting people's false beliefs can backfire especially when the correction contradicts their worldview.

❖ Information that evokes negative emotions like anxiety and disgust leave stronger impressions and tend to spread more quickly.

❖ Neurological processes underpin the cognitive biases that make individuals susceptible to believing false information.

❖ Social influences and social pressure may increase an individual's susceptibility to believing false information.

There is a popular conception that the persistence of false information in society today is largely driven by the Internet and social media. Indeed, the rise of the Internet and social media contributes to this problem in unprecedented ways. First, the Internet enables false information to spread rapidly and widely, often through amplification by social bots. Second, the algorithms used by social media platforms tend to create filter bubbles and echo chambers which then become breeding grounds for false information to persist (see *Section 4: The Persistence of False Information: Media Factors*).

However, there are also individual-level cognitive and socio-psychological explanations that account for why people believe and persist in believing false information.

Research conducted in the field of information processing differentiates systematic (or central) information processing from heuristic (or peripheral) information processing. In *systematic information processing*, individuals attempt to comprehend and evaluate a message's argument and truthfulness by assessing its characteristics, such as the presence or absence of high-quality arguments (Chaiken, 1980; Fridkin, Kenney & Wintersieck, 2015).

However, given the deluge of information that confronts individuals every day, people are not able to assess every piece of information in this manner. Instead, they rely on heuristics and social cues to assess the information they encounter. In *heuristic information processing*, people rely on cognitive shortcuts, such as the perceived trustworthiness and attractiveness of the information source, individual past experiences, as well as what others think. As a result, individuals typically do not interpret information in a rational, neutral and objective manner, but succumb to their own biases when

processing information they encounter (Flanagin & Medders, 2010; Lau & Redlawsk, 2001; Metzger, Tversky & Kahneman, 1975). For example, people tend to uncritically favour information that confirms their existing beliefs and dismiss counter-attitudinal information regardless of its truthfulness (see *Section 3.1: Confirmation Bias and Motivated Reasoning*).

This section focuses on the individual-level cognitive, psychological, as well as social factors that contribute to the persistence of false information in society. Why do certain people believe certain information but not others? Why do corrections sometimes fail after people have acquired and believed false information? Furthermore, why do attempts at correction at times backfire and strengthen people's belief in false information?

3.1

Confirmation Bias and Motivated Reasoning

❖ *Confirmation bias* refers to people's tendency to embrace information consistent with their preexisting beliefs and reject information that contradicts them. The more consistent the new information is with information an individual already assumes to be true, the more likely the new information will be accepted as true.

❖ *Motivated reasoning* refers to the notion that people's motivations may cause them to seek justifications for their desired conclusions, thus permitting them to believe what they want to.

❖ An individual's political beliefs and identity contribute to motivated reasoning, which can increase one's susceptibility to believing false information. It also plays a role in an individual's rejection of the validity of a scientific source.

Confirmation Bias

Research has found that humans are biased information seekers and processors — we tend to assess new information based on its logical compatibility with our preexisting beliefs. The more consistent the new information is with information an individual already assumes to be true, the more likely the new information will be accepted as true. This is known as *confirmation bias*, which refers to people's tendency to embrace information consistent with their preexisting beliefs and reject information that contradicts them. In other words, information that is compatible with a person's preexisting beliefs is more likely to be accepted as true even though it may be false (Holton & Pyszczynski, 1989; Nickerson, 1998).

A well-known experiment conducted by Lord, Ross and Lepper (1979) at Stanford University sought to understand whether people accept evidence that was in line with their preexisting beliefs at face value, while subjecting evidence conflicting with their preexisting beliefs to greater skepticism. In this study, the researchers gathered two groups of students – students who were in favour of capital punishment because they believed that it deterred crime, and students who were against it because they believed it had no effect on crime.

Both groups of students were exposed to two made-up studies — one providing data in support of capital punishment and crime deterrence, and the other providing data against it. As predicted, students who were in favour of capital punishment rated the pro-deterrence data as highly credible while rating the anti-deterrence data as unconvincing. The reverse was true for students who were against capital punishment — students who were against capital punishment rated anti-deterrence data as highly credible while rating pro-deterrence data as unconvincing.

In fact, the researchers also found that the students who were originally in favour of capital punishment became more supportive of it, and students who were originally against it became more opposed to it. This phenomenon is similar to the *backfire effect*, where people tend to hold on to their false beliefs even more strongly when presented with corrective information, especially when it challenges their worldview (see *Section 3.3: The Backfire Effect*).

Since then, other studies have also demonstrated the same phenomenon. For example, a study which sought to understand the effect of one's prejudices on processing scientific information found that individuals with higher prejudice towards homosexuals perceived fictitious scientific information that confirmed homosexual stereotypes as more convincing than individuals with lower prejudice. The reverse was true for individuals with lower prejudice towards homosexuals — individuals with lower prejudice towards homosexuals perceived fictitious scientific information that disconfirmed homosexual stereotypes as more convincing than individuals with higher prejudice (Munro & Ditto, 1997).

Confirmation bias and selective exposure also affect how people seek information. In a study by Brock and Balloun (1967), research subjects listened 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. Furthermore, other studies demonstrate that confirmation bias and selective exposure extend to people's information seeking behaviours online as well (Johnson, Bichard & Zhang, 2009; Knobloch-Westerwick *et al.*, 2015)

Besides seeking information that confirms people's preexisting beliefs, people also tend to reject information that contradict their preexisting beliefs. For example, Edwards and Smith (1996) found that arguments incompatible with one's preexisting beliefs were scrutinised for longer periods of time, subjected to more extensive refutational analyses, and eventually regarded as weaker arguments than arguments which were compatible with one's preexisting beliefs.

In short, the presence of confirmation bias in information processing suggests that individuals may have a tendency to uncritically accept false information as true simply because it is consistent with their preexisting beliefs, and resist or dismiss corrections that contradict them.

Motivated Reasoning

While existing research highlights the irrationality of confirmation bias, confirmation bias is an important coping mechanism that helps individuals make rapid decisions, while reducing the discomfort and mental effort required to hold conflicting beliefs. This is known as the *theory of cognitive dissonance*, developed by social psychologist Leon Festinger in the 1950s.

One of the ways individuals try to reduce cognitive dissonance is by engaging in motivated reasoning. *Motivated reasoning* refers to the notion that people's motivations may cause them to seek justifications for their desired conclusions, thus permitting them to believe what they want to. According to Kunda (1990), motivations can be grouped into two broad categories – 1) the motive to arrive at an accurate conclusion or 2) the motive to arrive at a directional conclusion based on one's biases.

In his book *When Prophecy Fails (1956)*, Festinger and his colleagues followed a doomsday prophet Dorothy Martin and her cult of followers who believed that aliens (known as "the boys upstairs") would come in flying saucers to save them from an impending flood.

However, when neither the flood nor the aliens came, Martin revised her prophecies repetitively to reinforce her beliefs despite all the evidence that proved her wrong (Festinger, 1956). In short, people are motivated to reduce cognitive dissonance in the face of conflicting evidence by seeking information that are consistent with their preexisting beliefs and by avoiding, ignoring or devaluing information that contradicts their preexisting beliefs.

Apart from reducing cognitive dissonance, individuals also engage in motivated reasoning to preserve their self-identity and group identity. False beliefs held by individuals are often attached to a group to which individuals belong. Thus, 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. This also explains why some people are more inclined to believe false information that others easily dismiss (Flynn, Nyhan & Reifler, 2017; Sunstein, 2014b).

A classic study in psychology demonstrates this point. Researchers screened an American football game between Princeton and Dartmouth to students from both colleges, and asked them to tally up the number of fouls observed to determine if the game was "clean and fair" or "rough and dirty". The study found that the Princeton students tend to overlook fouls committed by the Princeton team but were quick to point out the fouls made by the Dartmouth team. The reverse was true for the Dartmouth students. Researchers concluded that *the "game" was actually many different "games", and that each version of the events that transpired was just as "real" to a particular person as other versions were to other people* (Hastrof & Cantril, 1954: 132).

Political Identity and Motivated Reasoning

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 (Hartman & Newmark, 2012; Swire *et al.*, 2017; Taber & Lodge, 2006; Uscinski, Klofstad & Atkinson, 2016). For example, Republicans are more likely than Democrats to believe that President Obama was born outside the US, while Democrats are more likely than Republicans to believe that President Bush was complicit in the 9/11 attacks (Cassino & Jenkin, 2013).

In a study by Kahan *et al.* (2012), students were shown a video of a political demonstration and asked whether the protestors intended to persuade or intimidate members of the public. Some students were told that it was an anti-abortion (pro-life) protest outside an abortion clinic, while other students were told that it was a (gay-rights) protest outside a military recruitment office against the "Don't ask don't tell" policy. The study found that students who were liberal generally felt that the gay-rights protestors intended to persuade rather than intimidate members of the public, but felt that the pro-life protestors intended to intimidate rather than to persuade. Students who were conservative held the opposite view – pro-life protestors intended to persuade rather than intimidate members of the public, while gay-rights protestors intended to intimidate rather than to persuade. In short, how the students reacted depended on the congruence of the protestors' positions with their own cultural values. This suggests that same piece of false information may appeal differently to, and be interpreted differently by, different groups of people due to their political beliefs and identity.

Motivated reasoning and one's political identity also play a role in an individual's rejection of the validity of a scientific source. For example, perceptions and attitudes towards scientific evidence on climate change have been repeatedly found to be strongly associated with political beliefs and identity (Heath & Glifford, 2006; Lewandowsky, Oberauer & Gignac, 2013; Stenhouse *et al.*, 2014;

Sunstein, Lazzaro & Sharot, 2016). One study found that Republicans were more likely than Democrats to reject a scientist as a "trustworthy and knowledgeable expert" if his or her position stated that global warming is real and human-caused (Kahan, Jenkins-Smith & Braman, 2011). Another study by Pew Research Center also found that a high level of scientific knowledge did not make Republicans more likely to say that they believed in climate change, though it did for Democrats (Funk & Kennedy 2016).

In short, motivated reasoning explains that people accept and reject false information differently, depending on how the information supports or contradicts their deeply held beliefs and identity. As University of Virginia psychologist Jonathan Haidt said, "*We think we are scientists discovering the truth, but actually we are lawyers arguing for positions we arrived at by other means.*" (Haidt, 2014).

3.2

The Mechanics of Belief, Belief Perseverance, Illusory Truth Effect

❖ An increased cognitive load may increase people's susceptibility to believing false information, which in turn affects their decision making.

❖ Familiarity through repetition of false information leads to false statements being rated as true. Repeated false information feels more familiar and truer even if it goes against what an individual already knows.

While confirmation bias and motivated reasoning explain why people may believe false information, they do not explain why attempts at correcting false beliefs sometimes fail.

How Belief Systems Work

One psychological explanation is linked to the processes and mechanics behind how humans believe. Earlier theories posit that the comprehension (understanding) of a piece of information precedes and is separate from the assessment of whether it is true or false. However, social psychologist Daniel Gilbert and his colleagues argued that the acceptance of a piece of information as true is part of the automatic comprehension of that information. Its rejection only occurs, in the face of counter-evidence, subsequently to, and more effortfully than, its acceptance (Gilbert, 1991). For example, Gilbert argued that if comprehension precedes and is separate from acceptance, young children should be able to believe and disbelieve with equal ease. Instead, young children are generally gullible and prone to accept propositions uncritically because they have yet to master the intricacies of doubt. Indeed, empirical studies have demonstrated that young children are particularly vulnerable to suggestive influences about bodily touching and emotional events – see Bruck and Ceci (1999) and Loftus and Davies (1984) for comprehensive reviews on such studies.

Gilbert (1991) also argued that the rejection of a piece of information is more effortful than its acceptance. Studies have shown that cognitive resource-depleted individuals have a markedly reduced ability to reject information. For example, beliefs are most easily instilled in political prisoners deprived of sleep or when individuals

are told to perform several tasks simultaneously. In an experiment conducted by Gilbert, Taforadi and Malone (1993), subjects were asked to watch a criminal incident being reported on screen, where statements about the crime would crawl across the bottom of the screen like an emergency bulletin. The subjects were subsequently asked to play the role of a judge and recommend a prison term for the perpetrator based on the crime report they have seen. However, there was a catch – the subjects were told that the crime report would consist both true statements about the crime (which appeared in black), as well as false statements that exacerbated the nature of the crime (which appeared in red). The subjects were also separated into two groups – one group was tasked with an additional "digit search", where they were told to press a button whenever the digit "5" appeared on screen. The other group was not tasked with this "digit search".

The study found that subjects who had to perform the additional task were more affected by the false statements they read, and generally recommended prison terms which were twice as severe as the prison terms recommended by subjects who did not have to perform the additional task. Thus, an increased cognitive load may increase people's susceptibility to believing false information, which in turn affects their decision making.

In short, the mechanics behind how humans believe suggest that the comprehension of false information may automatically mean an acceptance of it as true, and that it requires subsequent effortful correction to reject the falsehood.

<u>Belief Perseverance</u>

Another factor which accounts for the persistence in beliefs of false information is a common phenomenon known as *belief perseverance*, where individuals retain newly created beliefs even after being informed that the initial information on which the beliefs were based was incorrect (Anderson, Lepper & Ross, 1980; Green & Donahue, 2011; Greitemeyer, 2014; Johnson & Seifert, 1994).

In an experiment by Ross, Lepper and Hubbard (1975) at Stanford University, researchers presented subjects with pairs of suicide notes and were told to distinguish the genuine suicide notes from fake ones. After the task, one group of subjects were told that they correctly identified the genuine suicide note 24 out of 25 times, while the other group of subjects were told that they only correctly identified the genuine suicide note 10 out of 25 times. However, this was a deception – the subjects who were told that they were right 24 out of 25 times were on average no more discerning than the subjects from the other group. Subsequently, the researchers revealed the deception to the subjects, and asked them to estimate how many genuine suicide notes they actually managed to identify. Interestingly, the subjects who were told that they were right 24 out of 25 times gave significantly higher estimates than subjects who were told they were only right 10 out of 25 times, even though they were told that the initial score given was a deception.

In other words, 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 retract or blunt its influence. While fact checking efforts can meticulously point out falsehoods, they may be limited in their ability to undo the beliefs and impressions created when people first encounter the false information.

Illusory Truth Effect

Psychological research has also established that repeated exposure to false information can influence people to believe that a falsehood is true (Arkes, Hackett & Boehm, 1989; Polage, 2012). Politicians (for example those who made the accusation that Britain sends £350 million a week to the European Union) and marketers are often great manipulators of this effect. The *illusory truth effect* was first observed by Hasher, Goldstein and Toppino (1977), who found that research subjects rated repeated statements as truer than new statements. The researchers concluded that *"repetition of a plausible statement increases a person's belief in the referential validity or truth of that statement"* (Hasher, Goldstein & Toppino, 1977: 111).

The illusory truth effect was also supported in a study by Begg, Anas and Farinacci (1992) which confirmed that familiarity through repetition of false information leads to false statements being rated as true. A more recent study on fake news on social media during the 2016 US Presidential Election also found that the illusory truth effect was a factor in the spread of fake news on social media, as the degree of belief increases with the frequency of exposure to fake news stories (Goodwin-Ortiz de Leon, 2017).

While Hasher, Goldstein and Toppino (1977) demonstrated that familiarity increases perceived accuracy of plausible information, another study on the effect of fake news on the 2016 US Presidential Election showed that the illusory truth effect extends to highly implausible and partisan statements as well. In this study, researchers found that a single exposure to a fake news headline was sufficient to lead to an increased perception of accuracy. A second exposure to a fake news headline led to an even greater perception of accuracy with the effect compounding over time. Furthermore, the increased perceptions of accuracy for familiar fake news headlines occurred despite the presence of explicit warning labels that indicated that the story was contested by fact checkers (Pennycook, Cannon & Rand, 2017). Research also found that the illusory truth effect happens even to people with knowledge about the topic (Fazio *et al.*, 2015).

Cumulatively, these research findings suggest that exposing people to false information will increase belief in the false information as people rely on familiarity as heuristic in their cognitive processing. Repeated false information feels more familiar and truer even if it goes against what an individual already knows. As a result, exposure to false information may have long term effects while corrections may unfortunately be short lived (see *Section 4.3: Bots and False Amplifiers* on how amplification on the Internet and social media can exacerbate the illusory truth effect).

3.3

The Backfire Effect

❖ The *overkill backfire effect* occurs when relatively simple and straightforward false beliefs are debunked with over complicated corrective information. This is because information that is succinct and easy to understand is more likely to be accepted as true. Thus, we should keep corrective information lean, mean and easy to read.

❖ The *worldview backfire effect* is particularly strong when it involves corrective information that challenges people's sense of cultural identity and worldviews, i.e. their fundamental beliefs about how society should operate. Thus, we need to present corrective information in a manner that does not trigger a defensive emotional reaction if we want people to accept new evidence.

The standard measure used to counter or correct people's belief in false information is to persuade via fact checking, evidence, and argument. However, head-on attempts to correct people's belief in false information can sometimes trigger a *backfire effect*, where people not only fail to change their minds in the face of facts, but instead become even more committed to the false information (Lewandowsky *et al.*, 2012; Nyhan, Reifler & Ubel, 2013). This suggests that attempts at corrections such as fact checking erroneous statements may cause people to endorse their misperceptions even more strongly.

Cook and Lewandowsky (2011) identified different types of backfire effects.

Familiarity Backfire Effect

Refuting a false belief often involves mentioning the false information itself, thus making people more familiar with the falsehood. Similar to the illusory truth effect (mentioned in the earlier subsection), the *familiarity backfire effect* is based on the idea that familiarity towards a piece of information increases its chances of being accepted as true. Thus, the act of debunking false information may reinforce it in people's minds. This is demonstrated by a study conducted by Skurnuk *et al.* (2005), 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.

Overkill Backfire Effect

The *overkill backfire effect* occurs when relatively simple and straightforward false beliefs are debunked with over complicated corrective information. This is because information that is succinct and easy to understand is more likely to be accepted as true (Schwartz *et al.*, 2007). Thus, simple and straightforward false beliefs may become cognitively more attractive to people if the counter-arguments presented are overly complicated (Cook & Lewandowsky, 2011). Thus, we should keep corrective information lean, mean and easy to read.

Worldview Backfire Effect

Research has shown that the *worldview backfire effect* is particularly strong when it involves corrective information that challenges people's sense of cultural identity and worldviews, i.e. their fundamental beliefs about how society should operate. In a study conducted by Nyhan and Reifler (2010) for example, researchers presented subjects with a fake newspaper article that confirmed that Iraq possessed weapons of mass destruction prior to the US invasion in 2003. Subjects were then presented with a corrective news article which clarified that Iraq did not possess weapons of mass destruction. The study found that conservatives became even more likely to believe that Iraq possessed weapons of mass destruction after reading the correction news article. Liberals, on the other hand, accepted the correction and rejected the old article. Another study also found that Republicans' support for climate change policies decreased after being presented with messages highlighting the adverse health effects of climate change, whereas Democrats increased their support instead (Hart, Nisbet & Shanahan, 2011).

Backfire effects are not only limited to the correction of false information, but also found in other forms of communication as well. For example, campaigns that aim to reduce smoking and promote positive health behaviours can also backfire and result in an increase in smoking rates (Byrne & Hart, 2009). The same phenomenon was

demonstrated by Nyhan *et al.* (2014) which showed that pro-vaccine messages do not always work as intended, and that the effectiveness of those messages may vary depending on parental attitudes toward vaccines.

In short, people's belief in false information is often resistant to correction, especially when the correction challenges one's worldview. Thus, we need to present corrective information in a manner that does not trigger a defensive emotional reaction if we want people to accept new evidence. One promising approach is derived from the *inoculation theory* where people are prepared for potential false information *a priori* (see *Section 6.2: Medium-term Measures: Inoculation, Leveraging Machines and Improving Journalism*).

3.4

Emotions

> ❖ Studies have shown that information that evokes negative emotions is processed more thoroughly, leaves a stronger impression, and also more resistant to disconfirmation than information that evokes positive emotions.
>
> ❖ Research has also found that information that evokes an emotional response is more likely to be shared and passed on irrespective of its believability.

Information processing is not a purely cognitive process that is devoid of emotions. Emotions play a key role in how people process false information. For example, research has shown that emotions like anger and anxiety have different effects on whether people believe in false information. Weeks (2015) found that in the face of false information, anger encourages people to process information in a partisan manner, resulting in beliefs that reinforce their party affiliation. On the other hand, anxiety encourages less reliance on partisanship during information processing and more on the information environment, resulting in people being more open-minded to views outside of their political beliefs instead.

Research also found evidence of a *negativity bia*s, where information that evoke negative emotions in general has a stronger impact than information that evoke positive emotions. Studies have shown that information that evokes negative emotions is processed more thoroughly, leaves a stronger impression, and also more resistant to disconfirmation than information that evokes positive emotions (Baumeister *et al.*, 2001; Ito *et al.*, 1998).

Emotions also influence the sharing and dissemination of information. While the believability of a piece of information is a determining factor of whether it will be shared, research has found that information that evokes an emotional response is more likely to be shared and passed on irrespective of its believability (Berger, 2011). For example, Heath, Bell and Sterberg (2001) found that in the case of rumours, the emotion "disgust" (evoked by the rumours) encourages people to share them. Contemporary urban legends are also more likely to spread on the Internet if they evoke disgust. Other studies have also pointed to the role of fear and anxiety in influencing belief and spread of rumours and conspiracy theories (Rosnow, 1980; Rosnow, Esposito & Gibney, 1988; Uscinski, Parent & Torres, 2011).

Evidence for the phenomenon of negativity bias was also found in the sharing and transmission of information. Bebbington *et al.* (2017) found that the negativity bias favours the social transmission of unambiguously negative story events over unambiguously positive ones.

Together, these findings explain why false information involving child abuse, deviant sexual behaviour and flesh-eating bacteria are more "successful" or endemic than others. Information that produce strong negative emotions like disgust, anger, and fear are more likely to be spread by people, regardless of whether they are true or false.

3.5

Neurology

❖ Research suggests that there may be a neurological basis underlying confirmation bias, and that using confirmation bias to make decisions make us feel good in the same way people experience the positive effects of alcohol or opiate.

❖ Neuroimaging studies found that when people 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.

While the preceding sections explained the individual-cognitive and psychological factors behind why some people believe in false information and why corrections sometimes fail, this section provides the neurological explanations underlying some of these psychological factors.

As mentioned in *Section 3.1: Confirmation Bias and Motivated Reasoning*, confirmation bias is a necessary cognitive technique that allows individuals to make rapid decisions, and reduce the discomfort and mental effort required to hold to conflicting beliefs. However, research suggests that there may in fact be a neurological basis underlying confirmation bias, and that using confirmation bias to make decisions make us feel good in the same way people experience the positive effects of alcohol or opiate. Individuals get a rush of dopamine (the neurotransmitter that is released whenever humans anticipate or experience a rewarding or pleasurable stimulus) when they find confirming data, similar to when they eat chocolate, have sex, or fall in love. Thus, people tend to focus on information that supports their confirmation bias and ignore information that contradicts their beliefs. *In other words, it feels good to "stick to our guns" even if we are wrong* (Gorman & Gorman, 2016: 135).

Another study published in *Nature* demonstrated that challenging an individual's political beliefs activates the areas of the brain that are involved in personal identity and emotional response to threat. Researchers used neuroimaging to investigate the neural systems involved in maintaining political beliefs in the face of counter-evidence. They presented 40 liberals with arguments that contradicted their strongly held political beliefs. 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. However, activation in those areas of the brain was lower in subjects who changed their

minds upon evaluating counter-evidence (Kaplan, Gimbel & Harris, 2016). Thus, this suggests that humans may in fact be neurologically "hardwired" to engage in motivated reasoning and hold onto our preexisting beliefs in the face of counter-evidence.

3.6

Social Influence

> ❖ In *conformity cascades*, people tend to go along with the majority (regardless of their private views and doubts) in order to conform to the positive expectations of others and maintain membership in the group.
>
> ❖ In *informational cascades*, people tend to believe in a rumour because a significant number of others appear to believe it as well.

Besides the psychological and neurological factors mentioned in the earlier subsections, social influence and social pressure may also increase people's susceptibility to believing false information.

As Sunstein (2014b) argued in his book *On Rumours*, rumours often spread through conformity cascades (i.e. normative social influence) and informational cascades (i.e. informational social influence).

In *conformity cascades*, people tend to go along with the majority (regardless of their private views and doubts) in order to conform to the positive expectations of others and maintain membership in the group. The effect of conformity cascades are especially strong in social networks of tightly knit groups, and with a strong stake in the group beliefs.

In *informational cascades*, people tend to believe in a rumour because a significant number of others appear to believe it as well. When people lack information of their own, they tend to rely on the words, opinions, and deeds of others as valid evidence of reality (Price, Nir & Cappella, 2006). This is also partly because of the *false consensus effect* where people tend to think that something is probably true if they hear it from many other people (Ross, Greene & House, 1977). In fact, this false consensus effect is often amplified on the Internet and social media, where volume can speak volumes (see *Section 4.3: Bots and False Amplifiers*).

In short, people might be more susceptible to believing false information if many others appear to believe as well, or if they are willing to overlook their own skepticisms in order to affirm and demonstrate their membership in a group.

4

THE PERSISTENCE OF FALSE INFORMATION:

MEDIA FACTORS

❖ Source-layering on the Internet and social media complicates individuals' assessment of source credibility, encouraging them to rely on cognitive heuristics even more.

❖ Filter bubbles and echo chambers act on individuals' cognitive biases, further entrenching their worldviews.

❖ Bots and false amplifiers amplify the propagation of false information on the Internet and social media.

❖ Decline in trust in journalism and media encourages people to seek alternative sources of information, thus increasing their susceptibility to believing false information.

As mentioned in *Section 1: Executive Summary*, false information is not a new problem. Propaganda, misinformation and disinformation have been used throughout history to influence public opinion. The research highlighted in *Section 3: The Persistence of False Information: Human Factors* demonstrates how individual-cognitive and psychological biases in information seeking and processing may increase people's susceptibility to believing and spreading false information. However, the emergence of the Internet and social media has presented novel challenges in terms of dealing with the persistence of false information, and has exacerbated the negative effects of existing human biases.

The Internet and social media have reduced the barriers of entry into the media industry. While such changes in infrastructure have contributed to the breakdown of authoritarian regimes (as seen in the case of Arab Spring), this new structure has simultaneously allowed individuals with the technical and socio-political know-how to produce and distribute false information easily for various motivations (see *Section 5: Consumers and Propagators of False Information*). Social media platforms have also allowed false information to reach a wide audience with relative ease — active Facebook and Twitter users have reached 1.8 billion and 400 million per month respectively (Allcott & Gentzkow, 2017). Furthermore, this new infrastructure has also allowed false information to spread very quickly across different platforms, further complicating efforts to counter the spread of false information.

The emergence of the Internet and social media has also shifted news consumption to online and social media platforms. This has disrupted traditional business models of journalism and impaired

audiences' trust in the legitimacy of longstanding media institutions. This is largely because the proliferation of information sources has made assessing the credibility of information on the Internet and social media platforms increasingly difficult.

With an abundance of information sources, people rely even more heavily on heuristics and social cues to determine the credibility of information they encounter. Combined with the tendency for people to follow like-minded people, this results in the creation of filter bubbles and echo chambers, which are in turn reinforced by the algorithms used by many social media platforms. Bots and false amplifiers also contribute to the propagation of false information on the Internet and social media. These provide an ideal breeding ground of false information to spread and persist.

This section focuses on how the Internet and social media contribute to the rise of filter bubbles and echo chambers, bots and false amplifiers, source-layering and declining trust in journalism, and how all these factors in turn increases the persistence of false information in our society.

4.1

Source-layering and Source Credibility

❖ Information presented in traditional media typically features one salient source, which makes assessing source credibility relatively straightforward. However, the information people encounter online tends to come with multiple source cues presented on the interface.

❖ Research has shown that the most proximate source cue (often our friends) exerts the greatest influence (in terms of credibility assessment) than other distant source cues also present on the interface.

❖ Research has also shown that people are less skeptical of information that they encounter in customised and personalised environments.

The emergence of the Internet and social media has revolutionised the information environment and created one where information is diverse and abundant. However, this increased diversity and abundance of information sources have also complicated the standard ways of assessing the credibility or the believability of information, hence challenging traditional notions of credibility accorded to central authorities like governments and mainstream media (Callister Jr., 2000). With lowering costs and complexities involved in producing information, Internet and social media users have shifted from being passive consumers of information, to being active content producers on platforms such as Facebook, Twitter and blogs. This change in the nature of information production and consumption has prompted research on its impact on credibility evaluation. While findings are mixed and inconclusive, some studies show that online sources enjoy similar levels of credibility as traditional sources. For example, a study that surveyed blog users found that blog users rated blogs as highly credible, in fact even more credible than traditional sources of information (Johnson & Kaye, 2004).

In general, the proliferation of information sources online has caused individuals to rely more on cognitive heuristics (rather than systematically processing information) to assess the credibility of information sources (Metzger, Flanagin & Medders, 2010). Often, these "shortcuts" used to assess the 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 of a website such as the overall visual appeal, layout, typography, font size and colour schemes to assess a website's credibility (Fogg *et al.*, 2002; Wathen & Burkell, 2002). The reliance on such peripheral cues for

credibility assessment of information may increase one's susceptibility to believing false information, especially since many fake news websites are designed to look and function in a professional manner.

Another unique feature of consuming information on the Internet and social media is that of *source-layering*. In traditional mass media, the information presented typically features one salient source, which makes assessing source credibility relatively straightforward. However, the information people encounter online tends to come with multiple source cues presented on the interface. As Pennsylvania State University Media Studies Professor S. Shyam Sundar said, *"Imagine checking your Facebook news feed and seeing something your friend has shared: a politician's tweet of a newspaper story. Here, there's actually a chain of five sources (newspaper, politician, Twitter, friend and Facebook). All of them played a role in transmitting the message, obscuring the identity of the original source."* (Sundar, 2016).

Sundar's research found that given such a chain of sources, readers generally do not process all visible source cues when assessing the credibility of the information. Instead, the most proximate source cue (e.g. the friend) exerts the greatest influence (in terms of credibility assessment) than other distant source cues (e.g. the politician) also presented on the interface. Furthermore, only people who are highly involved in the information topic would be more likely to systematically analyse the chain of source cues during credibility assessment (Kang *et al.*, 2011).

Given the findings of his research, Sundar further argued that because our friends tend to be perceived as the most proximate source cue, we tend to be less skeptical about the information shared by our friends because we trust them. A more recent study by Jun, Meng and Johar (2017) supported this idea and found that the perceived social presence on social media platforms reduced the likelihood of people's fact checking efforts. Thus, this can in turn increase people's chances of uncritically accepting false information as accurate.

Research has also shown that people are less skeptical of information that they encounter in customised and personalised environments.

One example of a customised and personalised site is Google News, where individuals can tailor the information presented to them according to their interests. Kang and Sundar (2016) found that people who customised their own news website tend to process information that they encountered less systematically and more heuristically. This is because of a "self-as-source" effect, where customisation increases users' perceived control and perceived identity. This may predispose users to be less skeptical of the information they encounter and be more easily persuaded by the message content.

4.2

Filter Bubbles and Echo Chambers

❖ Algorithms used by social media platforms are essentially prediction machines – they look at the things people have done, the things people have liked, and the things people's friends have liked, and try to extrapolate this information.

❖ Algorithms are also being used to personalise information flow. This results in an algorithm-driven filter bubble that surrounds us with ideas consistent with our preexisting beliefs, making us overconfident in our mental frameworks and dramatically amplifying our confirmation biases.

The term *filter bubble* was coined and popularised by Eli Pariser in his book titled *The Filter Bubble* (2011). It refers to how social media platforms like Facebook and Twitter tend to lock users into personalised feedback loops, each with its own news sources and political leanings, thus fundamentally altering the way people encounter ideas and information. This results in isolated online communities where people largely consume information that reinforces their worldview in the absence of conflicting ideas (Pariser, 2011).

Pariser argued Internet giants like Google and Amazon use algorithms to customise and personalise advertisements as part of their business strategy so that users are more likely to buy the products they offer. The algorithms used are essentially prediction machines – they look at the things people have done, the things people have liked, and the things people's friends have liked, and try to extrapolate this information. However, personalisation has gone beyond targeted advertising.

Algorithms are now being used to personalise information flow as well. For example, websites like *Yahoo News* and *The New York Times* funded a start-up called *News.me* to cater their headlines to audiences' interests and desires. Facebook and Twitter offer personalised news feeds, which are increasingly becoming people's primary news source. A 2016 Pew Research Center study found that 62 percent of US adults get their news on social media and 18 percent do so often (Gottfried & Shearer, 2016). Ultimately, this turns into a highly homogenous echo chambers where individuals largely discuss similar views with like-minded people, and fail to penetrate into other filter bubbles for alternative views. (Halberstam & Knight, 2016; Pariser, 2011).

While information filtering is not a new phenomenon (as mentioned in *Section 3.1: Confirmation Bias and Motivated Reasoning*), the algorithm-driven filter bubble surrounds us with ideas consistent with our preexisting beliefs, making us overconfident in our mental frameworks and dramatically amplifying our confirmation biases (Pariser, 2011). A study on the spread of misinformation on Facebook found that the homogeneity of echo chambers was the primary driver of misinformation online (Del Vicario *et al.*, 2015). Filter bubbles may also increase group polarisation, where deliberation among like-minded people often entrench false information such as false rumours (Allport & Postman, 1947; Sunstein, 2014b).

However, data scientists at Facebook published a study that looked at how 10.1 million of the most partisan American Facebook users interacted with socially shared news, and argued that the filter bubble/echo chamber effect may not be as potent as many people feared it to be. The study found that an average of 29 percent of the news stories displayed by Facebook's news feed presented views that conflicted with the user's ideology. In addition, the study also showed that individuals' choices of what information to consume had a stronger effect than Facebook's filtering algorithm (Bakshy, Messing & Adamic, 2015).

The results of the study drew strong criticism against Facebook for downplaying the echo chamber effect (Manjoo, 2015). Despite the contentions, the findings of the study demonstrate that the filter bubble/echo chamber effect is real (even though its effect may be smaller than expected), and that Facebook's algorithm increases people's chances of encountering information (both true and false) that reinforces their worldview.

In short, algorithm-driven filter bubbles work on people's cognitive biases like steroids, optimising the information they encounter as well as associating them with like-minded people and views such that their personal biases are further entrenched.

4.3

Bots and False Amplifiers

❖ Bots and false amplifiers can contribute to the spread of false information by automatically sharing or retweeting false content.

❖ As a result, bots and false amplifiers allow certain views and voices to be spoken more loudly than others, and manufacture a false consensus by creating an illusion that a particular idea is highly popular and endorsed by many.

Bots are software that create content on social media and interact with people. In Facebook's report titled *Information Operations and Facebook*, these are referred to as false amplifiers — inauthentic accounts created with the intention to manipulate political discussion (rather than financial gains). Some of these false amplifiers are automated, while others are driven by coordinated groups of people who are committed to operating inauthentic accounts. Regardless, networks of bots and false amplifiers have the ability to share content at high volumes, often to promote or disparage specific causes or issues (e.g. propaganda), sow distrust in political and social institutions, or simply to spread confusion (Marwick & Lewis, 2017; Weedon, Nuland & Stamos, 2017).

Bots can be used in a variety of ways. For example, bots that automatically aggregate content across various sources to produce news feeds are generally benign and helpful. However, because bots are also cheap and easy to deploy, they can be leveraged by governments and individuals to spread information (as well as false information) quickly and widely. For example, bots have been used by governments and political elites in countries like Azerbaijan and Venezuela to disseminate pro-government messages and demobilise opposition (Forelle *et al.*, 2015; Woolley & Howard, 2016). Bots have also been employed to spread false information and negative information about opposition candidates during elections (Arsenault, 2016; Misener, 2016).

Bots and false amplifiers can contribute to the spread of false information by automatically sharing or retweeting content without verifying the facts or checking the credibility of the source (Ferrara *et al.*, 2016). This problem is most serious on Twitter because the platform offers users anonymity (unlike Facebook which requires users to use their real names), thus allowing people to use online tools to easily create thousands of Twitter bots (Manjoo, 2017). For

example, a study done on the spread of fake Twitter content relating to the Boston Marathon blasts found that 29 percent of the viral content was fake, and that the propagation of fake Twitter content was driven by inauthentic Twitter accounts (Gupta, Lamba & Kumaraguru, 2013). Furthermore, bots can be difficult to be distinguished from real accounts by the average user because these inauthentic accounts often use pictures of real people, natural language algorithms, and even engage in conversations with other users in order to "look and sound human". This also makes bot detection an increasing challenge (Ferrara *et al.*, 2014).

In short, bots and false amplifiers allow certain views and voices to be spoken more loudly than others, and manufacture a false consensus by creating an illusion that a particular idea is highly popular and endorsed by many (see *Section 3.6: Social Influence*). In the long run, this may drown out valid and accurate information.

4.4

Declining Trust in Media and Journalism

> ❖ Low public trust in the mainstream media helps to create the conditions for the emergence of "media manipulators" who produce false information for a variety of reasons.
>
> ❖ People who do not trust the media are more likely to have more partisan and less accurate beliefs about the world.
>
> ❖ Low levels of trust in institutions can cause people to resort to conspiracy theories answers and alternative narratives.

Several media experts and observers have also commented that the growing problem of false information that we currently face is merely a symptom of greater social ills.

Writer and editor of *The New York Times* Serge Schmemann said that *"fake news is just a symptom of the crisis that traditional journalism is experiencing, such as the lowering of journalistic standards and quality due to financial pressures and time constraints"* (Bouteldja, 2017).

Professor of Journalism at City University of New York Jeff Jarvis said that the real problem underlying fake news is one of trust and manipulation. Low public trust in the mainstream media helps to create the conditions for the emergence of "media manipulators" who produce false information for a variety of reasons (Jarvis, 2017; Madden, Lenhart & Fontaine, 2017).

Professor of Communications and New Media at the National University of Singapore Mohan Dutta echoed this idea and said that *"fake news is not a product of lack of regulation of new and digital public spheres, but rather a product of a climate of backhanded control of the public sphere by the power elite, [which has] contributed to public insecurities and public mistrust … thus further feeding the production, consumption and circulation of fake news."* (Dutta, 2017).

Many studies have indeed found a decline in trust in the media. A Gallup poll released in September 2016 found that Americans' trust in the mass media *"to report the news fully, accurately and fairly"* was at 32 percent, the lowest in Gallup's polling history (Swift, 2016). A Data and Society report also found that most teens express distrust of the news and assume much of it to be biased (Madden, Lenhart & Fontaine, 2017). The 2017 Edelman Trust Barometer also revealed that trust in the media is at an all-time-low (43 percent), and that the media is distrusted in more than 80 percent of the countries it surveyed. In fact, people now generally trust leaked information

much more than traditional news sources, and trust algorithms over human editors ("2017 Edelman Trust Barometer", 2017).

Research has shown that people who do not trust the media are more likely to have more partisan and less accurate beliefs about the world. This is because people who distrust the media generally update their beliefs in response to current affairs less frequently, and instead rely on their partisanship to form beliefs about events (Ladd, 2011). Another study also found that low levels of trust can cause people to resort to conspiracy theories answers and alternative narratives (Morris, n.d.). Similarly, Sunstein (2014a) argued that if institutions are generally trustworthy and embedded in an open society with free flow of information, conspiracy theories would usually be unjustified by its people. On the other hand, individuals living in closed societies (e.g. lacking a free press) would have stronger reasons to distrust institutions, thus making conspiracy theories more believable.

In the context of the US, Allcott and Gentzkow (2017) suggested that Republicans' decline of trust in the mainstream media may be due to the perception that the mainstream media is pro-Clinton, thus increasing their demand for non-traditional sources of information and susceptibility to false information as well (also see *Section 5.1: Who are Susceptible to False Information?*).

5

CONSUMERS AND PROPAGATORS OF FALSE INFORMATION

- ❖ People who hold extreme political beliefs are more susceptible to believing false information.

- ❖ People who perceive loss of power and social exclusion are more susceptible to believing false information.

- ❖ Education is a double-edged sword when it comes to resisting false information.

- ❖ Anecdotal evidence suggests that there may be a relationship between age and susceptibility to believing and sharing false information.

- ❖ Governments, organisations and individuals with vested interests, and Internet subculture groups contribute to the production and propagation of false information.

While the factors discussed in the earlier sections (*Section 3: The Persistence of False Information: Human Factors* and *Section 4: The Persistence of False Information: Media Factors*) suggest that anyone who is influenced by them may be equally susceptible to believing false information, research has shown that certain groups of people may be more susceptible to believing false information than others.

In this section, we focus on the characteristics of these groups of people – their political traits, perceptions of power and threats, and education level.

We also look at the other side of the story – false information producers. Who produces and propagates false information, and what are some of their motivations?

5.1

Who are Susceptible to False Information ?

❖ People with extreme political beliefs tend to be more distrustful of social and political institutions than people with moderate political beliefs, which may encourage them to endorse conspiracy theories or other forms of false information.

❖ Conspiracy theories resonate with communities that feel disempowered and socially excluded, suggesting that belief in false information in general may follow the same trend.

Political Traits

As mentioned in *Section 3: The Persistence of False Information: Human Factors*, research has found that a person's political beliefs is an important factor in contributing to motivated reasoning, thus increasing one's susceptibility to believing false information (Taber & Lodge, 2006; Uscinski, Klofstad & Atkinson, 2016). For example, Republicans are more likely than Democrats to believe that President Obama was born outside the US, while Democrats are more likely than Republicans to believe that President Bush was complicit in the 9/11 attacks (Cassino & Jenkin, 2013). Similarly, one's political beliefs also influence a person's receptivity to fact checking efforts, and thus influencing the likelihood of one's false beliefs being successfully corrected. For example, Republicans are more receptive than Democrats to fact checking results that authenticate claims in favour of Republicans, and vice versa (Fridkin, Kenney & Wintersieck, 2015).

In general, people with extreme political beliefs tend to be more distrustful of social and political institutions than people with moderate political beliefs, which may encourage them to endorse conspiracy theories or other forms of false information. Research has demonstrated this relationship between political extremism and belief in conspiracy theories. A study conducted in the US and The Netherlands asked participants to classify themselves on a political left-right dimension and respond to questions pertaining to conspiracy beliefs. The study found that people who hold extreme political beliefs (at both the left and the right of the political spectrum) were most susceptible to conspiracy beliefs. The researchers also suggested that certain ideologies may drive certain conspiracy theories – people at the extreme left of the political spectrum might be more susceptible to conspiracy theories on capitalism and multinationals, while people at the extreme right of the political

spectrum might be more susceptible to conspiracy theories on science and climate change (Swami, 2012; Van Prooijen, Krouwel & Pollet, 2015). Another study also found that people who are highly knowledgeable about politics and lacking in trust institutions are most susceptible to endorsing conspiracy theories (Miller, Saunders & Farhart, 2015).

Research has also shown that susceptibility to believing false information is currently a predominant problem of the political right. In a study on fake news and social media in the 2016 US Presidential Election, researchers found that Republicans were less likely to recognise fake articles as false, and that Democrats were more likely to correctly identify true versus false articles. In fact, the study also found that Republicans were generally more credulous of both true and false articles. Researchers also found that fake news stories that were widely shared were heavily tilted in favour of Donald Trump, with 115 pro-Trump fake news stories being shared on Facebook 30 million times, as compared to 41 pro-Clinton fake news stories being shared 7.6 million times (Allcott & Gentzkow, 2017).

Anecdotal evidence also appeared to support this idea, leading Jestin Coler, fake news creator and Founder of Disinfomedia, to conclude that fake news works best with Trump supporters, and that efforts to produce fake news for liberals hardly ever worked because it *"will get debunked within the first two comments and then the whole thing just kind of fizzles out"* (Sydell, 2016).

Winners and Losers; Insiders and Outsiders

While there might be evidence to suggest that belief in false information is predominantly a pathology of the political right, there is also evidence to suggest a reversal of the situation where liberals are increasingly falling for false information, especially after President Trump's inauguration. A survey administered by Survey Sampling International in 2016 found that Republicans were more likely to agree with conspiracy-minded statements before the election, but Democrats became more likely to do so after the election (Nyhan,

2017). Brooke Binkowski (managing editor of fact checker *Snopes*) also observed an increase in the amount and popularity of fake news targeted at liberals, such as (what she termed as "wishful-thinking type") stories that suggest Trump's impeachment or that government insiders are planning to rebel against Trump (Meyer, 2017).

The above phenomena can be explained by studies on conspiracy theories, power dynamics and threat perception. According to Uscinski and Parent (2011), conspiracy theories are used as coping mechanisms by vulnerable groups to manage their anxieties about perceived threats and loss of power. They analysed letters to the *The New York Times* editor from 1897 to 2010 for conspiracy talk and found that conspiracy theories about the right increased when a Republican was president and decreased when a Democrat was president, while conspiracy theories about the left increased when a Democrat was president and decreased when a Republican was president. In other words, conspiracy theories resonate with communities that perceive threats and feel disempowered, suggesting that belief in false information in general may follow the same trend.

Research has also shown that the feeling of being socially excluded is associated with the endorsement of conspiratorial and superstitious beliefs, as they provide a way for individuals to make meaning of their environment (Graeupner & Coman, 2017). This suggests that false information may thrive better within communities that feel socially excluded and alienated from the rest of society.

Education

Research has found that individuals' education level also plays a role in influencing one's susceptibility to believing false information. A study by Van Prooijen (2017) found that people with high education are less likely than people with low education to believe in conspiracy theories. The reason was people with high education were less likely to believe in simple solutions for complex problems, felt more powerful and thus more in control of their circumstances. This also

confirms the earlier point that disempowerment may increase people's likelihood to believe in conspiracy theories.

However, other research has also shown that education may have an opposite effect on an individual's susceptibility to believing false information — individuals with high education or high levels of knowledge tend to be better equipped to resist and counter-argue against information conflicting with their preexisting beliefs (Flynn, Nyhan & Reifler, 2017).

Age

Anecdotal evidence has also suggested that the elderly might be particularly susceptible to believing and sharing false information. The elderly are increasingly using smartphones and becoming active on social media platforms like Facebook, and Instant Messaging platforms like WhatsApp. However, being digital non-natives, they may lack the digital literacy to point out inaccuracies in information they encounter on these platforms, and hence share false information uncritically. Chairman of Singapore's Media Literacy Council Lok Wai Han said that the elderly might often be *"guilty of propagating unverified information such as political gossip or dubious health tips"* (Au-Yong, 2017c; Seow, 2017).

However, no empirical study on the relationship between age and susceptibility to false information was found at the time of writing this literature review.

5.2

Who Propagates False Information?

❖ Governments and politicians may propagate politically or ideologically motivated false information in society as an attempt to change people's political beliefs or influence public opinion on an issue.

❖ Individuals and organisations may have vested interests to produce and propagate false information, as seen from the examples of Macedonian teenagers and the tobacco industry.

❖ Internet subculture groups like trolls, hate groups and conspiracy theorists manipulate the online media and produce false information to achieve their respective goals.

Governments and Politicians

Governments and politicians may propagate politically or ideologically motivated false information in society as an attempt to change people's political beliefs or influence public opinion on an issue.

One example of propagating politically motivated false information was during the Bush administration's "War on Terror" after the 9/11 attacks. Prior to the invasion of Iraq by the US in 2003, the government proclaimed that Iraq undoubtedly possessed weapons of mass destruction, and implied that the US had intelligence that linked Iraq to terrorist group Al-Qaeda. However, both claims turned out to be untrue (Lewandowsky *et al.*, 2012). Scholars have argued that the cultural construction and political rationale behind the framing of this "War on Terror" security policy eventually laid the groundwork for the invasion of Iraq (Reese & Lewis, 2009).

A more recent example of false information being propagated by politicians is the "Birther Conspiracy" started by Donald Trump. In 2011, Trump began to question the birthplace of President Obama during television interviews, accusing him of not being born in the US. Even after Obama had released his long-form birth certificate, Trump continued to question the authenticity of his birth certificate, suggesting that it was a fraud. Trump also went on to falsely accuse his political rival, Hillary Clinton, of starting the "Birther Conspiracy" movement during the 2016 US Presidential Election (Babaro, 2016; Greenberg & Qiu, 2016).

Russian disinformation is another prime example of politically motivated false information propagated by governments. As mentioned in *Section 2: Defining "Fake News"*, disinformation refers to "non-accidentally misleading information" created with the intent to sow discord between enemy ranks (Fallis, 2015). One example of

Russian disinformation began during the 1980s when the Los Angeles Centre for Disease Control started witnessing increasing cases of AIDS primarily among homosexual men. Soviet intelligence made use of this opportunity then and carried out *Operation Infecktion*, where they systematically disseminated false information that AIDS was developed by an American scientist in a secret biological weapons laboratory (Grimes, 2017). More recently, the surprising result of the 2016 US Presidential Election also raised questions about the impact of Russian disinformation on the outcomes of the election.

Vested Interests: Individuals and Organisations

Individuals may stand to gain financially through the production and propagation of false information, especially by running fake news websites. In November 2016, *Buzzfeed* reported that at least 140 political websites in the US could be traced to the Macedonian town of Veles. Many of these websites had American-sounding domain names (e.g. *DonaldTrumpNews.co* and *USADailyPolitics.com*) and published pro-Trump content targeted at conservatives and Trump supporters. Most of the people behind these websites were Macedonian teenagers who ran the sites for economic incentives.

A high traffic to websites like the above is rewarded by automated advertising engines such as Google's *AdSense* and *Revcontent*. Facebook's self-service ad technology also generates revenue for these sites whenever fake news articles were clicked on or shared. In order to generate traffic these websites, the Macedonian teenagers published sensationalised and false content that appealed to Trump supporters (which they found generated more advertising revenue than left-leaning content). Many of them also said that running fake news websites was an easy way to make money; one teenager said that he made almost $16,000 off his two pro-Trump websites, when the average monthly salary in Macedonia is only $371 (Silverman & Alexander, 2016; Subramanian, 2017).

Organisations may also have vested interests to produce and propagate false information. For example, Professor Robert Proctor,

a science historian from Stanford University looked into how tobacco companies attempted to spread confusion about whether smoking indeed caused cancer among their consumers. Proctor's research found that the tobacco industry spent billions of dollars obscuring the truth behind the health effects of smoking, and purposefully propagated ignorance among their consumers. The tobacco industry tried to dismiss scientific evidence that linked smoking to cancer, saying that the experiments were done on mice and the same effects may not occur in humans (Kenyon, 2016; Rose & Barros, 2017). The tobacco industry also attempted to hire medical historians to make it seem as though experts disagreed on the issue of whether smoking caused cancer (Proctor, 2004).

Another study also found a relationship between environmental skepticism and conservative think tanks. Jacques, Dunlap and Freeman (2008) analysed 141 English-language environmentally-skeptical books published between 1972 and 2005, and found that over 92 percent of these books were linked to conservative think tanks. They also found that 90 percent of conservative think tanks interested in environmental issues adopted environmentally-skeptical positions. The researchers concluded that promoting environmental skepticism was a coordinated strategy by conservative think tanks to undermine scientific evidence that supports environmental movements.

Internet Trolls, Hate Groups and Conspiracy Theorists

In a Data and Society report titled *Media Manipulation and Disinformation Online*, Marwick and Lewis (2017) identified various Internet subcultures – trolls, hate groups and conspiracy theorists – that manipulate information on the Internet and in the social media ecosystem for their gains.

Internet trolls refer to people who deliberately bait other users to elicit an emotional response, often finding humour through sowing discord. They can either manipulate information online for mischievous behaviours without an intent to cause distress, or to ruin

the reputation of others. Hate groups like Men's Rights groups also manipulate online information to portray women as emotional and inferior, and men as logical and superior. Lastly, conspiracy theorists tend to produce their own amateur-made conspiracy "documentaries" and upload them onto websites like YouTube and forums. This allows other users to dissect these videos and form conspiracy theories that align with their worldviews.

While these subcultures might have their own agendas, they often use similar tactics like producing false information to achieve their goals.

6

SOLUTIONS AND POLICIES

❖ Fact checking is important in combating false information, but can be insufficient due to its limitations.

❖ Inoculation theory offers a promising approach to "psychologically vaccinate" individuals to confer protection against false information.

❖ Solutions against false information must also involve improving quality and trust in journalism, as well as improving news literacy of the general public.

❖ There is general consensus among experts that combating false information through regulatory approaches are impractical, in addition to infringing freedom of expression.

As presented in the earlier sections, the persistence of false information in society is a complex problem. People have individual-cognitive biases that make them susceptible to believing false information, and the unique characteristics of the Internet and social media amplify and exacerbate the negative effects of these biases. Correspondingly, the solutions and approaches to combat the problem must be multi-pronged.

This section focuses on the near-, medium-, and long-term solutions to tackle the persistence of false information in society.

When false information goes viral, fact checking and debunking are essential but reactive strategies to correct people's misperceptions. This is because fact checking tends to lag behind the spread of false information, and damage is often done before any debunking can come into effect. Also, as reviewed in *Section 3.3: The Backfire Effect*, corrective information that conflict with people's worldviews may backfire, thus further limiting the effectiveness of fact checking.

Fortunately, studies on inoculation theory (which draws on the idea of "psychological vaccination") have found promising results to suggest that people's resistance to believing false information can be built up over time. Other medium-term solutions to combat the false information include leveraging technology and machines, and re-building people's trust in quality journalism.

Finally, long-term strategies to combat the persistence of false information will involve improving news literacy of the wider public.

This section also looks at how governments around the world have responded to the increasing problem of false information, and the response from different stakeholders such as the non-governmental sector and academia.

6.1

Near-term Measures:

Fact Checking and Debunking

❖ The official statement of a reputable organisation (e.g. governments) influences people's perception of the accuracy of a rumour.

❖ Stories debunked by two or more fact checkers will appear on Facebook with a "Disputed" tag. Facebook also prompts users about the article's accuracy if they insist on sharing "disputed" articles.

❖ Google revised its algorithms to surface more authoritative pages and demote low quality content.

❖ News organisations and civil society have also stepped up their fact checking efforts to combat false information.

❖ Negative fact checking that debunk inaccuracies are more powerful than positive fact checking that reinforce the message.

The idea behind fact checking was first introduced in 1989 by *The Washington Post* political writer, David Broder, when he challenged his fellow journalists to take a more active role in monitoring the accuracy of televised political advertisements (Broder, 1989). This eventually led to the first incarnation of fact checking efforts known as the "ad watch", which was designed to help voters evaluate how truthful claims made by politicians in political advertisements were during the 1990 US Presidential Election (O'Sullivan & Geiger, 1995). Since then, fact checking websites such as *FactCheck.org* (by *PolitiFact*), *The Washington Post*'s *Fact Checker*, and *Snopes* have emerged, and fact checking quickly became an essential feature of modern political campaigns. The 2012 US Presidential Election was said to be one of "*the most fact checked election in history*" (Carr, 2012). Fact checker *PolitiFact* reported that it produced more than 800 fact checks on the presidential campaign, and that traffic to its website reached about 1 million page views on some days (Adair, 2012).

The demand for fact checking was given a further boost during the 2016 US Presidential Election, largely due to Donald Trump. *The Washington Post*'s *Fact Checker* reported that unique visitors to their site was five times higher in 2016 than in 2012, and *PolitiFact* reported that their fact check page on Trump reached about 2.7 million page views (Stelter, 2016). The surprising results of the Brexit Referendum and the 2016 US Presidential Election also contributed to casting the spotlight on fake news, leading various sectors of society – public, private, and people – to ramp up their fact-checking efforts to combat the spread of false information.

Public: Government-led Fact Checking

In certain countries, governments have taken the lead to provide fact checking.

One example is Singapore, where the Ministry of Communications and Information launched *Factually* in 2012. *Factually* aims to dispel and clarify false information that has gained sufficient public attention. *Factually* is part of the government's efforts in ensuring that Singaporeans possess accurate information on critical issues. Minister for Communications and Information Yaacob Ibrahim said that the government conducts periodic public engagement exercises to identify issues that are widely misconceived by the public or incorrect assertions that can harm Singapore's social fabric (Lee, 2017). More recently, *Factually* debunked WhatsApp rumours that claimed that Singaporeans' CPF savings will be transferred to their nominee's Medisave account by default. Earlier in April, the DSO National Laboratories also announced that it is developing an artificial intelligence system that can determine the authenticity of new stories within two hours using crowd sourcing, content analysis and source profiling (Cheng, 2017).

In an effort similar to Singapore's, the Malaysian Communications and Multimedia Commission also launched a fact checking website called *Sebenarnya* (which means "in actual fact") to fight the spread of false information. The Multimedia and Communications Minister said that such a move was necessary because *"Malaysians had the habit of spreading information without verifying the news"* (Kaos Jr., 2017).

Another example of a government-led fact checking effort is found in Czech Republic, where the Czech government set up a specialist "anti-fake news" unit to fight Russian disinformation campaigns. The unit scrutinises disinformation and debunks it via a dedicated Twitter account (@CTHH_MV), as well as via a new section on the interior ministry website (Tait, 2016) (also see *Section 6.4: Policies and Responses*).

Research has shown that the official statement of a reputable person (or in this instance a reputable organisation) influences people's perception of the accuracy of a rumour. A study by Allport and Postman (1947) found that after President Franklin Roosevelt gave an official statement that allayed rumours about the losses at the 1941 Pearl Harbour bombing, fewer students believed in the rumours about losses being greater than the officially stated numbers.

This suggests that government-led fact checking has the potential to play a key role in fighting the persistence of false information in society. However, there is also a danger when government-led fact checkers show a clear pro-government bias which can lead to a rejection alternative views, as seen in the example of *Fact Checking Turkey*, where almost every fact check seems to support the Turkish government (Jackson, 2017b).

Private: Industry-led Fact Checking

The private sector – comprising technology companies and news organisations – has also stepped up in its fact checking efforts, often in collaboration with other players.

Post 2016 US Presidential Election, Facebook faced widespread criticisms that the proliferation of false information on the platform (via its "Trending Topics" algorithm) may have helped Donald Trump get elected. In response, Facebook partnered third party fact checking organisations (signatories of Poynter's International Fact Checking Code of Principles) such as *PolitiFact* and *Snopes* to curb the spread of false news and misinformation on the platform (Mosseri, 2017). If two or more fact checkers debunk an article, the article will appear with a "Disputed by 3rd Party Fact-Checkers" tag on users' Facebook interface. Third party fact checkers can also link to their own articles explaining why the article was debunked. Facebook will also prompt users about the article's accuracy if they still insist on sharing the article, and penalise disputed articles such that they will appear lower in users' news feeds. In addition, Facebook has also improved its interface to make it easier for users to report a

fake news story, emphasising the importance of leveraging community efforts to fight the spread of false information (Mosseri, 2016).

However, a report by *The Guardian* raised questions about the effectiveness of Facebook's fact checking efforts. *The Guardian* said that Facebook's "disputed" labels were regularly ineffective and seem to have minimal impact. It said that articles that were debunked by Facebook's fact checking partners frequently remain on the platform without a "disputed" label, and that the label often comes after the story has gone viral, which suggests that damage had already been done (Levin, 2017).

Google, another technology giant, has also rolled out "fact check" labels in Google News. These labels help users identify articles that contain information that has been verified by fact checking organisations. Google has since extended this feature to Google Search, where it presents richer information from fact checking websites like *PolitiFact* and *Snopes* when they show up in users' searches (Kosslyn & Yu, 2017). Google believes that making fact checks more visible will help users better assess information that they encounter and make informed opinions. Google has also revised its algorithms to *"surface more authoritative pages and demote low quality content"* in order to improve the quality and reliability of its search results (Hern, 2017a). Lastly, Google Chrome extensions that perform fact checking functions have also been developed. One example is *RealDonaldContext* which is developed by *The Washington Post*'s Fix team to add more context or corrections to Donald Trump's tweets (Bump, 2016). Google News Lab also collaborated with non-profit organisation First Draft to develop *Cross Check*, which is an international fact checking project that aims to investigate false information during the recent French election (Loughran, 2017).

News organisations themselves have also stepped up their fact checking efforts. For example, BBC started its own fact checking service called *Reality Check*. According to the BBC, its *Reality Check* team will focus on content that is clearly fabricated and misleads the

public into thinking that the content is produced by a reputable news organisation. It also aims to use a variety of formats – online, television and radio – to ensure that facts reach out to a wider audience than false information. *Reality Check*'s focus on false information is part of BBC's "slow news" initiative, where it aims to use in-depth analysis to provide the public with more context on important issues (Jackson, 2017a). *Le Monde*'s (one of the biggest French newspapers) fact checking unit also developed a web extension called *Decodex*, which sends users a pop-up warning that says "warning this is a fake news site" when users stumble across one. *Decodex* is linked to *Le Monde*'s database which contains a compilation of real, fake and satirical sites (Davies, 2017a; Wendling, 2017).

Other fact checking initiatives leverage technology to debunk false information. One example is Storyful, a company founded by Irish broadcast journalist Mark Little, which aims to bring technology to newsrooms to improve the process of verifying content. Storyful also released *Verify*, a free Google Chrome add-on that tells users whether the videos they are watching have been verified by their team. *Verify* uses automated bots to comb through Twitter and YouTube for videos which are high in popularity, and incorporates the Storyful team's own reporting so that users know whether the video has been vetted. So far, Storyful has vetted 250,000 videos (Herbert, 2017).

Another example is Full Fact's (an independent fact checking organisation in the UK) automated end-to-end fact checking initiative, which aims to inform readers if a claim has previously been reported as inaccurate, and automatically fact check claims in real time using natural language processing and statistical analysis (Babakar & Moy, 2016). Full Fact has also been given a grant from Google's *Digital News Initiative* scheme, which aims to help groups create new tools that could make it easier for journalists to do their jobs (Burgess, 2016). A similar initiative is *ClaimBuster*, an Australian project that also uses natural language processing to help journalists identify factual claims within a text for subsequent verification (Robbins, 2017).

People: Civil Society-led Fact Checking

The spread of false information has also become a problem for Indonesia, where false information often riles up racial tensions and anti-Chinese sentiments. While the Indonesian government has responded to this (see *Section 6.4: Policies and Responses*), a significant pushback also came from civil society.

Turn Back Hoax is a crowdsourcing-based app created to curate false information and hoaxes circulating on the Internet and social media. It was developed as a Google Chrome extension that allows users to flag websites, chain messages and pictures they encounter as either "hoax" or "not hoax". Users' inputs are collated on the *Turn Back Hoax* website where users can comment and discuss. This allows users to use *Turn Back Hoax* as a reference to verify whether information they encounter online is accurate or not. The database of false information collected by *Turn Back Hoax* can also be used to identify patterns of the spread of false information ("New app allows public", 2016). Another example of a ground-up debunking initiative in Indonesia, the *Anti-Hoax Indonesian Citizens* movement, was started by social media activist Septiaji Eko Nugroho. The movement merges four social media communities (*Forum Anti Fitnah, Hasut and Hoax*, *Indonesian Hoax Buster*, *Indonesian Hoaxes*, and *Facebook group Sekoci*) that are particularly active in countering the spread of false information online (Sapiie, 2016).

Another example of a grassroots fact checking effort comes from Lithuania. Citizens in Lithuania have come together to fight against Russian disinformation, branding themselves as "elves" to take on Internet "trolls" who spread disinformation that could destabilise the nation. These "elves" coordinate their efforts through Facebook and Skype; they police social media platforms to counter Russian disinformation by exposing fake accounts and writing articles that debunk these falsehoods. Their efforts have been recognised by the Lithuanian government and the North Atlantic Treaty Organisation (NATO) (Au-Yong, 2017a; Dapkus, 2016; Weiss, 2016).

Effectiveness of Fact Checking

Research has found mixed findings regarding the effectiveness of fact checking.

Fridkin *et al.* (2015) found that exposing people to fact checking that assert the truthfulness of a political advertisement resulted in positive assessments of the advertisement's accuracy, while exposing people to fact checking that debunk the truthfulness of the advertisement resulted in poorer ratings of the advertisement's accuracy. This suggests that fact checking is indeed effective at influencing people's views about the accuracy of political advertisements. The researchers also found that negative fact checking that debunk inaccuracies in the advertisement were more powerful than positive fact checking that reinforce the message in the advertisement (Fridkin *et al*., 2015). Min (2002) also found similar results, and further demonstrated that the tone of an ad watch analysis (i.e. fact check or corrective information) influenced individuals' assessment of the political advertisement. Lastly, Ecker *et al.* (2011) found that repeating retractions (i.e. corrective information) alleviates the effects of misinformation, even though it does not eliminate it.

However, research has also demonstrated the limitations of fact checking and corrective information. For example, Nyhan and Reifler (2010) found that corrective information generally does not change people's beliefs, especially when the new information conflicts with their preexisting beliefs. As mentioned in *Section 3.3: The Backfire Effect*, Nyhan and Reifler (2010) found that conservatives became even more likely to believe that Iraq possessed weapons of mass destruction after reading a correction news article which clarified that Iraq did not actually possess weapons of mass destruction prior to the US invasion of Iraq in 2003. In fact, many of Nyhan's and Reifler's work paint a dismal picture about the effectiveness of fact checking, largely due to the backfire effect (Nyhan & Reifler, 2012; Nyhan & Reifler, 2013; Nyhan & Reifler, 2015).

Furthermore, Lim (2017) evaluated the performance of fact checkers and found that there was a lack of agreement among major fact checkers, which may limit the effectiveness of fact checking as a whole. According to Lim, two main factors contribute to the effectiveness of fact checking as a whole – 1) overlap in claims that were evaluated by fact checkers and 2) high degree of agreement in the outcome of evaluated claims among major fact checkers. However, after conducting an inter-reliability test between *PolitiFact* and *Fact Checker*, Lim (2017) found that the two major fact checkers rarely verify the same statement during their fact checking. And even when they do so, there is little agreement in the factual accuracy of the evaluated statement.

In short, while the increase in fact checking efforts is an essential component to combating the spread and persistence of false information in society, that in itself is insufficient due to the limits of fact checking and the prevalence of the backfire effect.

6.2

Medium-term Measures:

Inoculation, Leveraging Machines and Improving Journalism

❖ Inoculation theory suggests that individuals can be prepared for potential false information by exposing logical fallacies *a priori*.

❖ An inoculation message consists 1) a "threat" component that signals that one's position is susceptible to persuasion and 2) a "refutation" component that provides information one can mobilise to resist persuasion.

❖ Twitter identifies bot-based manipulation and shuts down bots on its platform Facebook also uses machine learning to identify and eliminate fake accounts.

❖ Studies show that "accuracy", "timeliness", and "clarity" are the top three factors people use to evaluate the trustworthiness of a news organisation.

"Pre-bunking" through Inoculation

As highlighted in the earlier subsection, the belief in false information is often resistant to correction for a variety reasons, especially if the correction challenges a person's worldview, which may cause the correction to backfire.

One promising alternative solution to this problem is derived from the *inoculation theory*, which prepares individuals for potential false information by exposing logical fallacies *a priori*. The theory is a biological metaphor first suggested by McGuire (1961), who proposed that an individual's beliefs could be inoculated against persuasive attacks the same way our immune system could be immunised against viruses.

Similar to how vaccines work – by injecting a weakened form of a virus into the body so that the body produces antibodies that will protect itself from a stronger form of the virus in future – McGuire (1961) argued that exposing individuals to information containing weakened arguments (even though it may conflict with individual's preexisting beliefs) can help individuals develop resistance against more persuasive attacks in future. Hence, this application of inoculation theory functions like a "psychological vaccination", where individuals may be inoculated against believing false information.

Inoculation messages involve two main components – 1) an explicit warning about an impending threat, and 2) a refutation of a pre-empted argument. The "threat" component signals to the individual that his or her position on an issue is susceptible to persuasion and

change, and the "refutation" component provides information that they can mobilise to strengthen their attitudes and resist the persuasion (Compton, Jackson & Dimmock, 2016). For example, an inoculation message might include – 1) explicit warning that there are attempts to cast doubt about the scientific consensus on climate change, and 2) that one of the methods of doing so is by referencing fake experts to feign a lack of consensus. Thus, the false information (i.e. lack of consensus on climate change) is being delivered to individuals in a weakened form (van der Linden *et al.*, 2017).

In a similar area, research in consumer psychology also examined the effects of one-sided and two-sided non-refutational and two-sided refutational messages on consumer response (Kamins & Assael, 1987; Rucker, Petty & Brinol, 2008). O*ne-sided appeal* presents only claims that are supportive of the product or brand, while *two-sided non-refutational appeal* includes the limits of product or brand performance on attributes of minor significance to the consumer, in addition to presenting positive claims. A *two-sided refutational appeal* is one that incorporates refutations or counter-arguments to the presented limitations. According to the inoculation theory, a two-sided refutational appeal will lead to greater acceptance of the communicator's position than one-sided and two-sided non-refutational appeals. Generally, two-sided refutational appeals were more effective in increasing perceived source truthfulness and believability.

Research has demonstrated the effectiveness and potential of inoculation theory at combating false information. In a study by van der Linden *et al.* (2017), researchers tried to understand if it was possible to inoculate people's beliefs about the level of scientific consensus against misinformation. The study involved three different messages – 1) a consensus message which communicated scientific consensus on climate change, 2) a counter-message which communicated no consensus on climate change, and 3) an inoculation message which warned individuals that some politically and financially motivated groups were trying to convince the public that there was no scientific consensus on climate change. The study found that presenting subjects with the consensus message alone

resulted in a positive influence of perceived scientific agreement, whereas presenting the counter-message alone resulted in a negative influence of perceived scientific agreement. The study also found that presenting subjects with the inoculation message before the counter-message preserved much of the positive influence of the consensus message. Furthermore, detailed inoculation messages worked better than general inoculation messages. Another similar study by Cook, Lewandowsky and Ecker (2017) also found that inoculating messages were effective in neutralising the adverse effects of misinformation on the level of scientific consensus on global warming.

In short, the findings from the above studies suggest that inoculation is a promising approach to protect the public from false information, and that providing individuals with *"the 'cognitive repertoire' to preemptively refute"* false information is essential (van der Linden et al., 2017: 6).

Leveraging Machines

As mentioned in *Section 4.2: Filter Bubbles and Echo Chambers* and *Section 4.3: Bots and False Amplifiers*, many of the algorithms and artificial intelligence used on social media platforms play a role in contributing to the spread of false information. However, efforts to combat false information can also leverage machines to fight machines.

For example, research estimated that about 9 to 15 percent of active Twitter accounts are bots (Varol *et al.*, 2017), which allows certain voices to be disproportionately amplified on the platform. In response to this problem, Twitter said that its spam-detection team identifies bot-based manipulation, and also improves on tools that can be used to spot and shut down bots on its platform (Manjoo, 2017). Facebook has also used analytical techniques such as machine learning to identify and eliminate fake accounts as part of its efforts to fight against false amplifiers. As of mid-April, Facebook has managed to take action against 30,000 fake accounts in France alone (Weedon, Nuland & Stamos, 2017).

There are also digital tools that help people "escape their filter bubble/echo chamber". For example, Google Chrome extension *PolitiEcho* produces a visualisation of users' Facebook political bias based on the pages that users' friends have "liked". Another example is *FlipFeed*, which is a Twitter plug-in that allows users to replace their Twitter feed with that of a random and anonymous user with a different political worldview. *Escape Your Bubble* is a plug-in that seeds users' Facebook news feed with stories of opposing political views (Hess, 2017). Lastly, *The Times* launched a Facebook messenger bot called *Filter-bubble Buster*, which provided people with balanced information leading up to the 2017 UK elections (Davies, 2017b).

As mentioned in *Section 6.1: Near-term Measures: Fact Checking and Debunking*, fact checking efforts can often be ineffective and/or counter-productive. With this research problem in mind, Shao *et al.* (2016) developed *Hoaxy*, a platform designed to visualise the competition dynamics between the spread of false information and fact checks online. *Hoaxy* can reconstruct the diffusion networks of hoaxes and their corrections as they spread from person to person online, thus enabling researchers, journalists, and the public to understand the factors behind successful mitigation of false information online. For example, *Hoaxy* has found that the sharing of fact checks generally lags behinds the spread of false information by 10 to 20 hours, and that the sharing of false information is dominated by very active social media users while fact checking is a more grassroots activity.

Research has also shown encouraging results at developing technologies for fake news detection. Conroy, Rubin and Chen (2015) developed an innovative hybrid approach that combines linguistic cue and machine learning with network-based behavioural data, which they argue provides a basis for the development of a comprehensive fake news detection tool. Ruchansky, Seo and Liu (2017) proposed a model that combines three characteristics of fake news – the text, the response received, and the source users promote – to allow a more accurate and automated prediction of fake news.

Improving Quality and Trust in Journalism

Good journalism also plays a huge role in fighting the persistence of false information. Ensuring quality journalism and upholding journalism ethics helps re-build the trust in journalism, thus dis-incentivising people to seek less robust sources of information to validate their worldviews.

Various journalists and media experts have voiced their opinions on how journalists can and should play a role in re-building trust in the journalism. In an interview with *The Atlantic*, managing editor of *Snopes* Brooke Binkowski said that ensuring quality journalism (and thus quality content) is crucial because that will make it easier for people to find *"vetted, nuanced, contextual and in-depth information"* (Meyer, 2017).

Speaking at a forum on fake news, associate opinion editor of *The Straits Times* Lydia Lim said that journalists also need to help the public understand what the good media values practised by traditional media are so that people realise why it is important to consult verified sources for accurate information (Au-Yong, 2017a). Writer and editor of *The New York Times* Serge Schmemann said that *"reporters should return to the fundamentals of journalism, step outside the comfort of their newsrooms and privileged lives and speak to different audiences"* (Bouteldja, 2017). Professor of Communications and New Media at the National University of Singapore Mohan Dutta also emphasised on the importance of journalism ethics and ethical communication in fighting fake news (Dutta, 2017).

Research has also shed light on what influences people's trust in the news. A study conducted by The Media Insight Project (2016) showed that trust and reliability in the news can be broken down into specific factors which news organisations can work on to improve their trustworthiness among consumers. The study found that accuracy was the number one factor behind trustworthiness. 85

percent of survey respondents said that it is extremely or very important that news organisations get the facts right. The second-most valued factor was timeliness, with 76 percent of respondents saying that news reports should be up-to-date with the latest information. The third-most valued factor was clarity. 72 percent of respondents said that news reports should be concise and to the point ("A New Understanding", 2016). Previous research on developing a multi-dimensional scale that measures news trustworthiness also found that trust in news media depends on factors such as accuracy of depictions, and the selection of topics and facts (i.e. clarity and contexualisation) (Kohring & Matthes, 2007). Thus, these are specific areas that journalists and news organisations can work on to re-build people's trust in journalism.

Technology companies are also playing their part to contribute to the re-building of trust in journalism. For example, Facebook launched the *Facebook Journalism Project* in early 2017, which aims to work with journalists and newsrooms to ensure that good and healthy journalism thrives. As part of the project, Facebook works with news organisations to develop new products like new storytelling formats to engage consumers and suit evolving information consumption preferences. Facebook also offers training and tools for journalists – one example is the *Innovative Journalism Certificate*, which is a three-course certificate program launched in collaboration with The Poynter Institute that aims to help journalists make better use of Facebook for news gathering and audience engagement ("The Poynter Institute and", 2017). Another example is *Crowdtangle*, which is a tool that helps publishers track how their content spreads online (Simo, 2017)

New models of journalism are also being experimented with to see if they can strengthen the relationship between journalists and the wider public. One such example is *Wikitribune*, started by Wikipedia co-founder Jimmy Wales. *Wikitribune* aims to pair a small team of professional journalists with volunteer citizen journalists to write and edit news stories as they develop, alongside a community of readers who will fact check and sub edit published articles. Wikitribune also plans to fund the project through crowdfunding campaigns.

According to Wales, the absence of information gatekeeping, the shift from editorial control to community control, and increased transparency will help bring readers' trust back in news organisations and journalism (Harb, 2017; Hern, 2017b).

6.3

Long-term Measures:

Improving News Literacy

❖ Studies have demonstrated a general lack of news literacy among the public, suggesting the need to improve news literacy as a long-term strategy to combat false information.

❖ Research has shown that college students who took a news literacy course demonstrated significantly higher levels of news literacy as compared to students who did not.

❖ While promoting news literacy is often associated with the young, promoting news literacy among adults and the elderly are equally important.

Many studies have pointed to a general lack of news literacy among the public. According to a survey done by Pew Research Center, 64 percent of US adults said that *"fabricated news stories caused a great deal of confusion about the basic facts of current issues and events"* (Barthell, Mitchell & Holcomb, 2016). Ipsos Public Affairs conducted a survey for *BuzzFeed News* and found the same trend (Silverman & Singer-Vine, 2016). Another study conducted by Stanford History Education Group found dismaying results regarding middle-school, high-school and college students' ability to assess online sources of information despite being digitally savvy. The study found that students were not able to distinguish fake accounts from real ones, differentiate information from activist groups versus information from neutral sources, and tell apart advertisements from articles (Donald, 2016).

In the UK, a YouGov survey commissioned by Channel 4 found that only 4 percent of people could correctly identify fake news ("C4 study reveals only", 2017). Lastly, a government survey in Singapore also found that two in three Singaporeans could not discern fake news when they see it, and that a quarter of survey respondents shared information that they later discovered to be false (Ng, 2017). Thus, improving news literacy among the general public is another important strategy in combating the persistence of false information in society.

Various efforts have been done to promote news literacy and reduce people's susceptibility to believing false information.

For example, Facebook launched the *News Integrity Initiative*, which is a $14 million collaborative project that aims to improve news literacy of the general public so they can make informed decisions

about what they read and share online. The initiative is funded by Facebook and various non-profit organisations, and research grants will be given to support applied research on news literacy. The initiative will also facilitate collaboration with industry experts to develop solutions (Murgia, 2017). Ultimately, the initiative aims to develop specific measures among news organisations and social media platforms to improve news literacy among the public (Bilton, 2017). Other efforts from Facebook to promote news literacy include publishing a list of tips to spot false news on the platform (Price, 2017), and running a full-page newspaper advertisement on spotting false news in France before the election (Dillet, 2017).

Schools have also started incorporating classes on news literacy into their curricula (Barron, 2017; Lapowsky, 2017; Madison, 2017). Journalism lecturer at the University of Hong Kong Anne Kruger argued that news literacy classes should be made mandatory in order to successfully equip the young with the skills to verify facts and falsehoods (Au-Yong, 2017a). At the university's Cyber News Verification Lab, Kruger teaches her students to apply news literacy skills learnt in class to debunk false information that is being circulated. Kruger said that there was a noticeable difference in the way students checked and verified information they encountered online before and after the class ("New JMSC project focuses", 2017).

In fact, research has also shown that college students who took a news literacy course demonstrated significantly higher levels of news literacy as compared to students who did not. After conducting a survey among Stony Brook University students, researchers found that students who took a news literacy course scored better on a New Media Literacy Scale than students who did not take the course (Maksl *et al.*, 2017).

While promoting news literacy is often associated with the young, promoting news literacy among adults and the elderly are equally important. Based on analysis done using survey data collected by the UK Office of Communications, both younger and older people have lower levels of critical digital literacy compared to people aged 30 to

60, suggesting that literacy education should not just be for the young (Livingstone & Olafsson, n.d.).

However, some experts have also raised concerns about how news literacy is being promoted in response to concerns about false information. Emerson College Professor Paul Mihailidis argued that one common misconception is to promote media literacy as a tool or an educational curriculum that can solve the problem of susceptibility to false information. Media literacy tends to involve teaching individuals to interrogate the credibility of the information in order to differentiate the real news from Instead, Mihailidis and Viotty (2017) argued that initiatives that promote media literacy must be repositioned to respond to an era of partisanship and distrust, and not jump to solutions before people have understood the problem. Assistant Professor of Journalism at the University of Hong Kong Masato Kajimoto also supports the idea of repositioning news literacy curricula to the larger socio-political contexts. After comparing patterns of social media usage among students in three different Asian countries (Hong Kong, Vietnam and Myanmmar), Kajimoto's research concluded that region-specific issues need to be pedagogically integrated into the development of news literacy curricula (Fleming & Kajimoto, 2016).

6.4

Policies and Responses

❖ Experts discourage the introduction of new laws because the problem of fake news is due to a failure of enforcing existing laws and social responsibilities.

❖ Regulation would increase the financial and manpower burden of traditional media but not fake news websites, thus further tipping the balance in favour of the latter.

❖ Enforcing social media platforms to self-regulate will cause them to unnecessarily delete content that may not be illegal to avoid the prospect of punishments, thus undermining freedom of expression.

❖ Other non-regulatory solutions proposed include setting up people's tribunals to judge the accuracy of news articles, and strengthening existing self-regulatory frameworks instead of introducing statutory regulation.

The responses towards prospective regulation of false information have been mixed. Some countries have shown support for regulation of false information. For example, a recent survey done by the Singapore government found that 91 percent of Singaporeans are supportive of stronger laws to ensure the removal and correction of fake news (Ng, 2017). Earlier in April, *The Straits Times* also reported that some experts who track the issue of fake news welcome the government's plan to look into dealing with the problem (Au-Yong, 2017b).

However, governments' inclination to regulate social media in an attempt to fight false information has engendered backlash from various sectors of society. Censorship and the infringement of freedom of expression are among the main issues raised. Representative on the Freedom of the Media for the Organisation for Security and Co-operation in Europe Dunja Mijatovic said that societies should not overreact to fake news by having more regulations, and that censorship can often lead to self-censorship. Instead, the problem of fake news should be addressed with education, literacy, and upholding journalism ethics rather than imposing regulations and restrictions (Dutta, 2017; Mijatovic, 2016).

<u>United Kingdom</u>

In January 2017, the UK House of Commons Culture, Media and Sport Committee launched an inquiry into fake news. The inquiry was launched in response to the public's growing distrust of traditional sources of news and their shift towards the Internet and social media for information, and also to understand whether fake news has had a significant impact on democratic processes. The scope of the inquiry included:

1. What is fake news?

2. What is the impact of fake news on traditional journalism?

3. Are there differences in the way people of different ages, social backgrounds, genders et cetera respond to fake news?

4. How has advertising encouraged the growth of fake news?

5. What are the social responsibilities of technology companies in fighting fake news?

6. How can the public be educated on news literacy?

7. Are there differences between the UK and other countries in the degree to which people respond to fake news?

8. How have other governments responded to fake news?

The inquiry started on 30th January 2017 and closed on 3rd March 2017. A total of 79 pieces of written evidence were submitted. The submissions came from various stakeholders such as news organisations (e.g. BBC and The Guardian), academia (e.g. King's College London and University of Bristol), technology companies (e.g. Facebook and Google), non-profit organisations (e.g. FullFact and WikiMedia U.K.), and individuals. The full list of written evidences can be found here.

One of the organisations that responded is the News Media Association (NMA). NMA represents 1,100 newspaper titles and functions to promote their members' interests to the government.

Regarding the issue of government regulation, NMA said that they strongly oppose the introduction of any new laws, because the growing problem of fake news is a result of the failure of enforcing existing laws and social responsibilities.

For example, fake news websites have failed to observe copyright and defamation laws, and Facebook and Google have failed to live up to the social responsibilities of information dissemination. Furthermore, regulation would only increase the financial and manpower burden of traditional media but not fake news websites, thus further tipping the balance in favour of the latter. NMA also highlighted private sector-led initiatives like *Decodex* and *CrossCheck* (see *Section 6.1: Debunking: Fact-Checking*) as the most benign of approaches, and that the setting up of government agencies for regulation would easily lead to censorship ("Culture, Media and Sport Select", 2017).

Germany

In March 2017, Germany's Justice Minister Heiko Maas proposed a legislation that would fine social media companies like Facebook, YouTube and Twitter if they fail to remove online content deemed as hate speech or fake news within 24 hours. Known as the Network Enforcement Act (Netzwerkdurchsetzungsgesetz in German), the proposal was approved by the Cabinet in April, but have yet to come into force.

A failure to comply with the new legislation can result in a fine of up to €50 million on the company. The legislation would also oblige social media companies to produce quarterly reports on how many complaints are received and resolved, and how much manpower is deployed in complaint management (Chazan, 2017). Back in 2015, Facebook and Twitter took part in a voluntary commitment to delete criminal content from their platforms within 24 hours, but research by the government found that Facebook deleted 46 percent of such content, while Twitter deleted only 1 percent. Maas said that this was

also one of the reasons why increased pressure on social media companies is necessary (Southern, 2017).

Facebook responded to the Network Enforcement Act and said that the draft law is not appropriate to fight hate speech and false news. It also said that because draft law will impose on it a disproportionate threat of fines, it may encourage the deletion of content even when they are not clearly illegal. This effectively transfers the responsibility of making complex legal decisions from public authorities to private companies. Facebook added that it has also consulted with legal experts, who said that the draft law goes against German constitution and is non-compliant with European Union law (Shead, 2017). CEO of YouTube Susan Wojcicki also agreed that the legislation will put unnecessary pressure on social media companies, which could lead to unwarranted deletion of content on their platforms (Schaer, 2017).

Academics have also voiced concerns about the draft law. Dr. Stefan Heumann of the technology and Internet think tank Stiftung Neue Verantwortung said that platforms would unnecessarily delete content that may not be illegal to avoid the prospect of fines, which will ultimately undermine freedom of speech (Chazan, 2017). In fact, the chilling effect on freedom of expression is already felt as Twitter has started blocking accounts that suggest hate speech even of the slightest bit (Bittner, 2017). Professor of Media Studies at Humboldt University Wolfgang Mühl-Benninghaus doubted that fake news would have much impact on Germany's upcoming elections because Germany has much experience in dealing with fake news throughout their history, as seen during the Weimar Republic and the Third Reich (Toor, 2017).

Other critics like the German Journalists Association and the Global Network Initiative believe that Maas is pushing for the legislation to gain political advantage in the upcoming elections (Schaer, 2017). Chairman of the Association of German Magazine Publishers, Stephan Scherzer, said that the legislation would turn social media companies into "private opinion police" (Faiola & Kirchner, 2017).

Alternative solutions have also been proposed. Julia Krueger, media consultant at the Amadeu Antonio Foundation which fights

extremism and hate speech in Germany, said that targeting advertisements that drive the spread of fake news is a better solution as it does not imply censorship like the draft law does (Schaer, 2017). Professor of Media Law at the Technical University of Dortmund Tobias Gostomzyk said that having independent bodies acting as fact checkers would be more effective than using criminal law ("500,000 euros fine for", 2016).

Italy

In December 2016, Italy's Justice Minister Andrea Orlanda hinted at the possibility of taking legislative action against fake news, saying that politicians should play a role in "*disincentivising the affirmation of post truth*" (Mantzarlis, 2017). A few days later, Italy's anti-trust chief Giovanni Pitruzzella called for European Union member states to set up a network of public agencies (modelled on a system of anti-trust agencies) that are dedicated to spotting and removing fake news. Pitruzzella said that the regulation of fake news on the Internet was best done by the state than by social media companies, and that it would be inappropriate for social media companies to self-regulate because the job of controlling information is historically that of public powers and not of private ones. He also added that currently, the only way to deal with fake news in Italy is through the judicial system (Politi, 2016). Laura Boldrini, speaker of the lower house of Parliament, also said that fake news is a critical issue that has to be dealt with immediately. Boldrini met with Facebook in December to look into ways the platform could regulate hate speech and fake news (Horowitz, 2016).

Experts have expressed their concerns about the government's plans to deal with fake news. Professor at Bocconi University's School of Management Carlo Alberto Carnevale Maffé said that such proposals were illiberal and impractical. Fellow at the Nexa Center for Internet and Society Fabio Chiusi also said that the proposed regulations may result in censorship, and that truth and democracy would be undermined without the space for having well informed debates (Mantzarlis, 2017). Founder of opposition party Five Star Movement

proposed an alternative solution that involves setting up people's tribunals that would judge the accuracy of news articles.

However, some also questioned the viability of this alternative solution. News presenter at LA7 (an Italian private television channel) said that a people's tribunal is unworkable because it would at best give the kind of "audience juries" seen in song contests ("Italy's fake news battle", 2017). Managing editor of an online publication *Il Post* Luca Sofri said that the Italian media is poorly equipped to deal with fake news because the media culture is one that does not prioritise accuracy and reliability (Mantzarlis, 2017).

Czech Republic

The Czech government has set up a specialist anti-fake news unit (named Centre Against Terrorism and Hybrid Threats) that will combat falsehoods (largely about migrants) spread by Russian disinformation campaigns. The unit aims to combat Russian interference in the upcoming elections in October 2017. There are currently 20 specialists working full-time in the unit, whose job is to scrutinise disinformation and debunk it via a dedicated Twitter account (@CTHH_MV) as well as via a new section on the interior ministry website (Tait, 2016).

The setting up of this anti-fake news unit is not without criticisms, largely drawing accusations that it will lead to censorship, surveillance and an erosion of free speech. Czech President Milos Zeman (who is known to defend Russian foreign policy) also criticised the unit, saying that nobody has a monopoly on the truth (Cameron, 2017). In response to allegations on censorship, the specialist unit said that they were not in the business of closing down websites but instead trying to show very clearly what is false ("New Czech government unit", 2017).

South Africa

South Africa is also looking into the regulation of social media in light of the spread of fake news. In March 2017, South Africa's Safety and Security Minister David Mahlobo said at a press conference that the government is considering options to regulate social media to fight the spread of false information and misleading photoshopped images. Mahlobo added that regulation is the way to go, and that even the best of democracies regulate the space (Dolley, 2017).

Mahlobo's comments on regulating social media was seen as an attempt at censorship that would undermine freedom of speech, and sparked a #HandsOffSocialMedia protest hashtag on Twitter. Experts and observers have also voiced their concerns about prospective regulations. Media analyst and commentator Arthur Goldstuck said that regulation is the government's move on attempting to gain control over the speech in the online environment, and that regulation goes against freedom of expression which is enshrined in the Constitution. Mark Weinberg, from non-profit advocacy organisation Right2Know, said regulation implies that the government is trying to censor voices it does not want to hear. Political analyst Ralph Mathekga said that regulation would encourage the state to think that it can think on behalf of its citizens, which may put South Africa on the road to a dictatorship. He added that South Africans are discerning enough to recognise fake news when they see it. Director of the Freedom of Expression Institute Tusi Fokane proposed strengthening existing self-regulatory frameworks over introducing statutory regulation (Jadoo, 2017).

Indonesia

In response to the growing problem of fake news in Indonesia, the Indonesian government announced in January 2017 that it will be creating a cyber agency that combats fake news on social media deemed as slanderous, fake, misleading and hateful. Indonesia's security ministry will oversee the cyber news agency which will monitor news stories circulating online, as well as safeguard government institutions from malicious hackers. Indonesian Internet expert Nukman Luthfie expressed fears that the cyber agency would also be used to monitor public opinion, resulting in in a breach of privacy ("Indonesia to set up", 2017).

The Philippines

In February 2017, House Speaker Pantaleon Alvarez proposed the Social Media Regulation Act of 2017, which aims to regulate the use of social media by making social media companies responsible for ascertaining the veracity of the identity of a user. Individual users can also be held liable if they open a social media account that presents themselves as someone other than themselves (i.e. creating fake accounts). If passed, violators of the act can be punished with imprisonment between 6 to 12 years, and fined between P30,000 to P50,000. Alvarez said that fake accounts are dangerous because the anonymity encourages users to spread false information online (Luci, 2017).

Social media experts have voiced their concerns over the implications of this act. Social media strategist Tonyo Cruz said that the proposed act would restrict freedom of expression, and does not take into account how regulation would apply to Filipinos living and working abroad. Digital marketing specialist Carlo Ople said that it is almost impossible to regulate social media because many social media companies operate overseas and are not covered by Philippine law.

Furthermore, fake accounts are difficult to trace as they are often provide false personal information. Instead, Ople suggested that solutions should focus on cyber education and critical thinking instead of regulation (Cabato, 2017). Information and technology law expert JJ Disini also said that it is impossible for the Philippine government to regulate offshore social media platforms, and that regulation would make people who are critical of the government less likely to voice their views. Disini suggested that netizens themselves should take the initiative to discern the authenticity of an account by reviewing the profile ("The pros and cons", 2017).

South Korea

Policy Director Heo Kwangjun of Open Net Korea (a non-profit organisation for freedom and openness of South Korea's Internet) said at the *Straits Times-WAN-IFRA Keep It Real: Media Policy Workshop* that South Korea is also looking into possible regulations against fake news. South Korea's parliament is currently reviewing proposed laws that, if passed, would punish the production of fake news, oblige Internet providers to play a more active role in regulating fake news, and also make the deletion of any evidence of fake news production punishable.

However, no additional information on this could be found at the time of writing this literature review.

7

CONCLUSION

As evident from the literature review, fake news is part of an information ecology that contains a wide range of misinformation and disinformation. The term *fake news* used by industry (e.g. media practitioners) and academics adheres to a narrow definition. However, in public discourse, the term is used to refer to diverse types of false information.

As empirical work on fake news is still at a nascent stage, this review has incorporated research conducted in past decades on different types of false information, such as those on rumours, conspiracy theories and propaganda. Using different methodologies such as surveys, experiments and machine learning, these studies are instructive as they provide insights into the various factors that shape people's reception, resistance, and reactions to false information in different political and social contexts.

The research presented in the earlier sections sheds light on existing gaps pertaining to the study of fake news. The following, which are applicable not just to fake news but false information in general, are recommendations for further research relevant to Singapore:

1) Do Singaporeans respond differently to different types of false information? Are they more susceptible to certain types of false information over others and why? In other words, what types of false information do Singaporeans find more believable and why?

2) Linked to the above, what is the impact of different social media platforms on Singaporeans' susceptibility to fake news/false information? For instance, are fake news/false information

circulated on Instant Messaging platforms more believable than those circulated in blogs and social networking sites?

3) How do Singaporeans respond to fake news/false information? In other words, what actions do they adopt when they encounter fake news/false information?

4) What impact do people's demographics have on their responses to fake news/false information? Do race, age, education, socio-economic status affect how Singaporeans respond to and manage fake news/false information?

5) Linked to the above, which segments of Singaporeans are more prone to confirmation bias and motivated reasoning? Do demographics and group membership influence confirmation bias and motivated reasoning?

6) What impact do people's political traits (e.g. political orientation, political efficacy, and political knowledge) have on their susceptibility to fake news/false information? Do political traits affect how Singaporeans respond and manage fake news/false information?

7) What is the relationship between people's information diet and their susceptibility and response strategies when it comes to fake news/false information? Is there a relationship between the diversity of one's media consumption and one's response to fake news? (By information diet, we refer to the type and diversity of media one seeks information from, and how one engages with media content, passive versus active.)

8) What is the relationship between people's trust in media (mainstream and online) and their susceptibility and response strategies when it comes to fake news/false information?

9) Linked to the above, what is the relationship between people's trust in journalism and their susceptibility and response strategies when it comes to fake news/false information?

10) How effective are different communication sources (e.g. government, media, academia, non-governmental or non-profit organisations, family and friends) in debunking fake news/false information? How credible are these sources with Singaporeans?

11) Does message design inoculate Singaporeans from fake news/false information? What are the effects of different message designs?

12) Linked to the above, which segments of Singaporeans are more receptive to inoculation?

False information assumes many different forms. Misinformation and disinformation have pervaded societies prior to the invention of the Internet and social media. However, technology has catalysed their spread. While fake news is the pressing problem at present, false information is likely to persist and will continue to plague societies as human interactions grow. As such, it is more useful to develop a research agenda that addresses false information in general, and not just fake news. The above research questions, which can be further developed, have implications for policy intervention, not just in terms of regulation but also education and inoculation, and policy communication.

REFERENCES

2017 Edelman Trust Barometer reveals global implosion of trust. (2017, January 15). Retrieved from http://www.edelman.com/news/2017-edelman-trust-barometer-reveals-global-implosion/.

500,000 euro fines for fake news on Facebook in Germany?. (2016, December 16). *Deutsche Welle*. Retrieved from http://www.dw.com/en/500000-euro-fines-for-fake-news-on-facebook-in-germany/a-36806244.

A new understanding: What makes people trust and rely on news. (2016, April). Retrieved from http://www.mediainsight.org/PDFs/Trust/TrustFinal.pdf.

Adair, B. (2012, November 8). The value of fact-checking in the 2012 campaign. Retrieved from http://www.politifact.com/truth-o-meter/article/2012/nov/08/value-fact-checking-2012-campaign/.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election (No. w23089). *National Bureau of Economic Research*.

Allport, G. W., & Postman, L. (1947). The psychology of rumor.

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology*, 39(6), 1037.

Arkes, H. R., Hackett, C., & Boehm, L. (1989). The generality of the relation between familiarity and judged validity. *Journal of Behavioral Decision Making*, 2(2), 81-94.

Arsenault, A. (2016, October 20). Partisan Twitter bots distorting U.S. presidential candidates' popularity. *CBC News*. Retrieved from http://www.cbc.ca/news/world/twitter-bots-trump-clinton-1.3814386.

Au-Yong, R. (2017a, June 20). Call to empower citizens to combat fake news. *The Straits Times*. Retrieved from http://www.straitstimes.com/singapore/call-to-empower-citizens-to-combat-fake-news.

Au-Yong, R. (2017b, April 17). Experts welcome laws to curb fake news. *The Straits Times*. Retrieved from http://www.straitstimes.com/singapore/experts-welcome-laws-to-curb-fake-news.

Au-Yong, R. (2017c, June 25). Weak spot in the battle - the elderly who spread the gossip. *The Straits Times*. Retrieved from http://www.straitstimes.com/politics/weak-spot-in-the-battle-the-elderly-who-spread-the-gossip.

Babakar, M., & Moy, W. (2016, August). The state of automated fact checking [Report]. Retrieved from https://fullfact.org/media/uploads/full_fact_the_state_of_automated_factchecking_aug_2016.pdf

Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130-1132.

Barbaro, M. (2016, September 16). Donald Trump clung to 'birther' lie for years, and still isn't apologetic. *The New York Times*. Retrieved from https://www.nytimes.com/2016/09/17/us/politics/donald-trump-obama-birther.html.

Barron, J. (2017, March 20). In an era of fake news, teaching students to parse fact from fiction. *The New York Times*. Retrieved from https://www.nytimes.com/2017/03/20/nyregion/fake-news-brooklyn-middle-school.html.

Barthel, M., Mitchell, A., & Holcomb, J. (2016, December 15). Many Americans believe fake news is sowing confusion. Retrieved from http://www.journalism.org/2016/12/15/many-americans-believe-fake-news-is-sowing-confusion/.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4), 323.

Bebbington, K., MacLeod, C., Ellison, T. M., & Fay, N. (2017). The sky is falling: Evidence of a negativity bias in the social transmission of information. *Evolution and Human Behavior*, *38*(1), 92-101.

Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology*: General, 121(4), 446.

Berger, J. (2011). Arousal increases social transmission of information. *Psychological Science*, 22(7), 891-893.

Bilton, R. (2017, April 3). The News Integrity Initiative is taking a cross-industry approach to fixing the news trust problem. Retrieved from http://www.niemanlab.org/2017/04/the-news-integrity-initiative-is-taking-a-cross-industry-approach-to-fixing-the-news-trust-problem/

Bittner, J. (2017, June 21). Germany vs. Twitter. *The New York Times*. Retrieved from https://www.nytimes.com/2017/06/21/opinion/twitter-germany-free-speech.html.

Bouteldja, N. (2017, April 12). Fake news 'symptomatic of crisis in journalism'. *Aljazeera*. Retrieved from http://www.aljazeera.com/indepth/features/2017/04/fake-news-symptomatic-crisis-journalism-170412092154717.html.

Brock, T. C., & Balloun, J. L. (1967). Behavioral receptivity to dissonant information. *Journal of Personality and Social Psychology*, 6(4), 413-428.

Broder, D. S. (1989, January 19). Should news media police the accuracy of ads?. *The Washington Post*. Retrieved from https://www.washingtonpost.com/archive/politics/1989/01/19/should-news-media-police-the-accuracy-of-ads/42ef57f9-eb13-4fa0-8aa0-f809b0aae672/?utm_term=.8f0b2ac8c67a.

Bruck, M., & Ceci, S. J. (1999). The suggestibility of children's memory. *Annual Review of Psychology*, 50(1), 419-439.

Bump, P. (2016, December 19*).* Now you can fact-check Trump's tweets — in the tweets themselves. *The Washington Post*. Retrieved from https://www.washingtonpost.com/news/the-fix/wp/2016/12/16/now-you-can-fact-check-trumps-tweets-in-the-tweets-themselves/?utm_term=.0ae6a31ae1b9.

Burgess, M. (2016, November 17). Google is helping Full Fact create an automated, real-time fact-checker. Retrieved from http://www.wired.co.uk/article/automated-fact-checking-full-fact-google-funding

Byrne, S., & Hart, P. S. (2009). The Boomerang Effect — A synthesis of findings and a preliminary theoretical framework. *Annals of the International Communication Association*, 33(1), 3-37.

C4 study reveals only 4% surveyed can identify true or fake news. (2017, February 6). *Channel 4 News*. Retrieved from http://www.channel4.com/info/press/news/c4-study-reveals-only-4-surveyed-can-identify-true-or-fake-news

Cabato, R. (2017, March 7). Experts criticize proposed social media regulation. *CNN Philippines*. Retrieved from http://cnnphilippines.com/news/2017/03/04/experts-criticize-social-media-regulation.html.

Callister Jr, T. A. (2000). Media literacy: On-ramp to the literacy of the 21st century or cul-de-sac on the information superhighway. *Advances in Reading/Language Research*, *7*, 403-420.

Cameron, R. (2017, February 2). Fake news: Czechs try to tackle spread of false stories. *BBC News*. Retrieved from http://www.bbc.com/news/world-europe-38808501.

Carr, D. (2012, November 6). A last fact check: It didn't work [Blog post]. Retrieved from https://mediadecoder.blogs.nytimes.com/2012/11/06/a-last-fact-check-it-didnt-work/.

Cassino, D., & Jenkins, K. (2013, January 17). Conspiracy theories prosper: 25% of Americans are 'truthers.' [Press release]. Retrieved from http://publicmind.fdu.edu/2013/outthere/final.pdf.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, *39*(5), 752.

Chatwick, P. (2017, May 12). Defining fake news will help us expose it. *The Guardian*. Retrieved from https://www.theguardian.com/media/commentisfree/2017/may/12/defining-fake-news-will-help-us-expose-it

Chazan, G. (2017, March 14). Germany cracks down on social media over fake news. *Financial Times*. Retrieved from https://www.ft.com/content/c10aa4f8-08a5-11e7-97d1-5e720a26771b?mhq5j=e1.

Cheng, K. (2017, April 21). DSO engineers building system that detects fake news. *TODAY*. Retrieved from http://www.todayonline.com/singapore/dso-engineers-building-system-detects-fake-news.

Compton, J., Jackson, B., & Dimmock, J. A. (2016). Persuading others to avoid persuasion: Inoculation theory and resistant health attitudes. *Frontiers in Psychology*, *7*(122).

Cook, J., & Lewandowsky, S. (2011). The debunking handbook. Retrieved from https://www.skepticalscience.com/docs/Debunking_Handbook.pdf.

Cook, J., Lewandowsky, S., & Ecker, U. K. (2017). Neutralizing misinformation through inoculation: Exposing misleading argumentation techniques reduces their influence. *PloS one*, *12*(5), e0175799.

Culture, Media and Sport Select Committee 'fake news' inquiry: News Media Association response. (2017). Retrieved from http://www.newsmediauk.org/write/MediaUploads/Fake%20News/NMA_Submission_to_the_CMS_Select_Committee_'Fake_News'_Inquiry.pdf.

Dapkus, L. (2016, December 28). Lithuanian 'elves' combat Russian influence online. *Associated Press News*. Retrieved from https://apnews.com/27ce7f001bde4ccb9415ce4a0de74af1/lithuanian-elves-combat-russian-influence-online.

Davies, J. (2017a, January 25). Le Monde identifies 600 unreliable websites in fake-news crackdown. Retrieved from https://digiday.com/uk/le-monde-identifies-600-unreliable-websites-fake-news-crackdown/.

Davies, L. (2017b, May 30). UK pubs enlist bots to fight against filter bubbles ahead of the UK election. Retrieved from https://digiday.com/media/uk-pubs-enlist-bots-fight-filter-bubbles-ahead-uk-election/.

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, *113*(3), 554-559.

DiFonzo, N., & Bordia, P. (2007). Rumor, gossip and urban legends. *Diogenes*, *54*(1), 19-35.

Dillet, R. (2017, April 14). Facebook runs full page newspaper ads against fake news in France ahead of the election. Retrieved from https://techcrunch.com/2017/04/14/facebook-runs-full-page-newspaper-ads-against-fake-news-in-france-ahead-of-the-election/.

Dolley, C. (2017, March 5). Social media In South Africa could be regulated, says security minister. *The Huffington Post*. Retrieved from http://www.huffingtonpost.co.za/2017/03/05/social-media-in-south-africa-could-be-regulated-says-security-m_a_21873677/?utm_hp_ref=za-fake-news.

Donald, B. (2016, December 22). Stanford researchers find students have trouble judging the credibility of information online. Retrieved from https://ed.stanford.edu/news/stanford-researchers-find-students-have-trouble-judging-credibility-information-online.

Dutta, M. J. (2017, May 18). We need ethics, not regulation, to counter fake news. Retrieved from https://thewire.in/137089/fake-news-communication-ethics/.

Ecker, U. K., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review*, *18*(3), 570-578.

Edwards, K., & Smith, E. E. (1996). A disconfirmation bias in the evaluation of arguments. *Journal of Personality and Social Psychology*, 71(1), 5.

Faiola, A., & Kirchner, S. (2017, April 5). How do you stop fake news? In Germany, with a law. *The Washington Post*. Retrieved from https://www.washingtonpost.com/world/europe/how-do-you-stop-fake-news-in-germany-with-a-law/2017/04/05/e6834ad6-1a08-11e7-bcc2-7d1a0973e7b2_story.html?utm_term=.29e1d4c9cf52

Fallis, D. (2015). What is disinformation?. *Library Trends*, *63*(3), 401-426.

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

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, *59*(7), 96-104.

Festinger, L. (1956). *When prophecy fails: A social and psychological study of a modern group that predicted the destination of the world*. Harper.

Fleming, J., & Kajimoto, M. (2016). The freedom of critical thinking: Examining efforts to teach American news literacy principles in Hong Kong, Vietnam, and Malaysia. In *Handbook of Research on Media Literacy in the Digital Age* (pp. 208-235). IGI Global.

Flynn, D. J., Nyhan, B., & Reifler, J. (2017). The nature and origins of misperceptions: Understanding false and unsupported beliefs about politics. *Political Psychology*, 38(S1), 127-150.

Fogg, B. J., Soohoo, C., Danielsen, D., Marable, L., & Stanford, J. (2002, October 29). How do people evaluate a web sites credibility?: Results from a large study [Report]. Retrieved from https://dejanseo.com.au/media/pdf/credibility-online.pdf.

Forelle, M. C., Howard, P. N., Monroy-Hernández, A., & Savage, S. (2015). Political bots and the manipulation of public opinion in Venezuela [Research paper]. Retrieved from https://arxiv.org/ftp/arxiv/papers/1507/1507.07109.pdf.

Fridkin, K., Kenney, P. J., & Wintersieck, A. (2015). Liar, liar, pants on fire: How fact-checking influences citizens' reactions to negative advertising. *Political Communication*, 32(1), 127-151.

Funk, C. & Kennedy, B. (2016, October 4). The politics of climate. Retrieved from http://www.pewinternet.org/2016/10/04/the-politics-of-climate/.

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist*, 46(2), 107.

Gilbert, D. T., Tafarodi, R. W., & Malone, P. S. (1993). You can't not believe everything you read. *Journal of Personality and Social Psychology*, 65(2), 221.

Goodwin-Ortiz de Leon, C. (2017). Fake news on social media: Illusory truth and the 2016 presidential election [Bachelor thesis]. Retrieved from https://www.researchgate.net/publication/316418350_Fake_News_On_Social_Media_Illusory_Truth_and_the_2016_Presidential_Election.

Gorman, S. E., & Gorman, J. M. (2016). *Denying to the grave: Why we ignore the facts that will save us*. Oxford University Press.

Gottfried, J. & Shearer, E. (2016, May 26) News use across social media platforms 2016. Retrieved from http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016/.

Graeupner, D., & Coman, A. (2017). The dark side of meaning-making: How social exclusion leads to superstitious thinking. *Journal of Experimental Social Psychology*, *69*, 218-222.

Green, M. C., & Donahue, J. K. (2011). Persistence of belief change in the face of deception: The effect of factual stories revealed to be false. *Media Psychology*, 14(3), 312-331.

Greenberg, J., & Qiu, L. (2016, September 16). Fact-checking Donald Trump's claim Hillary Clinton started Obama birther movement. Retrieved from http://www.politifact.com/truth-o-meter/statements/2016/sep/16/donald-trump/fact-checking-donald-trumps-claim-hillary-clinton-/

Greitemeyer, T. (2014). Article retracted, but the message lives on. *Psychonomic Bulletin and Review*, 21(2), 557-561.

Grimes, D. R. (2017, June 14). Russian fake news is not new: Soviet Aids propaganda cost countless lives. *The Guardian*. Retrieved from https://www.theguardian.com/science/blog/2017/jun/14/russian-fake-news-is-not-new-soviet-aids-propaganda-cost-countless-lives.

Gupta, A., Lamba, H., & Kumaraguru, P. (2013, September). $1.00 per rt #bostonmarathon #prayforboston: Analyzing fake content on Twitter. *eCrime Researchers Summit (eCRS), 2013* (pp. 1-12). IEEE.

Haidt, J. (2014, June 12). The psychology behind morality [Video]. Retrieved from https://onbeing.org/programs/jonathan-haidt-the-psychology-behind-morality/.

Halberstam, Y., & Knight, B. (2016). Homophily, group size, and the diffusion of political information in social networks: Evidence from Twitter. *Journal of Public Economics*, *143*, 73-88.

Harb, Z. (2017, May 3). Can Wikitribune combat fake news?. *Aljazeera*. Retrieved from http://www.aljazeera.com/indepth/opinion/2017/05/wikitribune-combat-fake-news-170502084823082.html.

Hart, P. S., Nisbet, E. C., & Shanahan, J. E. (2011). Environmental values and the social amplification of risk: An examination of how environmental values and media use influence predispositions for public engagement in wildlife management decision making. *Society and Natural Resources*, 24(3), 276-291.

Hartman, T. K., & Newmark, A. J. (2012). Motivated reasoning, political sophistication, and associations between President Obama and Islam. *PS: Political Science and Politics*, 45(3), 449-455.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16(1), 107-112.

Hastorf, A. H., & Cantril, H. (1954). They saw a game; a case study. *Journal of Abnormal and Social Psychology*, 49(1), 129.

Heath, C., Bell, C., & Sternberg, E. (2001). Emotional selection in memes: the case of urban legends. *Journal of Personality and Social Psychology*, 81(6), 1028.

Heath, Y., & Gifford, R. (2006). Free-market ideology and environmental degradation the case of belief in global climate change. *Environment and Behavior*, 38(1), 48-71.

Herbert, D.G. (2017, January 12). A digital fact-checker fights fake news. *Bloomberg Businessweek*. Retrieved from https://www.bloomberg.com/news/articles/2017-01-12/a-digital-fact-checker-fights-fake-news.

Hern, A. (2017a, April 25). Google acts against fake news on search engine. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2017/apr/25/google-launches-major-offensive-against-fake-news.

Hern, A. (2017b, April 25). Wikipedia founder to fight fake news with new Wikitribune site. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2017/apr/25/wikipedia-founder-jimmy-wales-to-fight-fake-news-with-new-wikitribune-site.

Hess, A. (2017, March 3). How to escape your political bubble for a clearer view. *The New York Times*. Retrieved from https://www.nytimes.com/2017/03/03/arts/the-battle-over-your-political-bubble.html.

Holton, B., & Pyszczynski, T. (1989). Biased information search in the interpersonal domain. *Personality and Social Psychology Bulletin*, 15(1), 42-51.

Horowitz, J. (2016, December 2). Spread of fake news provokes anxiety in Italy. *The New York Times*. Retrieved from https://www.nytimes.com/2016/12/02/world/europe/italy-fake-news.html.

Indonesia to set up agency to combat fake news. (2017, January 5). *Yahoo Finance*. Retrieved from https://sg.finance.yahoo.com/news/indonesia-set-agency-combat-fake-news-110819835.html.

Italy's fake news battle. (2017, February 11). *Aljazeera*. Retrieved from http://www.aljazeera.com/programmes/listeningpost/2017/02/italy-fake-news-battle-170211115518930.html.

Ito, T. A., Larsen, J. T., Smith, N. K., & Cacioppo, J. T. (1998). Negative information weighs more heavily on the brain: the negativity bias in evaluative categorizations. *Journal of Personality and Social Psychology*, 75(4), 887.

Jackson, J. (2017a, January 12). BBC sets up team to debunk fake news. *The Guardian*. Retrieved from https://www.theguardian.com/media/2017/jan/12/bbc-sets-up-team-to-debunk-fake-news.

Jackson, J. (2017b, February 15). Fact-checkers are weapons in the post-truth wars, but they're not all on one side. *The Guardian*. Retrieved from https://www.theguardian.com/media/2017/feb/15/fact-checkers-are-weapons-in-the-post-truth-wars-but-theyre-not-all-on-one-side.

Jacques, P. J., Dunlap, R. E., & Freeman, M. (2008). The organisation of denial: Conservative think tanks and environmental scepticism. *Environmental Politics*, 17(3), 349-385.

Jadoo, Y. (2017, March 7). Attempts to regulate social media a step towards dictatorship – experts. Retrieved from http://citizen.co.za/news/news-national/1449025/attempts-to-regulate-social-media-a-step-towards-dictatorship-experts/

Jarvis, J. (2017). Our problem isn't 'fake news.' – Our problems are trust and manipulation. Retrieved from https://medium.com/whither-news/our-problem-isnt-fake-news-our-problems-are-trust-and-manipulation-5bfbcd716440

Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(6), 1420.

Johnson, T. J., & Kaye, B. K. (2004). Wag the blog: How reliance on traditional media and the Internet influence credibility perceptions of weblogs among blog users. *Journalism & Mass Communication Quarterly*, 81(3), 622-642.

Johnson, T. J., Bichard, S. L., & Zhang, W. (2009). Communication communities or "cyberghettos?": A path analysis model examining factors that explain selective exposure to blogs. *Journal of Computer-Mediated Communication*, 15(1), 60-82.

Jun, Y., Meng, R., & Johar, G. V. (2017). Perceived social presence reduces fact-checking. *Proceedings of the National Academy of Sciences*, 201700175.

Kahan, D. M., Hoffman, D. A., Braman, D., & Evans, D. (2012). They saw a protest: Cognitive illiberalism and the speech-conduct distinction. *Stanford Law Review*, 64(4), 851.

Kahan, D. M., Jenkins-Smith, H., & Braman, D. (2011). Cultural cognition of scientific consensus. *Journal of Risk Research*, 14(2), 147-174.

Kamins, M.A., & Assael, H. (1987). Two-sided versus one-sided appeals: A cognitive perspective on argumentation, source derogation, and the effect of disconfirming trial on belief change. *Journal of Marketing Research, 24*, 29-39.

Kang, H., & Sundar, S. S. (2016). When self is the source: Effects of media customization on message processing. *Media Psychology*, 19(4), 561-588.

Kang, H., Bae, K., Zhang, S., & Sundar, S. S. (2011). Source cues in online news: Is the proximate source more powerful than distal sources?. *Journalism & Mass Communication Quarterly*, 88(4), 719-736.

Kaos Jr., J. (2017, March 15). Government launches 'Tidak Pasti, Jangan Kongsi' to stop spread of false information. *The Star Online*. Retrieved from http://www.thestar.com.my/news/nation/2017/03/15/website-to-debunk-fake-news-government-launches-tidak-pasti-jangan-kongsi-to-stop-spread-of-false-in/.

Kaplan, J. T., Gimbel, S. I., & Harris, S. (2016). Neural correlates of maintaining one's political beliefs in the face of counterevidence. *Scientific Reports*, 6, 39589.

Kenyon, G. (2016, January 6). The man who studies the spread of ignorance. *BBC Future*. Retrieved from http://www.bbc.com/future/story/20160105-the-man-who-studies-the-spread-of-ignorance.

Keohane, M. (2017, March 7). What is the new definition of fake news?. Retrieved from https://www.buchananpr.com/2017/03/what-is-the-new-definition-of-fake-news

Knapp, R. H. (1944). A psychology of rumor. *Public opinion quarterly*, *8*(1), 22-37.

Knobloch-Westerwick, S., Mothes, C., Johnson, B. K., Westerwick, A., & Donsbach, W. (2015). Political online information searching in Germany and the United States: confirmation bias, source credibility, and attitude impacts. *Journal of Communication*, *65*(3), 489-511.

Kohring, M., & Matthes, J. (2007). Trust in news media: Development and validation of a multidimensional scale. *Communication Research*, *34*(2), 231-252.

Kosslyn, J., & Yu, C. (2017, April 7). Fact check now available in Google Search and News around the world [Blog]. Retrieved from https://www.blog.google/products/search/fact-check-now-available-google-search-and-news-around-world/.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480.

Ladd, J. M. (2011). *Why Americans hate the media and how it matters*. Princeton University Press.

Lapowsky, I. (2017, June 7). In a fake fact era, schools teach the ABCs of news literacy. Retrieved from https://www.wired.com/2017/06/fake-fact-era-schools-teach-abcs-news-literacy.

Lau, R. R., & Redlawsk, D. P. (2001). Advantages and disadvantages of cognitive heuristics in political decision making. *American Journal of Political Science*, 951-971.

Lee, P. (2017, March 2). Factually website clarifies 'widespread' falsehoods. *The Straits Times*. Retrieved from http://www.straitstimes.com/singapore/factually-website-clarifies-widespread-falsehoods.

Levin, S. (2017, May 16). Facebook promised to tackle fake news. But the evidence shows it's not working. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2017/may/16/facebook-fake-news-tools-not-working.

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction continued influence and successful debiasing. *Psychological Science in the Public Interest*, 13(3), 106-131.

Lewandowsky, S., Oberauer, K., & Gignac, G. E. (2013). NASA faked the moon landing—therefore, (climate) science is a hoax: An anatomy of the motivated rejection of science. *Psychological Science*, 24(5), 622-633.

Lim, C. (2017, May 16). Checking how fact-checkers check [Ph.D. dissertation]. Retrieved from https://drive.google.com/file/d/0B_wUaJ01JSddWFZZbE9TNmhKelE/view

Livingstone, S., & Olafsson, K. (n.d.). Children's commercial media literacy: New evidence relevant to UK policy decisions regarding the General Data Protection Regulation [Report]. Retrieved from

http://www.lse.ac.uk/media@lse/documents/MPP/Childrens-commercial-media-literacy-new-evidence-relevant-to-UK-policy-decisions-regarding-the-GDPR.pdf.

Loftus, E. F., & Davies, G. M. (1984). Distortions in the memory of children. *Journal of Social Issues*, 40(2), 51-67.

Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11), 2098.

Loughran, C. (2017, March 21). Fake news: How CrossCheck is debunking false election reporting. *BBC Academy Blog*. Retrieved from http://www.bbc.co.uk/blogs/academy/entries/708e52cf-0c13-45e2-904b-1a6c482e0d5b.

Luci, C. (2017, February 21). Alvarez wants to regulate social media use. *Manila Bulletin*. Retrieved from http://news.mb.com.ph/2017/02/21/alvarez-wants-to-regulate-social-media-use/.

Madden, M., Lenhart, A., & Fontaine, C. (2017, March 1). How youth navigate the news landscape. Retrieved from https://datasociety.net/output/how-youth-navigate-the-news-landscape/.

Madison, E. (2017, June 16). How a journalism class is teaching middle schoolers to fight fake news. Retrieved from https://theconversation.com/how-a-journalism-class-is-teaching-middle-schoolers-to-fight-fake-news-75043.

Maksl, A., Craft, S., Ashley, S., & Miller, D. (2017). The usefulness of a news media literacy measure in evaluating a news literacy curriculum. *Journalism and Mass Communication Educator*, *72*(2), 228-241.

Manjoo, F. (2015, May 7). Facebook use polarizing? Site begs to differ. *The New York Times*. Retrieved from https://www.nytimes.com/2015/05/08/technology/facebook-study-disputes-theory-of-political-polarization-among-users.html.

Manjoo, F. (2017, May 31). How Twitter is being gamed to feed misinformation. *The New York Times*. Retrieved from https://www.nytimes.com/2017/05/31/technology/how-twitter-is-being-gamed-to-feed-misinformation.html.

Mantzarlis, A. (2017, January 10). Regulation, the Inquisition, and popular tribunals: Italy's fake news debate just got real. Retrieved from http://www.poynter.org/2017/regulation-the-inquisition-and-popular-tribunals-italys-fake-news-debate-just-got-real/444200/

Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online [Report]. Retrieved from https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf.

McGuire, W. J. (1961). The effectiveness of supportive and refutational defenses in immunizing and restoring beliefs against persuasion. *Sociometry*, *24*(2), 184-197.

Mele, N., Lazer, D., Baum, M., Grinberg, N., Friedland, L., Joseph, K., & Mattsson, C. (2017, May 2). Combating fake news: An agenda for research and action. Retrieved from https://shorensteincenter.org/combating-fake-news-agenda-for-research/.

Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60(3), 413-439.

Meyer, R. (2017, February 3). The rise of progressive 'fake news'. *The Atlantic*. Retrieved from https://www.theatlantic.com/technology/archive/2017/02/viva-la-resistance-content/515532/.

Mihailidis, P., & Viotty, S. (2017). Spreadable spectacle in digital culture: Civic expression, fake news, and the role of media literacies in "post-fact" society. *American Behavioral Scientist*, 61(4), 441-454.

Mijatovic, D. (2016, December 1). Dunja Mijatović: Why bother? A quick take on lying on social media. Retrieved from https://www.indexoncensorship.org/2016/12/dunja-bother-quick-take-lying-social-media/.

Miller, J. M., Saunders, K. L., & Farhart, C. E. (2016). Conspiracy endorsement as motivated reasoning: The moderating roles of political knowledge and trust. *American Journal of Political Science*, 60(4), 824-844.

Min, Y. (2002). Intertwining of campaign news and advertising: The content and electoral effects of newspaper ad watches. *Journalism and Mass Communication Quarterly*, 79(4), 927-944.

Misener, D. (2016). Political bots spread misinformation during U.S. campaign. *CBC News*. Retrieved from http://www.cbc.ca/news/technology/political-bots-misinformation-1.3840300.

Morris, M. (n.d.). Trust and conspiracy theories [Report]. Retrieved from http://counterpoint.uk.com/trust-and-conspiracy-theories/.

Mosseri, A. (2016, December 15). News feed FYI: Addressing hoaxes and fake news [Press release]. Retrieved from https://newsroom.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/.

Mosseri, A. (2017, April 6). Working to stop misinformation and false news [Press release]. Retrieved from https://newsroom.fb.com/news/2017/04/working-to-stop-misinformation-and-false-news/.

Munro, G. D., & Ditto, P. H. (1997). Biased assimilation, attitude polarization, and affect in reactions to stereotype-relevant scientific information. *Personality and Social Psychology Bulletin*, 23(6), 636-653.

Murgia, M. (2017, April 3). Facebook launches $14m collaborative news literacy project. *Financial Times*. Retrieved from https://www.ft.com/content/4551ccf4-1793-11e7-9c35-0dd2cb31823a?mhq5j=e1.

New app allows public to report fake news. (2016, December 25). *The Jakarta Post*. Retrieved from http://www.thejakartapost.com/life/2016/12/25/new-app-allows-public-to-report-fake-news.html.

New Czech government unit seeks to monitor fake news. (2017, January 23). Retrieved from http://www.npr.org/2017/01/23/511267152/new-czech-government-unit-seeks-to-monitor-fake-news.

New JMSC project focuses on checking and verifying online sources. (2017, April 27). Retrieved from https://jmsc.hku.hk/2017/04/new-jmsc-project-focuses-on-checking-and-verifying-online-sources/.

Ng, K. (2017, June 19). Laws tackling fake news to be introduced next year: Shanmugam. *TODAY*. Retrieved from http://www.todayonline.com/singapore/new-laws-tackle-fake-news-be-introduced-next-year-shanmugam.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175.

Nyhan, B. (2017, February 15). Why more democrats are now embracing conspiracy theories. *The New York Times*. Retrieved from https://www.nytimes.com/2017/02/15/upshot/why-more-democrats-are-now-embracing-conspiracy-theories.html.

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303-330.

Nyhan, B., & Reifler, J. (2012). Misinformation and fact-checking [Research paper]. Retrieved from http://www.dartmouth.edu/~nyhan/Misinformation_and_Fact-checking.pdf.

Nyhan, B., & Reifler, J. (2013). Which corrections work [Research paper]. Retrieved from https://www.dartmouth.edu/~nyhan/nyhan-reifler-report-naf-corrections.pdf.

Nyhan, B., & Reifler, J. (2015). Estimating fact-checking's effects: Evidence from a long term experiment during campaign 2014 [Working paper]. Retrieved from http://www.americanpressinstitute.org/wp-content/uploads/2015/04/Estimating-Fact-Checkings-Effect.pdf.

Nyhan, B., Reifler, J., & Ubel, P. A. (2013). The hazards of correcting myths about health care reform. *Medical Care*, 51(2), 127-132.

Nyhan, B., Reifler, J., Richey, S., & Freed, G. L. (2014). Effective messages in vaccine promotion: a randomized trial. *Pediatrics*, 133(4), e835-e842.

O'Sullivan, P. B., & Geiger, S. (1995). Does the watchdog bite? Newspaper ad watch articles and political attack ads. *Journalism and Mass Communication Quarterly*, 72(4), 771-785.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you.* Penguin UK.

Pennycook, G., Cannon, T. D., & Rand, D. G. (2017). Prior exposure increases perceived accuracy of fake news [Research paper]. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2958246.

Polage, D. C. (2012). Making up history: False memories of fake news stories. *Europe's Journal of Psychology*, 8(2), 245.

Politi, J. (2016, December 30). Italy antitrust chief urges EU to help beat fake news. *Financial Times*. Retrieved from https://www.ft.com/content/e7280576-cddc-11e6-864f-20dcb35cede2?mhq5j=e1.

Price, R. (2017, April 7). Facebook is trying to teach its users how to spot fake news. *Business Insider*. Retrieved from http://www.businessinsider.sg/facebook-how-to-spot-fake-news-2017-4?r=US&IR=T.

Price, V., Nir, L., & Cappella, J. N. (2006). Normative and informational influences in online political discussions. *Communication Theory*, 16(1), 47-74.

Proctor, R. N. (2004). Should medical historians be working for the tobacco industry?. *The Lancet*, 363(9416), 1174.

Reese, S. D., & Lewis, S. C. (2009). Framing the war on terror: The internalization of policy in the US press. *Journalism*, *10*(6), 777-797.

Robbins, M. (2017, January 31). Fake news and fact-checking: Trump is demonstrating how to outsmart an AI. *The Guardian*. Retrieved from https://www.theguardian.com/science/2017/jan/31/fake-news-and-fact-checking-trump-is-demonstrating-how-to-outsmart-an-ai-artificial-intelligence

Rojecki, A., & Meraz, S. (2016). Rumors and factitious informational blends: The role of the web in speculative politics. *New Media and Society*, *18*(1), 25-43.

Rose, J., & Barros, M. (2017, January 20). Scientists have a word for studying the post-truth world: Agnotology. Retrieved from http://theconversation.com/scientists-have-a-word-for-studying-the-post-truth-world-agnotology-71542.

Rosnow, R. L. (1980). Psychology of rumor reconsidered. *Psychological Bulletin*, 87(3), 578-591.

Rosnow, R. L., Esposito, J. L., & Gibney, L. (1988). Factors influencing rumor spreading: Replication and extension. *Language and Communication*, 8(1), 29-42.

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 13(3), 279-301.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 32(5), 880.

Rubin, V. L., Chen, Y., & Conroy, N. J. (2015). Deception detection for news: three types of fakes. *Proceedings of the Association for Information Science and Technology*, *52*(1), 1-4.

Ruchansky, N., Seo, S., & Liu, Y. (2017). CSI: A hybrid deep model for fake news [Research paper]. Retrieved from https://arxiv.org/pdf/1703.06959.pdf.

Rucker, D.D., Petty, R.E., & Brinol, P. (2008). What's in a frame anyway? A meta-cognitive analysis of the impact if one versus two-sided message framing on attitude certainty. *Journal of Consumer Psychology, 18*, 137-149.

Sapiie, M. A. (2016. December 3). Hoax busters join hands to fight rampant fake news. *The Jakarta Post.* Retrieved from http://www.thejakartapost.com/news/2016/12/03/hoax-busters-join-hands-fight-rampant-fake-news.html.

Schaer, C. (2017, May 26). Fake laws against fake news. Retrieved from https://global.handelsblatt.com/politics/fake-laws-against-fake-news-772564.

Schwarz, N., Sanna, L. J., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, *39*, 127-161.

Seow, B. Y. (2017, June 21). Good news: Help in spotting fake news coming soon. *The Straits Times*. Retrieved from http://www.straitstimes.com/singapore/good-news-help-in-spotting-fake-news-coming-soon.

Shao, C., Ciampaglia, G. L., Flammini, A., & Menczer, F. (2016, April). Hoaxy: A platform for tracking online misinformation. In *Proceedings of the 25th International Conference Companion on World Wide Web* (pp. 745-750). International World Wide Web Conferences Steering Committee.

Shead, S. (2017, May 30). Facebook said Germany's plan to tackle fake news would make social media companies delete legal content. *Business Insider.* Retrieved from http://uk.businessinsider.com/facebook-says-germany-fake-news-plans-comply-with-eu-law-2017-5.

Silverman, C., & Alexander, L. (2014, November 4). How teens in the Balkans are duping Trump supporters with fake news. *BuzzFeed News.* Retrieved from https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm_term=.vk8aQd9eqe#.jmYjpymzRz.

Silverman, C., & Singer-Vine, J. (2016, December 7). Most Americans who see fake news believe it, new survey says. *BuzzFeed News*. Retrieved from https://www.buzzfeed.com/craigsilverman/fake-news survey?utm_term=.gxg1nkeB9B#.quQ6z5glWl.

Simo, F. (2017). Introducing: The Facebook Journalism Project [Press release]. Retrieved from https://media.fb.com/2017/01/11/facebook-journalism-project/.

Skurnik, I., Yoon, C., Park, D. C., & Schwarz, N. (2005). How warnings about false claims become recommendations. *Journal of Consumer Research*, *31*(4), 713-724.

Southern, L. (2017, March 21). What to know about Germany's fake-news crackdown. Retrieved from https://digiday.com/media/germanys-proposed-hate-crime-law/

Stelter, B. (2016, November 7). How Donald Trump made fact-checking great again. *CNN Media*. Retrieved from http://money.cnn.com/2016/11/07/media/donald-trump-fact-checking/index.html

Stenhouse, N., Maibach, E., Cobb, S., Ban, R., Bleistein, A., Croft, P. & Leiserowitz, A. (2014). Meteorologists' views about global warming: a survey of American Meteorological Society professional members. *Bulletin of the American Meteorological Society*, 95(7), 1029-1040.

Subramanian, S. (2017, February 15). The Macedonian teens who mastered fake news. Retrieved from https://www.wired.com/2017/02/veles-macedonia-fake-news/

Sundar, S. S. (2016). Why we fall for fake news: The psychology of online news consumption. *Newsweek*. Retrieved from http://www.newsweek.com/why-we-fall-fake-news-529777.

Sunstein, C. R. (2014a). *Conspiracy theories and other dangerous ideas*. Simon and Schuster.

Sunstein, C. R. (2014b). *On rumors: How falsehoods spread, why we believe them, and what can be done*. Princeton University Press.

Sunstein, C. R., Bobadilla-Suarez, S., Lazzaro, S. C., & Sharot, T. (2016). How people update beliefs about climate change: Good news and bad news [Research paper]. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2821919.

Swami, V. (2012). Social psychological origins of conspiracy theories: the case of the Jewish conspiracy theory in Malaysia. *Frontiers in Psychology*, *3*(280).

Swift, A. (2016, September 14). Americans' trust in mass media sinks to new low. Retrieved from http://www.gallup.com/poll/195542/americans-trust-mass-media-sinks-new-low.aspx.

Swire, B., Berinsky, A. J., Lewandowsky, S., & Ecker, U. K. (2017). Processing political misinformation: comprehending the Trump phenomenon. *Royal Society Open Science*, 4(3), 160802.

Sydell, L. (2016, November 23). We tracked down a fake-news creator in the suburbs. Here's what we learned. Retrieved from http://www.npr.org/sections/alltechconsidered/2016/11/23/503146770/npr-finds-the-head-of-a-covert-fake-news-operation-in-the-suburbs.

Taber, C. S., & Lodge, M. (2006). Motivated skepticism in the evaluation of political beliefs. *American Journal of Political Science*, 50(3), 755-769.

Tait, R. (2016, December 28). Czech Republic to fight 'fake news' with specialist unit. *The Guardian*. Retrieved from https://www.theguardian.com/media/2016/dec/28/czech-republic-to-fight-fake-news-with-specialist-unit

Tavernise, S. (2016, December 6). As fake news spreads lies, more readers shrug at the truth. *The New York Times*. Retrieved from https://www.nytimes.com/2016/12/06/us/fake-news-partisan-republican-democrat.html.

The Poynter Institute and Facebook launch Innovative Journalism Certificate. (2017, April 10). Retrieved from http://about.poynter.org/about-us/press-room/poynter-institute-and-facebook-launch-innovative-journalism-certificate.

The pros and cons of regulating social media. (2017, February 26). *ABS-CBN*. Retrieved from http://news.abs-cbn.com/focus/02/26/17/the-pros-and-cons-of-regulating-social-media.

Toor, A. (2017, January 15). Facebook rolls out fake news filter in Germany. Retrieved from https://www.theverge.com/2017/1/15/14277964/facebook-fake-news-filter-germany.

Tversky, A., & Kahneman, D. (1975). Judgment under uncertainty: Heuristics and biases. In *Utility, probability, and human decision making* (pp. 141-162). The Netherlands: Springer.

Uscinski, J. E., Klofstad, C., & Atkinson, M. D. (2016). What drives conspiratorial beliefs? The role of informational cues and predispositions. *Political Research Quarterly*, 69(1), 57-71.

Uscinski, J. E., Parent, J., & Torres, B. (2014). Conspiracy theories are for losers. In *American Conspiracy Theories* (pp. 131-153). Oxford University Press.

Van der Linden, S., Leiserowitz, A., Rosenthal, S., & Maibach, E. (2017). Inoculating the public against misinformation about climate change. *Global Challenges*, *1*(2).

Van Prooijen, J. W., Krouwel, A. P., & Pollet, T. V. (2015). Political extremism predicts belief in conspiracy theories. *Social Psychological and Personality Science*, *6*(5), 570-578.

Varol, O., Ferrara, E., Davis, C. A., Menczer, F., & Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization [Research paper]. Retrieved from https://arxiv.org/pdf/1703.03107.pdf.

Wardle, C. (2017, February 17). Fake news. It's complicated. Retrieved from https://medium.com/1st-draft/fake-news-its-complicated-d0f773766c79

Wathen, C. N., & Burkell, J. (2002). Believe it or not: Factors influencing credibility on the web. *Journal of the Association for Information Science and Technology*, *53*(2), 134-144.

Weedon, J., Nuland, W., & Stamos, A. (2017, April 27). Information operations and Facebook. Retrieved from https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf.

Weeks, B. E. (2015). Emotions, partisanship, and misperceptions: How anger and anxiety moderate the effect of partisan bias on susceptibility to political misinformation. *Journal of Communication*, 65(4), 699-719.

Weiss, M. (2016, March 20). The Baltic elves taking on pro-Russian trolls. *The Daily Beast*. Retrieved from http://www.thedailybeast.com/the-baltic-elves-taking-on-pro-russian-trolls.

Wendling, M. (2017, January 30). Solutions that can stop fake news spreading. *BBC News*. Retrieved from http://www.bbc.co.uk/news/blogs-trending-38769996

Woolley, S. & Howard, P. (2016, May 15). Bots unite to automate the presidential election. Retrieved from https://www.wired.com/2016/05/twitterbots-2/.

Zubiaga, A., Liakata, M., Procter, R., Hoi, G. W. S., & Tolmie, P. (2016). Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PloS one*, *11*(3), e0150989.

USEFUL RESOURCES

1. *A Field Guide to Fake News*, Public Data Lab and First Draft - https://fakenews.publicdatalab.org/download/SAMPLE-field-guide-to-fake-news.pdf

2. *Finding the Truth Among the Fakes*, Aljazeera - http://institute.aljazeera.net/mritems/Documents/2017/3/13/93ccf5e4cc834436999f71b11ce8ca53_100.pdf

3. Lies Damn Lies, Tow Center for Digital Journalism - http://towcenter.org/wp-content/uploads/2015/02/LiesDamnLies_Silverman_TowCenter.pdf

4. *Media Manipulation and Disinformation Online*, Data and Society - https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

5. Social Media – A Handbook for Journalists, Sverige Radio - https://www.ebu.ch/files/live/sites/ebu/files/Publications/Swedish_Radio-Social_Media-Handbook-for-Journalists.pdf

6. The Debunking Handbook, John Cook and Stephan Lewandowsky - https://www.skepticalscience.com/docs/Debunking_Handbook.pdf

7. *The Fake News Machine*, Trend Micro - https://documents.trendmicro.com/assets/white_papers/wp-fake-news-machine-how-propagandists-abuse-the-internet.pdf?_ga=2.117063430.1073547711.1497355570-1028938869.1495462143

8. *The State of Automated Fact Checking*, Full Fact - https://fullfact.org/media/uploads/full_fact-the_state_of_automated_factchecking_aug_2016.pdf

9. *The Verification Handbook*, Eurpoean Journalism Centre - http://verificationhandbook.com/downloads/verification.handbook_additional.pdf