From Russia with Fear: The Presence of Emotion in Russian Disinformation Tweets

July 2019

2338913 (UoG)
17116554 (DCU)
44840611 (Charles University)

Presented in partial fulfilment of the requirements for the Degree of
International Master in Security, Intelligence and Strategic Studies

Abstract

Russian disinformation continues to be an ongoing issue in the present security environment. International organisations (e.g. EU) and researchers highlight that emotional appeals (mostly related to fear, anger, and prejudice) in Russian disinformation are used to deepen social division and increase polarisation surrounding a particular issue (European Parliamentary Research Service, 2019; Sivek, 2018; Nisbet & Kamenchuk, 2019; Asmolov, 2018; Bennet & Livingston, 2018; Schmitt, 2018; Karlsen, 2016). While the acknowledgement of emotional appeals in the disinformation literature is common, research into these appeals is sparse. The present study provides an overview of both emotion and disinformation literature and aims to answer three research questions: what emotions are present in Russian disinformation, are some emotions more common than others, and does disinformation communicate specific topics through discrete emotions? Through emotion (sentiment) analysis we found presence of all 8 of Plutchik's emotions in a Russian disinformation tweet dataset; fear, anger, trust, anticipation, sadness, joy, disgust, and surprise. Within the tweet corpus, approximately 5% of tweets belonged to a discrete emotional frame, with fear and anger the most prevalent by a large margin. Specific emotion categories contained 'crime,' 'politics,' and 'patriotism' as prominent themes. This study paves the way for more research into the use of emotional appeals in Russian disinformation.

Table of Contents

Definitions and Abbreviations

EU: European Union

IRA: Internet Research Agency

KGB: Komitet Gosudarstvennoy Bezopasnosti (Soviet State Security Agency)

NATO: North Atlantic Treaty Organisation

US: United States of America

The 'West': concerning majority consensus of EU and NATO member states

Chapter 1 Introduction

In the present security environment, Russian disinformation continues to be an ongoing issue.

This week, Robert Mueller testified to US Congress regarding his investigation into the

possible collusion of US President Trump and the Russian government (Guardian, 2019). In Europe, Russian disinformation operations have been observed in Germany, France, the Netherlands, Italy, and the United Kingdom (Stelzenmüller, 2017; Institute for Strategic Dialogue, 2017; Dutch Intelligence Services, 2017; Alaphilippe et al., 2018; Bakamo, 2017). The European Union East StratCom Task Force has increased its budget by more than double the previous amount and new initiatives continue to develop within member states to promote media literacy and debunk disinformation (European Commission, 2018). Russian disinformation is a timely and pressing issue.

A recent report from the EU highlights that emotional appeals (mostly related to fear, anger, and prejudice) in Russian disinformation are used to deepen social division and increase polarisation surrounding a particular issue (European Parliamentary Research Service, 2019). Many academic researchers have also noted that an 'appeal to emotion' is a common strategy of Russian disinformation efforts (Sivek, 2018; Nisbet & Kamenchuk, 2019; Asmolov, 2018; Bennet & Livingston, 2018; Schmitt, 2018; Karlsen, 2016). Besides the acknowledgement of emotional appeals, the disinformation literature lacks research exploring this issue. The present study aims to fill this gap by linking emotional appeal research from social psychology and political communication to Russian disinformation. Before outlining the structure of the present study we provide a background and justification.

1.1 Background

Although disinformation is widely written about, many authors do not provide adequate definitions of disinformation and oftentimes if they do, definitions are too broad or too narrow. We argue that disinformation is 'misleading information that has the intentional

function to mislead,' a definition adapted from Fallis (2015). We view disinformation from a constructivist position and argue that it is socially and culturally mediated. Emotion, much like disinformation is also a difficult one to define. We adopt Scherer's (2009: 1) functional perspective of emotion as 'a cultural and psychobiological adaptation mechanism which allows each individual to react flexibly and dynamically to environmental contingencies.' This functional approach to emotion is beneficial for the purposes of this study as it directly links emotions to behaviour and judgement patterns.

Disinformation does not exist in a vacuum and therefore we must analyse it in the context of the current digital media environment. We characterise the changing digital environment by the presence of Web 2.0 technologies, the increasing collection of user data, and the digitisation of news. Web 2.0 technologies allow users to actively participate in their digital environment as opposed to being passive consumers. The nonstop collection of user data allows advertisers and news disseminators and other websites to personalise content, so that content is more likely engaged with. The current digital environment enables disinformation by the lack of truth verification, the relative ease in creating content online, plausible deniability, increased personalisation, and the speed at which information is spread.

Enabled by the current digital environment, Russian disinformation operations are polarised and emotive (Allen & Moore, 2018). As previously stated, many researchers have noted an 'appeal to emotion' strategy within studies of Russian disinformation (Sivek, 2018; Nisbet & Kamenchuk, 2019; Asmolov, 2018; Bennet & Livingston, 2018; Schmitt, 2018; Karlsen, 2016). They state that disinformation campaigns actively use emotional extremity in their messaging to influence the judgements and behaviours of viewers. The largest study to date

of online disinformation analysed the spread of 126,000 verified true and false news stories on Twitter between 2006 and 2017 (Vosoughi et al., 2018). Researchers of the study found that disinformation 'diffused significantly farther, faster, deeper, and more broadly' than accurate information, propelled by emotional reactions such as fear and disgust (Vosoughi et al., 2018).

The use of emotional appeals, forms of communication which intend to elicit an emotional response, are widely noted within public health education (Tannenbaum, 2015) and political campaigning (Brader, 2006) to have an effect on viewers' judgments and behaviours. Researchers within these disciplines have argued for over a decade that emotional appeals can 'influence the participation and choices of viewers in distinct ways' (Brader, 2006: 13). However, emotional appeals have not been studied within the context of or applied to Russian disinformation. As such, the justification for the present study has come from the results of various disciplines including public health communication, social psychology, media framing, and political communication.

For example, in a recent meta-analysis of over 50 years of emotional appeal research, researchers found that fear appeals are effective at influencing and changing attitudes, intentions, and behaviours (Tannenbaum et al., 2015). One of the lead researchers of the study stated in an interview that 'fear produces a significant though small amount of change across the board. Presenting a fear appeal more than doubles the probability of change relative to not presenting anything or presenting a low-fear appeal' (Dolores Albarracin in American Psychological Association, 2015). Furthermore, an experimental emotion manipulation has shown to affect attitudes about drunk driving; angry participants were more

likely to favour retributive policies while fearful participants preferred protective solutions (Nabi, 2003).

Having described the importance of emotion in disinformation and the lack of research surrounding it, the present study aims to fill a fraction of this research gap. Firstly, we assess which specific emotions exist in Russian disinformation. Subsequently we discover the emotions that are the most prominent in disinformation. Finally, we adopt an exploratory qualitative analysis to synthesise the content of disinformation present within each emotion. This study brings together research from social psychology, political communication, and security studies with the aim of further understanding Russian disinformation.

Chapter 2 Disinformation

Although disinformation is a headline issue, defining it does not come easy. Across the security literature, definitions of disinformation are frequently too narrow, too broad, or simply contradict each other. While many definitions rely on the presumption of falsehood and intention, we hone in on the philosophy of information to provide the present study's definition of disinformation. We also introduce constructivism as the epistemological approach which accepts that disinformation is socially, culturally, and historically mediated (Karlova & Fisher, 2012). As disinformation is inherently associated with truth and

falsehoods, we discuss our approach to the nature of truth and adopt the coherence theory of truth which argues something is true if it coheres with the 'truths' of other individuals, or one's own beliefs (Stahl, 2006). This first section lays the disinformation theoretical groundwork for the rest of the study.

Although many states have been aware of Russian disinformation since the end of the Cold War (e.g. Baltics), most Western (EU and NATO) governments have slowly realised the damaging potential of disinformation. Having defined disinformation and provided a theoretical approach, this chapter subsequently outlines the historical context of Russian disinformation and presents both a Russian and a Western perspective. Within this literature review, we argue that much of the tactics and aims of Russian disinformation are not new - but rather an extension of Soviet practices during the Cold War. What is new is the current media ecosystem. In a growing social media environment, the way disinformation is conducted and disseminated is not changing, but rather adapting. Social media sites, like Twitter, exacerbate the potentially damaging effects of disinformation with the advent of ad personalisation, boosted connectedness, and a lack of news verifiability. The chapter concludes with our justification for the study of disinformation on Twitter.

2.1 Defining Disinformation

Discussions of disinformation tend to suffer from either a lack of definitions or imprecise definitions, in part due to the interdisciplinary and topical nature of disinformation. Many definitions frame disinformation within the current political climate, rendering it too narrow.

For example, Bennet & Livingston (2018: 124) provisionally define disinformation as 'intentional falsehoods spread as news stories or simulated documentary formats to advance political goals.' If we were to take that definition, satire and news sites like The Onion may be regarded as disinformation - but that is not the case. Others more broadly define it as something false and intending to cause harm (Wardle, 2018; Edson et al., 2018). Both of these definitions are too narrow as disinformation is not always false; cherry-picking statistics to support a particular worldview is considered disinformation, although the content is technically true (Søe, 2018). Interestingly, political actors tend to define disinformation in relation to its intent regarding the public sphere. The European Commission (2019) understands disinformation 'as verifiably false or misleading information that is created, presented and disseminated for economic gain or to intentionally deceive the *public, and may cause public harm*.' The UK government (2018) defines it as 'the deliberate creation and sharing of false and/or manipulated information that is intended to deceive and mislead *audiences*, either for the purpose of causing harm, or for political, personal or financial gain.' These definitions encompass factually 'true' disinformation presented in a misleading or manipulative way. However, the stated intentions of disinformation are too narrow again- as the mislead actor is not one individual, but the general public. If disinformation misleads only one individual, it is still disinformation. The above definitions highlight a problem within the academic and political realms where definitions of disinformation tends to be imprecise or rely on broad categorisations of falsehood and intent.

Stepping back from the security and politics realm, there has been an ongoing debate within philosophy of information between Luciano Floridi (2005; 2011) and Don Fallis (2009; 2015) for about a decade regarding the definition of disinformation. However, before diving into this debate, it is important to understand how both scholars conceptualise the nature of

information, as it has consequences in how they address the definition of disinformation. Floridi (2004; 2011) argues information is semantically well-informed, meaningful, and *truthful* data. In his view, information inherently possesses meaning, and without it, it is considered 'data (Floridi, 2005).' However, for the present study, this notion is problematic. If information is true, then this suggests that it is possible to know what is objectively true, and as a consequence reality is not socially-mediated, but independent and objective. This view cannot explain the subjective position of information; some information may hold meaning for one individual and not hold any meaning for another. Within this theory, information becomes a static entity that exists independently from any individual observer (Adriaans, 2010; Karlova & Lee, 2011). For the present study, information is conceived as something that represents the world, a form of representational content, which may be *true or false.*

Our conception of information aligns with that of Fallis (2009; 2015). Ultimately, the differences in the two scholars' positions on information reflect differences in their ontological positions. Floridi (2003) in philosophical realism, accepts that the nature of the world, including information, exists outside of our conceptual schemes, perceptions, and beliefs. As such, information inherently possesses meaning and truth value. In contrast, Fallis (2015) and the position of the present study, accept philosophical constructivism which emphasises dependence on the *context* of information.

With differences in the conception of information, the two scholars have very different views when it comes to the definition of disinformation. However, the debate on the definition of disinformation isn't much of a debate when their ontological positions are taken into account. In Floridi's (2011) assessment, disinformation is false semantic information that is

disseminated in order to deceive its receiver. Because in realism something can be objectively false, this definition is completely valid. However, the present study is grounded in philosophical constructivism and will take Fallis's (2015) definition of disinformation; 'misleading information that has the function of misleading someone (Fallis, 2015: 413).' Whether information is considered as disinformation, depends on who receives the information, and renders disinformation a socially constructed concept (Fallis, 2015; Karlova & Fisher, 2012). Further, it is not only important that disinformation is itself misleading, but it is equally important to know how it became misleading. We further argue that disinformation's misleading function must be acquired through design (e.g. an actor's intention to mislead) rather than evolution (e.g. information becomes outdated). If a news program unwittingly disseminates a misleading story which misleads its viewers, it is not considered disinformation unless the news program intended to mislead its viewers. As a final note, although the function of disinformation is to mislead, it does not need to be its ultimate purpose, which may include advancing political goals or general harm to a foreign state.

2.1.1 Associated Terms

In addition to the difficulty in defining disinformation, there is also confusion about its associated terms - propaganda, information warfare, and misinformation. While disinformation and propaganda have been frequently used as synonyms of the same phenomena (Cunningham, 2001), others view disinformation as a subset of propaganda characterising disinformation as 'propaganda-plus' (Martin, 1982). Disinformation is a relatively new word, and was coined by the Soviet Union (dezinformatsiya) to define the dissemination of false reports intended to mislead public opinion (Taylor, 2016). In contrast,

the term propaganda originated in the 1600s and is used to denote broad political communication (Merriam Webster, 2019). As such, we take the view that disinformation is a specialised part of propaganda.

While information warfare denotes the use of information as a tool, target, or a domain of operations for hostile activities, it also does not require falsehood (Giles, 2016). To some, information warfare overlaps within disciplines of psychological operations, electronic warfare, counterintelligence, strategic communications, and disinformation while for other scholars, disinformation is a strategy of information warfare (Giles, 2016; Hellman & Wagnsson, 2017; Theohary, 2018). Bellamy (2001: 70) most succinctly defined information warfare as a grand strategy covering strategic deception, cyber war, and disinformation operations. Lastly, misinformation, although similar to disinformation, is broadly defined as unintentional but false information; examples include clickbait, satire, or misleading quotes and images (Wardle, 2018; Edson et al., 2018; Theohary, 2018).

2.2 Theoretical Considerations

Epistemologically, the present study adopts constructivism to the understanding of disinformation. Constructivism within international relations is not interested in the objective facts of the world, but rather their social meaning (Wendt; 2000; Slaughter, 2011). Meaning is constructed from a complex mix of history, ideas, norms, and beliefs which we need to understand if we are to explain state behaviours, such as disinformation operations (Slaughter, 2011). A constructivist view of disinformation is useful because it emphasises

social context (Karlova & Fisher, 2012; Hjorland, 2007; Fallis, 2009). Karlova and Fisher's (2012) 'social diffusion of information model' specifically highlights disinformation as a constructivist notion, with social, cultural, and historical context-awareness. They reason that as disinformation diffuses over time, its acceptance as valid information (or disinformation) may change, disappear, or emerge. In other words, true, accurate information may become misinformation or disinformation from a change in context and vice versa.

2.2.1 Truth

By defining disinformation as intentionally misleading content, this suggests that content itself is not misleading, and brings us to consider the very nature of truth. Among the various theories of truth, the coherence theory is most applicable due to our epistemological position. While the correspondence theory of truth states that what is true is what accurately describes the state of the world, the coherence theory states that something is true if it coheres with the 'truths' of other individuals, or one's own beliefs (Stahl, 2006). The only way we can access the truth is through the perceptions and interpretations of other people, and then corroborating those interpretations to discover truth. One clear example of the coherence theory in security and international relations is the importance of consensus (or coherence of 'truths') among state actors. This noticeably manifests over the recognition of state territory and sovereignty. For most states, recognition of their territorial integrity coheres with the view of other states. However, when states disagree over territorial integrity, they use consensus to decide what is true and once consensus is reached, a state's territory will become 'true.' Currently, there are numerous examples of disputed territories (e.g. Crimea, Abkhazia, South Ossetia), and they are disputed because of a lack of consensus among state actors.

Like most concepts in the social sciences, disinformation is difficult to define. Many definitions are imprecise, and rely on broad categorisations. The present study recognises that disinformation is a type of information (Fallis, 2015), and although some scholars believe that information has to be true (Floridi, 2011), we take the definition of information as something that represents the world, which may be true or false. What separates disinformation from information is that disinformation is misleading; it is 'likely to create false beliefs' (Fallis, 2015: 406). Secondly, disinformation must also possess the intentional function of misleading. The above characteristics of disinformation lead us to an adaptation of Fallis' (2015) definition of disinformation as 'misleading information that has the [intentional] function of misleading.' As the study is grounded in a constructivist epistemology, disinformaiton may be socially, culturally, and historically mediated.

2.3 Review of the Disinformation Literature

2.3.1 Historical Context

The term disinformation originated from the Russian 'dezinformatsiya' in the 1950s. During the Cold War, Russian disinformation was a small fraction of an extensive 'active measures' campaign. Under the direction of the KGB, 'active measures' included media manipulations, political assassinations, kidnappings, the establishment of front organisations, forgeries of official documents, and of course disinformation campaigns (Romerstein, 2001; Boghardt,

2009; US Department of State, 1987). These efforts were considered an indispensable part of Soviet ammunition to achieve its ideological and geopolitical goals (Abrams, 2016; Allen & Moore, 2018; Rosenstein, 2001). Soviet disinformation aims were consistent with the aims of active measures; to 'confuse the enemy and to cause him to take an action beneficial to the Soviet Union' (Romerstein, 2001, p. 54). One widely memorable disinformation campaign falsely spread information that the US created and disseminated the HIV virus as part of a failed biological weapons programme (Romerstein, 2001; Boghardt, 2009). As the disease spread, largely into black communities, further disinformation purported the conspiracy that AIDS was used directly against black Americans. Although the Soviets eventually retracted the disinformation, the campaign struck a societal cleave in the US and had lasting social and cultural impacts. As late as in a 2005 single, Kanye West wrote 'And I know the government administered AIDS / So I guess we just pray like the minister say.' This example highlights the pervasive and long-lasting nature of Soviet disinformation and its adhesion to the public when attached to a divisive social issue. These sorts of disinformation measures were so effective the U.S. set up a special 'active measures' working group in 1981 to investigate and counter Soviet disinformation efforts (Boghardt, 2009).

Today, many academics and security professionals notice a striking similarity between contemporary Russian disinformation campaigns and those during the Cold War (Ostrovsky, 2017; Abrams, 2016; Darczewska, 2014). The general aim seems to be the same - weaken the enemy and generate a beneficial position for Russia. The tactics that are employed today are also similar to ones employed during the Cold War. For example, creating false narratives from already existing social issues to sow discord, such as the divide between the American political elite and Black American communities. The current disinformation strategy around Black Lives Matter and police-related shootings seems to mirror AIDS disinformation

campaigns during the Cold War (Spangher et al., 2018). The parallels include the use of emotion-driven issues (AIDS and police racial bias) tied to social and political movements (civil rights campaign and Black Live Matter). Although targeting issues surrounding Black American discrimination is one of many social issues Russian disinformation tries to aggravate, it nonetheless illustrates how Cold War and contemporary disinformation campaigns concentrated on polarising issues with strong emotional responses.

2.3.2 Russian Perspective of Disinformation

By noting the similarities between the Cold War-era disinformation campaigns and disinformation today, we can observe that many general aims and tactics have not changed. However, to gain a full picture of Russian disinformation we must also examine the Russian perspective. As we are limited to literature sources translated to English, we turn to well-known Russian military thinkers whose work has been widely circulated and translated (e.g. Valery Gerasimov). These scholars do not write about disinformation specifically, but rather about the future of the nature of war and hybrid warfare operations of NATO and EU states (Gerasimov, 2013; Kartapolov, 2015). From these texts, we extrapolate what they say about information operations and the role disinformation plays in the future of war.

Before exploring the themes of Russian scholars vis-a-vis disinformation, there are two trends that must be pointed out. First, when discussing hybrid tactics, many Russian scholars frame them as something solely the West uses against Russia (Gerasimov, 2013; Darczewska, 2014). For example, Kartapolov (2015) notes that in the advent of the 'Arab Spring' uprisings and the 'colour revolutions', the West used the internet to affect the

consciousness of people to divide them among various lines (Kartapolov, 2015). As Thomas (2016) points out, the question then arises whether Russian scholars (like Kartapolov) are truly describing the actions of the West as they perceive them, or if they are describing their own actions through the use of a foreign model, as Soviet authors did. This question is outside of the means of this thesis to answer but is important to keep in mind while reading the following section. Secondly, Russian and Western approaches to cyber operations are conceptually different. In Russian thinking, cyber is not seen as a separate function or domain, but rather an extension of information operations. In contrast, Western intelligence agencies commonly have a separate Cyber Command division and a Strategic Communications department. Ultimately, Russian information operations encompass not only disinformation campaigns but also cyber operations and electronic warfare (Giles, 2016). Keeping these issues in mind, the next section delves into the Russian concept of 'new generation warfare' and the rising importance of information control.

New generation warfare. Many notable Russian military scholars argue that the very nature of war is changing, or has changed, as direct and indirect actions (political, economic, and psychological factors) converge (Gareev, 2010; Gareev, 2013; Gerasimov, 213). The Chief of Staff of the Russian Federation, General Valery Gerasimov, wrote with reference to the Arab Spring, that the objectives that previously were only attainable through military means now can be achieved by combining organised military violence and economic, political, and diplomatic activity - something he terms 'new generation warfare (Gerasimov, 2013).' Gerasimov's (2013) now-infamous article, 'The Value of Science is in the Foresight', has been cited as Russia's hybrid warfare playbook for the 2014 annexation of Crimea and dubbed by Western academics the 'Gerasimov Doctrine (Christensen, 2018).' This 'Gerasimov Doctrine' is a whole-of-government approach that blurs the delineations between

war and peace and emphasises the fusion of hard and soft power (Rumer, 2019). It highlights the use of hybrid tactics and states that non-military tactics are not just supplementary to the use of military force, but the preferred way to win. This way, non-military tactics may be below the threshold of activating an opponent's military response (e.g. NATO Article 5). By emphasizing a changing nature of war, disinformation campaigns and corresponding nonmilitary elements of war are increasingly perceived as important.

Increasing value of information. Chekinov and Bogdanov (2013) further converge with Gerasimov's (2010) idea of a new generation of warfare and write 'today the means of information influence reached such perfection that they can tackle strategic tasks'. They argue information operations are able to disorganise and deceive an opponent, create a desired public opinion, and organise anti-government protests (Chekinov and Bogdanov, 2011). Many other authors also agree that information is a key tool, and oftentimes a requirement for achieving strategic goals (Darczewska, 2014; Bogdanov & Gorbunov, 2009; Chekinov & Bogdanov, 2015; Chekinov and Bogdanov, 2013; Hellman & Wagnsson, 2017). The effects of information operations are, in some cases, able to become equivalent to the use of armed forces (Kartapolov, 2015; Chekinov & Bogdanov, 2010). According to some Russian military thinkers, information efforts have 'primacy in operations' while conventional military forces play a supporting role (Thornton, 2015; Giles, 2016). Although some scholars are uncertain of the final effects of information operations (Bogdanov and Gorbunov, 2009), others believe that without information superiority, a state may lose its political sovereignty, economic independence, and its status as a world leader (Chekinov & Bogdanov, 2011). Still, at other times it is not the coercive potential that achieves success, but the interaction of military and nonmilitary (political, psychological, ideological) factors (Chekinov and Bogdanov, 2010). Although there are slight disagreements among Russian

scholars to the extent of the effectiveness or the level of cruciality of information, there is still an overlying emphasis on its increasing value and the value in being able to control information.

Whether this increasing value of information is truly a new way of war, or a 'Gerasimov Doctrine' is up to debate. The 'Gerasimov Doctrine' may not be a doctrine after all, but rather a general Russian geopolitical strategy that has existed since the Cold War, and may have even preceded it (Galeotti, 2018). Mark Galeotti, a seasoned author of Russian domestic and foreign affairs, who came up with the term 'Gerasimov Doctrine' to describe Russia's hybrid operations has personally shut down the use of the term (Galeotti, 2018). This is because in Gerasimov's (2013) seminal paper, he described how Russia understands what the West did in the Arab Spring uprisings and the colour revolutions, rather than laying out a plan for Russian foreign aggression. Others argue that Gertosimov's Doctrine is really the operational component of the 'Primakov Doctrine', named after former foreign and prime minister Yevgeny Primakov (Rumer, 2019). Rumer (2019) maintains that a 'Primakov Doctrine' defines the concept of Russian foreign and defence policies; the Russian vision of a multipolar international system, insistence on Russia's primacy in the post-Soviet space, and opposition to NATO expansion. Doctrines aside, the general agreement among Russian military scholars is that information is becoming increasingly important. This is even reflected in Russian military doctrines which have included some form of information operations as a basic feature of modern war since 2000 (Darczewska, 2015). For Russia, disinformation campaigns and the control of information is regarded as imperative to achieve its political and strategic goals.

2.3.3 Western perspective on Russian disinformation.

To Western intelligence agencies, Russia is perceived as a main threat, and the leading disseminator of disinformation (Karlsen, 2019). For the first time in December 2018 the EU called out Russia as the main hostile actor of disinformation in its 'Action Plan Against Disinformation' (EU Commission, 2018). Since then, the EU's East StratCom Task Force has more than doubled its funding to address Russian ongoing disinformation campaigns and developed a 'Rapid Alert System' to coordinate member states' responses to disinformation (European Commission, 2018). NATO is also very active in addressing Russian information operations efforts. The NATO Strategic Communications Centre of Excellence in Riga and the European Centre of Excellence for Countering Hybrid Threats lead conversations in member states about Russian disinformation.

The aims and objectives of Russian disinformation are quite varied, and the literature presents many overlapping views as to Russia's overall intentions. In an analysis of 11 Western intelligence agency documents (including USA, 10 NATO countries and Sweden), the documents suggest Russia ultimately aims to ensure its regime security and power, ensure predominance in its 'near abroad,' and to secure Russian world power status (Karlsen, 2019). NATO argues that achieving Russian strategic and political goals while weakening the powers and capability of both NATO and the EU is an overall objective (Giles, 2016). Others argue Russia's overall aims are to exacerbate tensions in society, be they racial, cultural, or religious and its objectives vary depending on the region (Helmus, 2018). Here, we present 5 key themes with regards to the aim of Russian disinformation campaigns; strategic victory, control over its 'near abroad', achieving reflexive control, and weakening adversary societies. Many of these objectives are not solely reliant on disinformation operations but rather on a compound of multiple types of operations, and may overlap with other geopolitical aims.

Strategic victory. To achieve strategic victory in many parts of the world, Russian disinformation creates narratives depicting Russian as a powerful state capable of defending and advancing its interests. Another common strategy the promotion of the idea of Russia's 'ominous unpredictability.' Richey (2017) describes that these practices of changing the perception of Russia are not directed at the leaders of opposing states but rather the foreign public audience. Russia's disinformation campaigns raise the 'audience costs' in a democratic state and raises public constraints when threatening Russia with punitive measures (Richey, 2017). This is a crucial point, as Russian disinformation is not focused on state-to-state but rather state-to-people diplomacy.

Near abroad. Russian disinformation operations further aim to ensure dominance in its 'near abroad' (Karlsen, 2019). Near abroad countries include Ukraine, the Baltics, Belarus, and Moldova. Geographically, these countries serve as land barriers for Russia and increase Russia's strategic depth. Following the 2007 decision of the Estonian government to remove a Soviet-era monument, Estonia experienced a massive cyber attack on its government communications alongside the spread of disinformation in Russian media (European Parliament, 2018). Disinformation in the Baltics and surrounding areas has massively increased since then. The 'Twitter Revolution that worked too well' in Ukraine followed the annexation of Crimea by Russia (Mejias & Vokuev, 2017: 1035). Researchers argue that social media played a decisive role in Ukraine's Euromaidan movement as it exacerbated the distorted perception of public support for the protestors. There was a sense of unanimous support for the opposition which turned out to be an illusion created in part and further aggravated by well-orchestrated bot disinformation operations (Mejias & Vokuev, 2017:

1035). Events in Ukraine highlight the worst-case scenario of Russian disinformation operations.

Now-independent states who were once part of the Soviet Union contain a significant amount of ethnic Russians (Grigas, 2016). Within these countries, Western analysts argue that Russia aims to divide Russian-speaking and ethnic Russian populations between their host governments, NATO, and the EU (Grigas, 2016; Thornton, 2015). Under Russian law, these individuals are labelled as 'Compatriots Living Abroad' and are perceived as requiring the protection of Russia. Russia aims to instil a 'soft loyalty' of Russians living abroad to Russia through an emphasis on cultural, linguistic, and religious connections and leverages the historical memory of compatriots (Thornton, 2015). Leading researcher in this area Agnia Grigas (2016) documents Russian compatriot policies in which Russia attempts to provide protection and support for compatriot populations under the veil of aggressive foreign policy goals (e.g. occupation of South Ossetia in Georgia). One common tactic is to accuse Russia's near abroad countries with 'Russophobia' with the aim of reducing criticism to the Russian state to an irrational intolerance of Russia (EU vs Disinfo, 2018). Overall, achieving control (whether direct or indirect) of the information space in Russia's near abroad countries is extremely advantageous for Russia; it strengthens its strategic depth, increases geopolitical power, and clouds the judgement of international actors.

Reflexive control. Reflexive control is the practice of predetermining an adversary's decision altering key factors in the adversary's perception of the world (Giles, 2016). In public discussions, this is sometimes termed 'perception management'. As with many other objectives, it is not only achieved through disinformation but rather through a compound of operations which target decision makers and the general population in several ways (Giles et

al., 2018). One major component of reflexive control includes the act of legitimating artificially constructed 'facts on the ground' (Richey, 2017). One notable example of Russian disinformation was the 'Lisa Case' in Germany in 2016 (Giles & Seaboyer, 2018). Russian media and officials created hysteria around a false story of the alleged rape of a young girl named 'Lisa' by a Middle Eastern migrant. The political environment in Germany, as elsewhere in the EU, was tender to the issue of migration and this story served to antagonise the growing divide in Germany over the issue of migration. This resulted in a plunge for support for Chancellor Angela Merkel. This example highlights the objective of creating highly emotive and reactionary facts which people to respond to first, and question later. This example also bridges us nicely into another key aim; the weakening of adversary societies.

Weakening of adversary societies. Weakening adversary societies is one of the most highly cited aims of Russian disinformation (Giles, 2016). Russian disinformation operations attempt to achieve this objective by employing a 'divide and rule' approach (Karlsen, 2019). For example, disinformation operations spread narratives which invoke public distrust of democratic institutions. These institutions may be the mainstream media, the judicial system, or the government administration in general. The narratives further generate cynicism about politics, undermining international law and norms, and erode trust between populations and domestic and international political leaders. By promoting narratives that suggest current institutions are corrupt or unjust, Russian disinformation aims to divide the public from their representing and governing bodies.

A second approach involves polluting the information environment with an excess of information around a particular hot button issue wherein finding the truth about something becomes too cognitively draining. The most well-known example of this strategy is the

downing of Malaysian Airlines Flight MH17 (Giles & Seaboyer, 2018). Russian media were quick to pollute the media environment with multiple different stories of what happened. As a result, individuals who were not invested in the story did not dig to find information to critically assess.

Finally, Russian disinformation campaigns tend to aggravate both sides of divisive issues. For example, during Brexit, Russian disinformation operations disseminated narratives around immigration and Islam as part of a 'divide and rule' strategy (Demos, 2018). In the US, divisive issues tend to include racial discrimination, gun rights, immigration, and abortion. Studies in the disinformation landscape in the US have shown that disinformation operations tend to promote both sides of an argument, driving both sides to extremes.

2.3.4 Tactics

Many notable scholars and organisations note specific Russian disinformation tactics (Nimmo, 2015; Karlsen, 2019; Giles, 2016; Swedish Civil Contingencies Agency, 2018). A seminal paper by the RAND Corporation in 2016 summarised Russian disinformation as the 'firehose of falsehood' propaganda model (Paul & Mathews, 2016). The title of the paper refers to how the authors conceptualised the nature of Russian disinformation; multi-channel, high-volume, and lacking commitment to truth. They note Russian disinformation may contain a fraction of truth or may be wholly manufactured. Reports of fake actors portraying victims of crimes or fake on-the-scene reporting is not unprecedented and these stories tend to arouse strong emotions and act on social identities. They note Russian propaganda further exploits the peripheral cues readers use to identify reliable sources from unreliable ones; seemingly credible sources like RT and Sputnik, which visually look like reputable news

programs, disseminate disinformation (Paul & Mathews, 2016). The channels do not broadcast the same version of events and they do not shy away from changing their narrative. This is because people tend to overlook contradictions when a source considers a different perspective, making it seem like it has given a topic greater consideration

We would like to point out that within many Western publications, there is a tendency to conflate Russian disinformation with falsehood. The RAND paper in the previous section is one example that highlights this problem. Although Paul and Mathews (2016) acknowledge that Russian disinformation may contain a fraction of truth, they overemphasise the 'falsehood' dimension. Furthermore, the majority of efforts against disinformation have been 'debunking' or fact-checking initiatives. For example, the EU's East Stratcom Task Force created EUvsDisinfo, an online portal in which Russian disinformation stories are debunked. Other notable debunking initiatives include Ukrainian 'StopFake', Lithuanian 'Debunk', Atlantic Council's 'Disinfo Portal', the European Values think tank's 'Kremlin Watch' programme, and 'EU Fact Check.' Unfortunately, some of the most pervasive disinformation campaigns contain fractions of truth, or distort the context of information. These initiatives are a good start, but they only put a bandage over a major wound. This is why more research is needed into less semantic content-dependent disinformation measures and manipulation.

It is important to note that Russian disinformation campaigns do not work in a vacuum. They often support cyber operations (e.g. Macron email hack), military operations (e.g. Donbas in Ukraine), and/or political actions (e.g. criticism of EU sanctions towards Russia). This intricate web of hybrid operations makes it that much more effective, and harder to counter. In the next section we outline the way the current digital environment has changed the way disinformation is produced and disseminated.

2.3.4 Disinformation in a changing media environment

New digital technologies (especially social media) have greatly impacted the way disinformation is constructed, tailored, disseminated, and consumed (Čižik, 2017). Here, we characterise the 'digital age' by the presence of Web 2.0 technologies, the increasing importance of data, and the digitisation of news. Web 2.0 technologies are characterised by the notion of the 'web as a platform' and allow users to not only be consumers in their digital environment but to actively participate in it (Helmond, 2015). Social networking sites, like Twitter, are an enormous chunk of the 2.0 technologies. Defined by several broad characteristics social media is a platform in which users must create and interact with content as the modus operandi. Users create an account or some other identifying item and engage with other users via the creation of, engagement in, and /or dissemination of some sort of content.

Parallel in this social media universe is the constant, and far-reaching collection of user data; this include location, engagement with other users, age, photos, number of ad clicks, and a countless mirage of other seemingly inconsequential data points. However, like a mosaic, once these data points are combined, they create a wealth of inferred information about users' interests, habits, and the user themselves (e.g. age, location, political leaning, if not already disclosed). This mosaic of data is incredibly valuable to advertisers and companies looking to gain information about potential consumers. The increasing focus on the importance of data and information flows is termed 'dataism' in some academic circles and warns of the possibility of an over reliance and overconfidence of predictive algorithms (Dijk, 2014; Lohr,

2015). The user's cost to use these social media sites ends up being the data they produce about themselves. This exchange speaks to the enormousness of the data collection and retention industry and its impact on not only users but traditional media and governing bodies.

Alongside Web 2.0 and data, the digital age is also characterised by the digitisation of news (Casero-Ripollés, 2012). In the UK, over half of the population with internet access use social media as a form of news (Wakefield, 2016). Traditional news outlets have been slow to capitalise on the emerging data/advertising nexus. As users receive freely available news in favour for a bit of data, traditional news loses out. Whereas news sites can offer their content for free while being funded through advertising revenue, traditional news has been slow to adopt or integrate this type of model which make digital news so economical and efficient (Bakir & McStay, 2018). Traditional journalism has suffered from declining paying audiences as most current digital advertising revenue goes to five technology companies, four of which (Facebook. Google, Yahoo, and Twitter), integrate news into their feeds (Bakir & McStay, 2018). In the next section we outline how the digital age, and those technologies enable the construction, dissemination, and consumption of disinformation.

Truth verification. The digitisation of news has produced a plethora of news sources; online newspapers and magazines, Wikileaks, think tank publications, blogs, discussion threads, and social media to name a few. Citizen journalists and bloggers are now able to share their experiences and newsworthy events they witnessed first-hand (Edson et al., 2018). Journalists who initially treated social media as another platform to promote their stories, now use it to actually break stories and interact with audiences (Edson et al., 2018). Being simultaneously connected to users all over the globe, a sensational story - whether true or not - is able to

quickly generate reactions and following. Where a traditional media organisation is held accountable to tell the truth, stories on social media exist frequently without any reliable information sources (Edson et al., 2018). And here is the first 'enabler' - the lack of a verification body. As users are their own publishers, and are incentivised to create content worth sharing, there is no requirement to hold content on social media to a truth-standard. As such, disinformation is enabled by the advertising-centric model. As emotive, polarising content generates more clicks, this type of content is more readily produced.

Cheaper, faster, better? Advertising on social media has become incredibly cheap and easy. Further, as the user is assumed to be a publisher of public opinion, it is very easy for an actor to appear as a genuine user online. The requirements to get online are very minimal and usually an email and an internet connection. Moreover, the mere volume of information disinhibits users from finding the most accurate information as the immediacy of social media reduces opportunities for users to deliberate their information content analytically. As content is competing for clicks, disinformation oftentimes has a leg up as the 'first to the scene' reporting on major events, without needing any verification.

Plausible Deniability. Particularly relevant to the dissemination and consumption of Russian disinformation on social media is the problem of plausible deniability. Disinformation actors can easily mask themselves on social media to a point where it is difficult to distinguish genuine actors from those propagating disinformation. Russian actors have greatly benefitted from plausible deniability as Russian 'patriotic hackers' conduct the state's disinformation operations (Calamur, 2017).

2.3.5 Disinformation on Twitter

Twitter is one of the main social networking sites where the digital age forces converge; Web 2.0, data, and digitised news. A wide range of political actors are increasingly using social media as a form of policy and political communication, and because of its popularity and pervasiveness, it is ripe with disinformation. The Twitter platform is one of the more transparent social media platforms to allow researchers to access their API. Twitter also publishes frequent datasets on disinformation activities of foreign actors, available to the public. Accordingly, these factors render Twitter very beneficial to study disinformation. Although disinformation campaigns on Twitter cannot be entirely separated from traditional state-led media disinformation campaigns, the present study will focus on Russia disinformation efforts on Twitter.

Due to the above reasons, Twitter is a very worthwhile platform to study disinformation. Under Twitter's terms of service, anyone may create a Twitter account and interact with a global audience. The most common agents of influence on social media (especially Twitter) are bots and trolls. A bot is automated software written to 'gather information, make decisions, and both interact with and imitate real users online (Woolley, 2016, p. 1).' Security experts believe that bots generate more than half of all traffic online (Woolley, 2016). Social bots specifically engage in direct communication with human users on social media platforms; when their actions are deployed by political actors to subtly manipulate public opinion they are termed political bots (Woolley, 2016). Bots may work to give false impressions of popularity online or they may attack, hijack, or alter discourse on social media. One known technique is 'astroturfing' in which bot actions are concealed to make it seem that activity is genuine and/or supported by grassroots initiatives. In the past, state-

sanctioned Russian bots have been deployed to promote regime ideals and combat anti-regime speech (Helmus, 2018).

A troll is a fake social media account operated by a human; Russian state-sponsored trolls are frequently recruited to distribute the message of Russia's political leaders online (Aro, 2016; Helmus, 2018). In 2013, a Russian investigative journalist uncovered a now infamous 'troll factory' in St Petersburg called the Internet Research Agency (IRA) (Mejias & Vokeuv, 2017; Aro, 2016). Each working troll is required to produce hundreds of comments during a 12-hour shift targeting ordinary citizens, politicians and other public figures (Aro, 2016). Bots and trolls are generally deployed by governments and political actors during elections or moments of country-specific political conversation. The most powerful computational disinformation efforts are those where bots and trolls (automation and human curation) are working together (Morgan, 2018).

In addition to trolls and bots disinformation campaigns can be supplemented by hackers who conduct cyber operations such as defacing websites, denial-of-service attacks, and extraction of secret information (Helmus, 2018). In the past, when cases of Russia-linked hacking occur Russia promotes the idea of 'patriotic hackers'; individuals who identify with the Kremlin's goals and conduct cyber operations without any higher command (Helmus, 2018). There are numerous examples of Russian hacking, leaking, and even the insertion of fake information to troves of dumped documents online (Morgan, 2018).

One more important agent of influence, albeit it could be considered non-malign, is users of social media themselves. Users who benefit from sensationalised stories on Twitter may spread disinformation where it fits into their agenda, be it the far-left or the far-right.

Research has shown that although both groups have engaged with Russian trolls on Twitter, conservatives engaged to a much higher degree (Hjorth & Adler-Nissen, 2019). Other users who may engage include state-affiliated accounts and state-controlled media-affiliated accounts (such as the Twitter accounts of RT and Sputnik). Apart from agenda-driven users, a user may also unintentionally spread and engage with the content of disinformation - which is the ultimate aim of trolls. The dissemination and engagement with disinformation enables it to reach the online public sphere and makes the content seem as a legitimate discussion by the population.

This chapter outlined both the Russian and Western perspective of disinformation. According to Russian military scholars, information is a critical asset in war and with dominance over the information realm, strategic victory is unlikely. Among the Western scholars, Russian disinformation aims to achieve strategic victories, reflexive control, ensure dominance in its near abroad, and weaken their adversary societies. Disinformation in the modern digital age is exacerbated by technologies like social media. Without a verification body, news stories are under no obligation to report on the facts. Compared to traditional forms of media, spreading information on social media is cheaper, faster, and gives an actor plausible deniability. In the next chapter we introduce the concept of emotion and provide a review of the literature.

Chapter 3 Emotion

As with defining disinformation, defining emotion is also difficult. Discussions of emotion

tend to depend on the perspective one takes. As such, the present study approaches emotion

through a functional perspective argues that emotions are adaptive, intentional, and solutions

to problems (Scherer, 2009).

3.1 Defining Emotion and Theoretical Considerations

Although there has been an increasing interest into emotion and the way emotion impacts

decision making, there is profound disagreement on how emotion is best defined (Izard,

2009). Due to this difficulty, we will list five defining characteristics of emotion that are

generally agreed upon within the literature (Scherer, 2005; Nabi, 2010). These characteristics

can be thought of as the processes that need to occur for something to be considered an

emotion. Firstly, emotion involves a cognitive appraisal - a conscious or unconscious evaluation of an event, which may be external or internal to the individual. An example of an external event may be the loss of a job and an internal event may be the recalling of specific memories. Within this first characteristic there is an underlying assumption that the event in question needs to be relevant to the major concerns of the individual. This follows common sense as we do not generally have an emotional reaction to things or people we do not care about. Secondly, emotion involves the physiological component of arousal, such as changes to heart beat, breathing, or sweating. Third, emotions should motivate action, like a change in attention or stopping a current action. The fourth component of emotion involves motor expression, such as changes to facial expressions (smiling) or changes in one's voice. Finally, emotion involves a change in a subjective experience - we start 'feeling' happy or sad. Accordingly, we use Scherer's definition of emotion: 'a cultural and psychobiological adaptation mechanism which allows each individual to react flexibly and dynamically to environmental contingencies (Scherer, 2009, p.1).' Stemming from this evolutionary perspective, emotions are often high in intensity and short-lived because they must motivate an individual to quickly adapt their behaviour and attention to their surroundings (Holm, 2012; Hudlicka, 2011).

As emotions are increasingly studied in disciplines outside of psychology, inconsistencies arise even among these five characteristics. For example, Munezero et al (2014) discount the cognitive appraisal necessity, stating that emotions are not always cognitive appraisals, but may exist for no apparent reason. This contradicts much of psychological literature. As the imprecise use of terms is a growing limitation in interdisciplinary research, it is important to understand how emotion differs from other phenomena such as affect, feeling, and mood. Affect is defined as the positive or negative evaluation of something existing outside of

consciousness and is considered a predecessor to feelings and emotions (Munezero et al., 2014). Feelings are then a broad category of conscious affective sensations specific to an individual (Munezero et al., 2014). Mood differs from emotion by its lack of a cognitive appraisal and intensity. Moods may often occur without an apparent cause (they do not need to be linked to an event or specific appraisals), they are low in intensity, and may last over hours or even days (Scherer, 2005). While there is much overlap within the literature for defining emotion, the above five characteristics serve to highlight the phenomenon we aim to study; the presence of a cognitive appraisal, a physiological component, motivated response, motor expression, and a subjective experience (Scherer, 2005; Nabi, 2010).

3.1.1 Functional Perspective

By defining emotion as an 'adaptation mechanism' (Scherer, 2009), we approach emotion through a functional perspective. Although there are many different variations of functional approaches, the present study will focus on their shared assumptions. Most functional accounts of emotion appeal directly to evolution and natural selection as the explanation for why emotions developed in humans - they serve an adaptive function to promote reproduction and survival (Lench & Carpenter, 2018). Lench and Carpenter (2018) note two broad claims that bring together functional theories of emotion. First, emotions are elicited by particular events and situations that represented adaptive problems and second, emotions are 'organised responses to those problems that helped resolve the event or situation' (Lench & Carpenter, 2018, p.4). Emotions, then, are intentional states and solutions to problems, they do not come about on a whim and are directed toward one's environment (Lench & Carpenter, 2018). The functional approach of emotions is beneficial for the purposes of this study because of the approaches' direct linkage to behaviour and judgement. By serving an

adaptive function, emotions serve to orient someone to think or behave one way for the purpose of his/her survival. For example, fear is broadly a reaction to threat and related to avoidance behaviour when the threat is known (Brader, 2005). As emotive disinformation oftentimes aims to manipulate the judgement or behaviour of its receivers, the knowledge of likely behaviour outcomes of the emotion in question is particularly useful.

3.1.2 General approaches

Generally, there are two basic models of emotion that guide the vast majority of empirical research; dimensional and discrete theories (Nabi, 2010). Dimensional theories of emotion view emotion as a more generalised motivational state that can be represented within a continuous space defined by a small number of underlying dimensions (Fontaine et al., 2013). Typically these dimensions are valence (how positive or negative something is) and arousal (intensity) (Russel, 1980). Dimensional theories argue that boundaries for emotional categories are fuzzy and do not encompass the completeness of human emotion. In contrast, discrete models characterise emotions based on a limited number of qualitatively different, hardwired, and universal categories (Fontaine et al., 2013). The most well-known discrete model of emotion is Ekman's (1971) six basic emotions (anger, disgust, fear, joy, sadness, and surprise), which is based on consistent facial expressions of emotion across cultures. Depending on the theorist, the number of discrete emotions ranges from 6 to 14 (Scherer, 2005; Izard, 1992; Frijda, 1986; Plutchik, 1980). Plutchik's (1980) wheel of emotions is particularly interesting as he identifies 8 primary emotions which directly oppose each other; joy/sadness, anger/fear, trust/disgust, and surprise/anticipation (Figure 1). Within the wheel, the emotional intensity moves from the centre outwards. Throughout the history of discrete

emotion models, many theorists have expanded and condensed the number and type of

emotions. The present study will use Plutchik's wheel of emotions as the theoretical basis

because unlike Ekman's (1971) six basic emotions, they are not majority negative, and

Plutchik's wheel aligns with the study's methodology discussed in subsequent chapters.
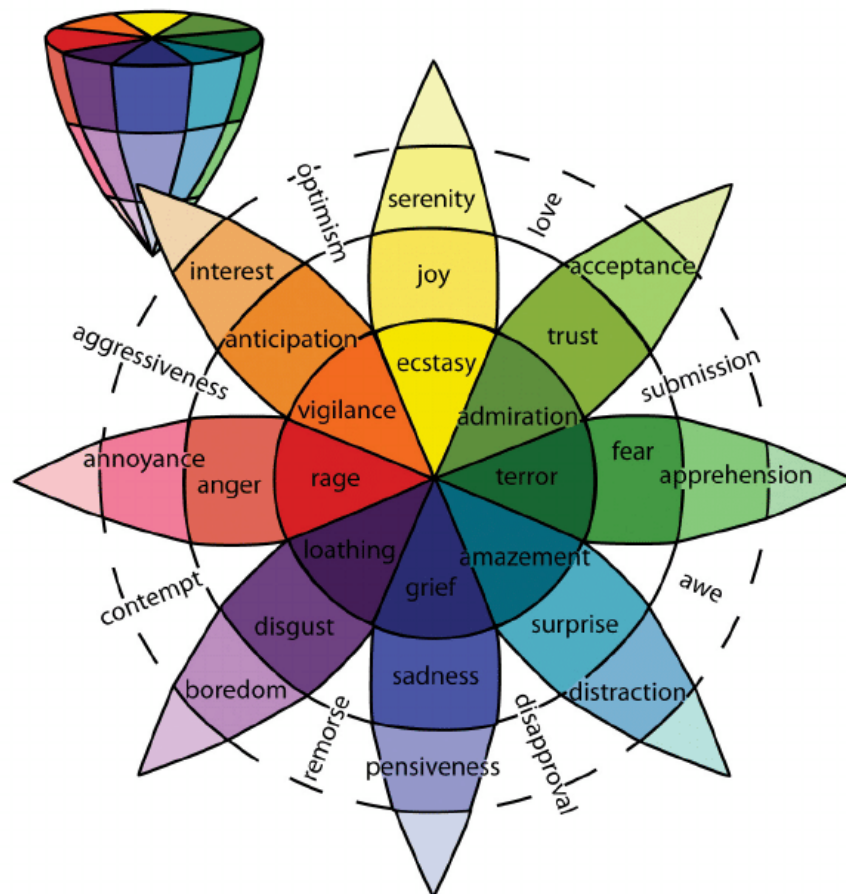
Figure 1. Plutchik's wheel of emotions

3.1.3 Effects of emotion

To gain understanding in how emotion impacts judgement, decision making, and behaviour, we will use the appraisal-tendency framework (Lerner et al., 2007; Lerner & Keltner, 2001). The appraisal-tendency framework generally states that the development of an emotional state derives from individual and subjective evaluations of a specific event/situation. The evaluations are termed appraisals, which come together in different ways/patterns to bring about emotion. The assumption is that different patterns of appraisals are linked to different emotions. Appraisals which are generally accepted include certainty, pleasantness, attentional activity, anticipated effort, control, and responsibility (see Table 1. for definitions) (Smith & Ellsworth, 1985). Patterns of appraisals and their dimensions (high-low) provide a basis for comparing and contrasting discrete emotions. For example, certainty, control, and responsibility are the central appraisals which distinguish anger from other negative emotions; anger arises from low self-responsibility for negative events, individual control, and a sense of certainty about what happened. The table below shows four discrete emotion examples and the degree of their corresponding appraisals. Within the study of disinformation, these appraisals may be useful in identifying the targeted emotion being evoked. From the corpora of recently released Russian state-sponsored tweets is the below example:

'Cops have killed 68 people in 22 days since #Kaepernick started protesting. 68 in 22 days...have no words #KeithLamontScott' (Twitter, 2018)

This tweet reveals a high degree of control (individual agency of the cops), high responsibility (of the cops), and displeasure. Therefore, we can assume that the above tweet is most likely to generate anger in contrast to other emotions. For a layperson, this act of

deriving emotion in terms of appraisals seems overly complicated as it is easily deducted that this tweet will arouse anger. However, emotions are subjective to the individual, and this framework provides a qualitative research approach that can be applied across multiple studies. Although this theory is not the most efficient way of deriving emotion from text, it serves an important explanatory purpose, as the deconstruction of emotion into appraisals provides insight into what specifically causes an emotional reaction.

Considerable evidence shows that following an appraisal pattern, distinct emotions elicit distinct 'action tendencies' (de los Santos & Nabi, 2019). An action tendency is a reactive component that triggers a certain action when an individual experiences a specific emotion. Once an emotion is evoked, its associated action tendency, which arises in response to the specific pattern of appraisal, serves to guide information processing, influencing what information is attended to and likely to be recalled, and what is ignored. The appraisals are further projected onto subsequent assessments and as such, emotions create a tendency to evaluate subsequent situations in an emotion-congruent way (Kuhne, 2014). There has been substantial theoretical and empirical work on identifying the appraisal patterns and action tendencies associated with different emotions (Fridja, 1987, Lazarus, 1991; from Nabi, 2003). This theory is extremely useful in understanding the way emotions may have an impact on decision making and behaviour. Even more so, this theory sheds light on how the manipulation of specific appraisals and emotions may impact behaviour in communication contexts. For example, after a company public relations disaster which makes its consumers very upset and angry, a response which focuses on the unpredictability of the event and situational agency theoretically would have the greatest effect on appeasement (as demonstrated in Kim & Cameron, 2011).

Table 1. Appraisal-tendency framework, adapted from Lerner & Keltner (2000)

Appraisal	Example of negative emotions		Example of positive emotions	
	Anger	**Fear**	**Pride**	**Surprise**
Certainty Degree to which future events seem predictable and comprehensible (high) vs. unpredictable and incomprehensible (low)	High	Low	Medium	Low
Pleasantness Degree to which one feels pleasure (high) vs. displeasure (low)	Low	Low	High	High
Attentional Activity Degree to which something draws one's attention (high) vs. repels one's attention (low)	Medium	Medium	Medium	Medium
Anticipated Effort Degree to which physical or mental exertion seems to be needed (high) vs. not needed (low)	Medium	High	Medium	Medium
Control Degree to which event seems to be brought about individual agency (high) vs. situational agency (low)	High	Low	Medium	Medium
Responsibility Degree to which someone or something other than oneself (high) vs. oneself (low) seems to be responsible	High	Medium	Low	High

Appraisal/action Tendency	Perceive negative events as predictable, under human control, and brought about by others	Perceive negative events as unpredictable & under situational control	Perceive positive events as brought about by self	Perceive positive events as unpredictable & brought about by others

3.1.4 Action Tendencies of Plutchik's Emotions

As the present study adopts Plutchik's (1980) wheel of emotions as its theoretical model, it is important to understand the cognitive appraisals and action tendencies of each of the eight emotions; fear, anger, surprise, anticipation, sadness, joy, disgust, and trust. Research on negative emotions (e.g. fear and anger) typically dominate the emotion literature, which is why emotions like anticipation and surprise receive little attention in the next section.

Fear. Evolutionarily, fear is a reaction to a threat and thus, central to survival (Ohman, 2008). The threat to which one feels fearful of, is an obvious (albeit may not be clearly perceived) danger that must be dealt with (Ohman, 2008). As an aversive emotion, fear motivates attempts to cope with threats to one's survival and well-being (Sander & Scherer, 2014). The action tendency for fear may include constructive action to deal with the threat, withdrawal, avoidance, and/or immobility, all depending on the individual and the situation (Sander & Scherer, 2014; Brader, 2005). Research within emotional appeals has largely focused on fear appeals, 'persuasive messages that attempt to arouse fear by emphasizing the potential danger and harm that will befall individuals if they do not adopt the messages recommendations' (Tannenbaum et al., 2015: 1178). Although there is a lot of disagreement

among the literature on the effectiveness of fear appeals and their process of action (Kok et al., 2018; Nabi & Moyer-Guse, 2013), research suggests a positive relationship between the use of fear appeals and behaviour intention, and behaviour change. In a meta-analysis of 127 studies, researchers aimed to investigate the effectiveness of fear appeals in influencing attitudes, intentions, and behaviours in a public health context (Tannenbaum et al., 2015). They found that overall, fear appeals are effective, and they become more effective when the message communicates high amounts of fear, stresses severity and susceptibility, and recommends one-time only behaviours (Tannenbaum et al., 2015). Within politics, political campaign ads using fear appeals are more persuasive, more likely to be recalled, and garner more support (Brader, 2005). Overall, fear is an aversive emotion which prompts avoidance and immobility. When used in a communication context, fear appeals are effective in influencing attitudes, intentions, and behaviours.

Anger. Anger is generally a response to goal blockage, frustration, or an unjust act and typically consists of other-person agency (e.g. someone else responsible), and high control potential (Roseman, 2018). In addition to the above appraisals, injustice and unfairness also correlate with instances of anger (Roseman, 2018). The action tendency of anger is readiness to engage in aggressive behaviours, removing the obstruction causing anger, correcting some injustice, or getting revenge (Roseman, 2018; Haidt, 2003). Anger has been revealed as a functional effect to safeguard physical survival by removing threats to the self (Keltner & Haidt, 2001), deterring transgressions (Fessler, 2010), and decreasing willingness to cooperate (Tooby & Cosmides, 2008).

Anger is extremely relevant within discussions of politics (Roseman, 2018). When anger is felt toward political actors they are perceived as responsible for one's harm or blockage of a

goal. Among many political researchers, anger is linked to political action (Roseman, 2018), and even dubbed as 'the prototypical protest emotion' by van Stekelenburg and Klandermans (2013, p.175). In a study of elections from 1984 to 2008, anger has been consistently linked to four measures of political participation; talking to people about how they should vote, attending a rally, donating to a campaign, and wearing a campaign button/sticker (Groenendyk & Banks, 2014). Interestingly, increases of anger, contempt, and disgust in the speeches of multiple political leaders in various countries has preceded acts of aggression (e.g. war, invasion, revolution) but not acts of resistance (e.g. non-violent protests) (Matsumoto et al., 2014). Anger appeals within an advertising context works very well. In a study looking at antismoking advertisements, advertisements which depicted a high level of control in participants and high anger were the most persuasive, and talked about (Ilakkuvan et al., 2018). This suggests that messages which incite a high level of anger and high degree of personal control have powerful persuasive effects (Ilakkuvan et al., 2018). Anger appeals also have been shown to be effective in promoting clean indoor air policies (Quick et al., 2009), and are widely used in political campaigning (Brader, 2006; Ridout & Searles, 2011). Ultimately, anger is appraised by perceptions of goal blockage and frustration, followed by actions which remove threats to the self, engaging in aggressive behaviours, and/or getting revenge. In communication contexts, anger appeals are shown to be effective in persuasion and communication of a message (Ilakkuvan et al., 2018; Quick et al., 2009).

Surprise. Surprise is intentional, or object-directed (one is normally surprised about something) and caused by the cognitive appraisal of unexpectedness and novelty (Reisenzein et al., 1996; Reisenzein et al., 2019). Unexpected events cause an automatic interruption of ongoing mental processes which is followed by a shift in attention (Reisenzein et al., 2019). Work on surprise is quite fragmented, as researchers study it in different contexts often using

distinctly different methodologies (Munnich et al., 2018). Surprise is often a condition for learning and cognitive development and an instigator for spontaneous causal search (Munnich et al., 2018; Reisenzein et al., 2019). Cognitively, surprise causes interruption and an attentional shift. Behavioural evidence for an attentional shift was observed in a series of eye-tracking experiments and showed that surprising stimuli attract gaze in a visual search task and events are looked at longer when they are surprising versus not (Hostmann & Herwig, 2015, 2016 in Reisenzein et al., 2019). This may be why surprising events are recalled much better, the phenomenon of 'flashbulb memories' (very detailed and vivid memories in which a surprising and consequential event occurred) exceptionally illustrates this effect.

Anticipation. Anticipation as an emotion is the least written about in the literature. Anticipation is the cognitive appraisal of an uncertain future which may be positive or negative (van Boven & Ashworth, 2007). Anticipation of positive future events is an often-cited emotion in other theories and described as hope, which is much more written about in the literature. Hope is a future-orientated emotion which motivates behaviour by focusing on the possibility of rewards and avoidance of punishments. In order for hope to exist, one must appraise a positive future outcome as possible, and some authors argue that hope can only arise from current threatening situation (Lazarus, 1999 in Chadwick, 2014). Hope appeals are messages which evoke hope and create an opportunity to take advantage of an opportunity (Chadwick, 2014). In a study looking at hope appeals of climate change, feelings of hope predicted interest in climate protection, suggesting important implications for information persuasion and disinformation (Chadwick, 2014).

Disgust. Disgust was first characterised by Darwin (1872) in 'The Expression of the Emotions in Man and Animals' as 'something revolting, primarily in relation to the sense of taste, as actually perceived or vividly imagined; and secondarily to anything which causes a similar feeling through the sense of smell, touch, and even of eyesight (p. 253).' Evolutionarily, threats of disease and infection shaped the disgust response in humans (Sherman & Haidt, 2011). It may seem that disgust only applies to material things, but it follows the law of contagion - contact with disgusting things makes one disgusting (Sherman & Haidt, 2011). The emotion itself is a mechanism for tracking negative social value, eliciting revulsion, and desires for social distance (Rozin et al., 2008).

According to some scholars, disgust may be further divided into physical disgust and moral disgust - characterised by a feeling of repulsion to moral violations such as racism, sexism, or betrayal. Although some argue that moral disgust is rather an extension of anger (Nabi, 2002), there is evidence that cases of moral disgust are discreetly disgust (Rozin et al., 2008). Moral disgust also follows the law of contagion. Contact with people who have committed moral offences (such as murder) is highly aversive, to about the same extent as similar contact with someone with a serious contagious disease (Rozin et al., 2008). Therefore contact with other people can elicit disgust. This is sometimes termed interpersonal disgust and has been characterised as an aversion to four identifiable components: strangeness, disease, misfortune, and moral taint. For example, a sweater worn once by a healthy stranger and then cleaned is still less desirable than an unworn sweater (aversion to strangeness) (Rozin et al., 2008).

Moreover, disgust serves some explanations for the concept of dehumanisation. Haslam (2006) distinguishes between two types of dehumanisation, 'animalistic dehumanisation'

which makes others less human by making them more like animals and 'mechanistic dehumanisation' which makes others less human by denying them uniquely human emotions and traits making them more like machines. Haslan (2006) suggests that disgust is the emotional reaction to animalistic dehumanisation and is likely to be felt for groups perceived as animal-like, low-status, and who are dissimilar (Rozin et al., 2008).

Joy and Trust. Joy arises from feelings where a person feels safe, familiar, requiring little personal effort. Joy occurs when people have made progress on their goals (Sander & Scherer, 2014). Joy also broadens people's attention and thinking which is thought to be supported by the 'do anything' action tendency (Sander & Scherer, 2014). Trust is a fundamental emotion, part of social and economic life. Trust is indispensable in friendship, love, family, and organisational relationships. Trust is evolutionarily valuable as it saves a lot of time to have a fair order of reliance on other people's word (Roseman, 2018).

Sadness. Sadness is often characterised by passivity and behavioural inhibition, however it serves an adaptive function and directs the challenging task of reconstructing goals and beliefs when one feels loss (Karnaze & Levine, 2018). Sadness arises in response to a loss of a valued state (e.g. bereavement, damage to valued possessions, missed job opportunities), and the perception that goal failure is irrevocable is an important component of sadness (Karnaze & Levine, 2018). Further, an appraisal of low control over the situation also commonly precedes sadness.

One common action tendency with sadness is an increase in politeness and generosity (Karnaze & Levine, 2018). Karnaze and Levine (2018) argue that this may serve to elicit reciprocal support and generosity from others or alleviate sadness indirectly by contributing

to the well-being of others. Furthermore, studies suggest suggest that sadness may both narrow and broaden information processing (Karnaze & Levine, 2018). Evidence for more narrow information processing shows that induced sadness results in a reduction in the tendency to use heuristics and make broad judgements, (i.e. relying less on top-down processes). Sadness can promote systemic, detailed, and effortful information processing and has experimentally shown to improving memory (Karnaze & Levine, 2018).

3.1.5 Conclusion

Although there is value in looking at both dimensional and discrete models of emotion, the present study will apply the discrete emotion approach. Discrete emotions allow for a more precise prediction of actions as they produce nuanced effects consistent with underlying action tendencies. Psychological research has demonstrated that discrete emotions (even those of the same valence) have different effects on perceptions of risk, attitudes, and decisions (Nabi, 2003). Further, the discrete approach is more widely used in political communication research, making it easily applicable to disinformation.

The differences and unique characteristics of each of Plutchik's (1980) 8 basic emotions suggest how different emotions may elicit different behaviour patterns. Fear, evoked when there is a threat to survival generally follows a withdrawal or avoid behaviour. When someone appraises goal blockage, or experiences an unjust act, this typically evokes anger, which follows approach behaviours like aggression. Surprise occurs after an appraisal of unexpectedness or novelty. Psychological research within surprise illustrates that memory and attention around surprising events is significantly better. Anticipation, closely linked

with more common hope is a future-orientated emotion which motivates behaviour by focusing on the possibility of rewards and avoidance of punishments. Disgust, which tracks negative social value, elicits revulsion and desire for social distance has been linked with dehumanisation. Joy broadens people's attention and thinking and trust is indispensable in friendship, family, and organisational relationships. Finally, sadness serves an adaptive function and directs the challenging task of reconstructing goals and beliefs when one feels loss (Karnaze & Levine, 2018). Understanding the properties of the above discrete emotions and their associated action tendencies helps an actor who aims to disinform.

3.2 Emotion Literature Review

While much of the literature asserts that the use of emotion in disinformation is rampant and a significant feature of persuasion and influence, there is limited research fully exploring the use of discrete emotions in disinformation. In this chapter we turn to disciplines of political psychology, marketing, and political communication to demonstrate the importance of emotion in news consumption, news dissemination, and decision making. In disinformation, emotion may manifest in many ways; text, images, and video are all means of communication that may present different emotions. For the purposes of this study, we will solely focus on emotion in text, as most disinformation analyses focus on text and it is most methodologically appropriate for the study of emotions on Twitter. This chapter outlines the use of emotion in disinformation campaigns and further describes how emotion manipulation can play a role in misleading both the general public and individual citizens.

3.2.1 The appeal to emotion within disinformation

The notion of an appeal to emotion within disinformation is widely mentioned in the academic literature (Sivek, 2018; Nisbet & Kamenchuk, 2019; Asmolov, 2018; Bennet & Livingston, 2018; Schmitt, 2018; Karlsen, 2016). There is a general consensus that disinformation campaigns actively use emotional extremity in their messages for strategic effect; to influence an audience's emotions, motives, and objective reasoning (Garret, 2017; Lin & Kerr, 2017; Allen & Moore, 2018; Asmolov, 2018). Some scholars have termed the use of emotions in a news story as 'sensationalism', aiming to trigger emotional reactions in the reader (Mourao & Robertson, 2019). Contrary to the popular saying 'if it bleeds, it leads', there isn't consensus whether audiences are more likely to respond to sensational news stories within established online news organisations (Kilgo et al., 2018). However, news stories and other imagery content on social media that strike strong emotions are more likely to be shared (Sivek, 2018; Vosoughi et al., 2018). The largest study to date of online disinformation analysed the spread of 126,000 verified true and false news stories on Twitter between 2006 and 2017. The researchers found that disinformation 'diffused significantly farther, faster, deeper, and more broadly' than accurate information, propelled by emotional reactions such as fear and disgust (Vosoughi et al., 2018).

Case studies of Russian disinformation especially highlight the use of emotion. In an analysis of 3,500 Facebook ads allegedly purchased by the Russian government, Dutt et al (2018) found that the most effective ads tended to have a less positive sentiment. Researchers noted that all 8 of Plutchik's emotions, excluding surprise, were observed to be significantly pronounced in the more effective ad campaigns (Dutt et al., 2018). Most recently, Miller

(2019) analysed tweet sentiments from 3,814 Twitter accounts associated with the Russia-based Internet Research Agency and found that the content of these tweets increasingly used aggressive language during and immediately preceding the 2016 election. These Russian disinformation cases show that emotional messages are used frequently on social media, and they are likely to be read and passed along. Unfortunately, these studies do not delve deeply into the content associated with specific emotions, but they do provide empirical evidence of increased emotionality in disinformation on social media.

In addition to the academic literature, international organisations agree that emotion is extremely important in the diffusion of information (European Parliamentary Research Service, 2019; UNESCO, 2016; NATO StratCom Centre of Excellence, 2018). RAND's now infamous report 'The Russian "Firehose of Falsehood" Propaganda Model' argues how information that arouses emotion can be particularly persuasive as it is more likely to be passed on whether true or not (Paul & Mathews, 2016). CEPA's Stratcom program which aims to expose Russian disinformation in Central and Eastern Europe further notes that an appeal to emotion is one of the most commonly used techniques (CEPA, 2019). Although these organisations have not directly studied the role of emotion in disinformation, there is consensus among the research community about the importance of emotional messaging.

While there is consensus that the appeal to emotion in disinformation is important and has an effect on the dissemination of disinformation, there is a lack of research that explores this issue beyond the detection of specific emotions (eg. Miller (2019) and Dutt et al (2018)). How specific emotions within disinformation impact individuals and the public as a whole remains understudied. In the next section we will amalgamate research from political psychology, journalism, marketing, and most importantly political communication to argue

how emotion in messages affects an individual's decision making and the distribution of information. The next section will justify the exploration of discrete emotions in disinformation as a tool of manipulation.

3.2.2 Effects of discrete emotion in messages

The presence of emotion in messages influences humans at the individual level, affecting their decision making and opinion formation (Nabi, 2003; Lecheler & de Vreese, 2019). Experiments which demonstrate this influence are conducted within communication disciplines, predominantly media framing, public health communication, marketing, and political campaigning. The use of emotional appeals, forms of communication which intend to elicit an emotional response, are widely noted within public health education (Tannenbaum, 2015) and political campaigning (Brader, 2006) to have an effect on viewers judgments and behaviours. Brader (2006) argues that within the context of politics, emotional appeals 'can influence the participation and choices of viewers in distinct ways [emphasis mine]' (Brader, 2006: 13). In this section, we will demonstrate the effects of emotion in information through research in media framing.

Framing. Framing, 'the way in which information is presented, or the perspective taken in a message' also influences the responses individuals will have to an issue at hand (Nabi, 2003). Framing does not have one single agreed upon definition, although some scholars specify several criteria that must be met. A frame must have identifiable conceptual and linguistic characteristics, should be commonly observed in journalistic practices, must possess representational validity, and must not be merely a figment of the researcher's imagination

(Cappelle & Jamieson, 1997: 47, 89; from Lecheler & de Vreese, 2019). Frames may be issue-specific or general, and some of the most common generic frames include 'conflict', 'human interest', 'attribution of responsibility', and 'morality' (Semetko & Valkenburg, 2000; from Lecheler & de Vreese, 2019). Unfortunately within disinformation, news frames are seldom explored and occur within the literature as part of a larger case study (e.g. the 'War on Terror' news frames in Lewandowsky et al., 2013). However as disinformation is a type of information, and frequently imitates news stories, general news framing theory is easily applicable to disinformation.

3.2.3 Emotion-as-frames model

News frames have been shown to affect the interpretation of issues, cognition, attitudes, and behaviour (Lecheler & de Vreese, 2019; Chong & Druckman, 2007). While much of the literature on news framing effects has focused on cognitive effects, recent research has documented the existence and influence of emotionally evocative frames (Nabi, 2003; Kuhne, 2014; de los Santos & Nabi, 2019). Nabi's (2003) emotion-as-frames model (EFM) proposes that emotions are themselves frames, incorporated into messages via information associated with particular appraisal patterns. Her model draws heavily from the appraisal tendency framework discussed in the previous chapter and generally follows two steps: appraisal to emotion and emotion to action tendency (Lerner & Keltner, 2000; Nabi, 2003). For example, fear may be elicited through a news story if there is a perceived threat to an individual's safety or if the story lacks a causal factor for an event. Once the emotion associated with the specific appraisal pattern is evoked, the emotion's associated action tendencies guide information processing, influencing what information is accessible, attended

to, and relevant for attitude formation and decision making. Kuhne (2014) put forward a more complex model incorporating more cognitive elements alongside the emotional aspects of framing effects. In addition to Nabi's (2003) two steps, Kuhne includes a preliminary step in which news frames produce certain appraisals. Further, he includes moderators of cognitive framing effects such as prior knowledge and attitudes towards specific issues. Although cognition and emotion are difficult to separate as they are two halves of the same coin, for the purposes of this study we will use Nabi's (2003) emotion-as-frames model. Because the focus of this study is an exploratory analysis of emotion in disinformation and not an analysis of framing effects, the EFM is most appropriate due to its simplicity and focus on discrete emotions.

Recent empirical framing studies support the emotion-as-frames model (EFM); news frames which elicit emotions and corresponding emotional reactions are found to influence opinion formation (Nabi, 2003; Kim & Cameron, 2011; Goodall et al., 2013; Hasell & Weeks, 2016; Kühne & Schemer, 2015). Most emotion framing experiments follow a general design. First, participants in an experimental group engage in some sort of emotion-inducing activity, in most studies they read a passage that is supposed to evoke a discrete emotional response. Control participants will engage in some sort of neutral activity. Afterwards, participants respond to an emotion manipulation check to make sure that the emotion in question is indeed experienced by the participant. Next, participants in both the experimental and control groups engage in an activity (generally reading another passage) and respond to questions regarding their specific judgements and decisions. As a test of the EFM, attitudes about drunk driving were affected by an experimental emotion manipulation; angry participants were more likely to favour retributive policies while fearful participants preferred protective solutions (Nabi, 2003). Subsequent studies have found significant differences in information

processing and opinion formations following the manipulation of different emotional frames (Kühne & Schemer, 2015; Kim & Cameron, 2011). Most often, negatively valenced emotional frames are explored (e.g. anger, fear, sadness), but researchers have found that empathy (Gross, 2008), humour (Skurka et al., 2018), enthusiasm (Lecheler et al., 2013) and hope (Nabi & Prestin, 2017) as emotional frames possess varying effects on policy preferences and opinion formations.

3.2.4 Conclusion

While much of the literature asserts that the use of emotion in disinformation is rampant and a significant feature of foreign influence, there is limited research fully exploring the use of discrete emotions in disinformation. Russian disinformation, especially on Twitter is rife with emotional content. In news framing theory, emotion evoked from a message may itself be considered a frame, and as such, the emotion-as-frames model is applied to the study of Russian disinformation.

Chapter 4 Russian Disinformation and the Emotion-as-Frames Model

4.1 Tweets as Emotions

As this study analyses Russian disinformation on Twitter, we need to determine whether tweets can be emotional frames. Empirical evidence of the emotion-as-frames model has

generally manipulated 'news' text of around 300 words as their emotional frame (e.g. Nabi, 2003; Kim & Cameron, 2011). However recent studies show that emotional frames also exist in shorter text, such as tweets. In case studies of specific events such as the Vancouver riots and the Egyptian uprising, researchers outline how emotion and news framing often overlap, and some news frames exist as emotions themselves (Burch et al., 2015; Meraz & Papacharissi, 2013; Harlow & Johnson, 2011). Chilek (2017) found that tweets that use fear, sadness, surprise, or trust positively influence 'pass along' behaviour. A further study demonstrates that positive and negative emotions are able transfer to other individuals on social media (Kramer et al., 2014). In a now ethically-dubious experiment, Kramer et al (2014) manipulated the levels of positive and negative content on individual Facebook newsfeeds of 689K individuals. They identified that when positive expressions were reduced, people produced more negative posts and fewer positive posts; and vice-versa when negative expressions were reduced. Although tweets are 140 characters long and seemingly inconsequential, research suggests that tweets containing emotion do affect their readers. Accordingly, the emotion-as-frames model is a useful tool to identify emotions in Russian disinformation tweets and how this may manipulate opinion formation and decision making.

4.1.1 Operationalising disinformation

The present study defines disinformation as misleading information which has the intentional function to mislead. To cover the full extent of this definition, tweet information should be misleading, *and* come from a source which intends to mislead. However in practice Russian disinformation, especially on Twitter, is difficult to detect and label. Tweets which are misleading to one user, may not be towards another. Rather than trying to normatively judge what information is objectively misleading, the present study will focus on user intents to mislead.

As such, to operationalise disinformation on Twitter the present study will assess tweet information from a source who's intentional function is to mislead - the Internet Research Agency (IRA). The IRA is a Russian company of 'professional trolls' based in St. Petersburg (US National Intelligence Council, 2017). As early as 2014, the organisation's strategic goal was to 'sow discord in the U.S. political system, including the 2016 US presidential election (US v. Internet Research Agency, 2018).' The IRA sought to, among other things, to conduct 'information warfare against the United States of America through fictitious U.S. personas on social media platforms and other Internet-based media' (US v. Internet Research Agency, 2018). According to the February 2018 indictment by the Special Counsel's office the IRA was 'organised into departments, including a graphics department, a data analysis department, a search-engine optimisation department, and an information-technology

department (Spangher et al., 2018). The US intelligence community reports that the IRA operations were ordered directly from the President of the Russian Federation and was also in part focused on undermining Western democratic institutions (US National Intelligence Council, 2017). The overall assessment of the US intelligence community is that the IRA is tied to the Russian President through its financier who is a 'close Putin ally with ties to Russian intelligence (US National Intelligence Council, 2017).'

Tweet information from the IRA has been previously used as a proxy for studying Russian disinformation (Miller, 2019; Zannettou et al., 2019; Xia et al., 2019), Russian 'fake news' (Badawy et al., 2018) and Russian information operations (Arif et al., 2018). Since Twitter has made IRA content openly available to researchers, the academic community has begun to characterise the IRA and its disinformation strategies.

4.1.2 IRA Strategy on Twitter

IRA tweets have been found to be linguistically distinct from the tweets of genuine users on Twitter, and tweets are carefully constructed in an intentionally deceptive manner (Boyd et al., 2018). A lot of studies note one general strategy of IRA or other Russian-state sponsored accounts to pretend to be genuine local news outlets (Farkas & Bastos, 2018; Zannettou et al.,

2018). In an IRA dataset released by Twitter in 2017, the most prevalent topics covered were

local affairs (encompassing news related to specific cities or municipalities), politics, crime,

economy, and entertainment (Farkas & Bastos, 2018). The local news content had a bias

towards news items in the crime section and issues surrounding public safety (Farkas &

Bastos, 2018). Researchers further noticed that the fake local news outlets were biased

towards amplifying issues around public security, (particularly crime) but also fatal accidents

and natural disasters (Farkas & Bastos, 2018). In a dataset released by Twitter in 2018,

researchers analysed tweets containing hyperlinks and found that the IRA relied heavily on

genuine local news sources when sharing content (Yin et al., 2018). 30% of all URLs the IRA

posted were linked to local media outlets (Yin et al., 2018). The researchers note that this

may be due to the fact that Americans tend to trust local media more than any other media

type.

Topics covered. Miller (2019) looked at the topics that were discussed in tweets from 3,814

Twitter accounts associated with the Russia-based Internet Research Agency (Miller, 2019).

In the 2018 dataset, Miller (2019) found that the majority of topics referenced political

concepts. Of the topics mentioned, it is unsurprising Trump and his campaign was featured

heavily in this data (just under 20% of the data). However, Trump-related messages were not

the majority of topics indicating that the common perspective of IRA meddling as being

solely pro-Trump is a simplified outlook (Miller, 2019). Interestingly, the number of topics

related to Clinton and Trump were relatively equal, although Clinton-related messages

contained overwhelmingly negative word content (Miller, 2019). The IRA tweet corpus

contained a host of both conservative and progressive political topics, but there was

significantly higher levels of conservative hashtags use than liberal/progressive hashtags use

(Miller, 2019). Other topics of interest contained topics related to refugees and border issues,

former President Obama, black lives matter movement/black power movements, Christain

language, and topics concern Russia and/or Putin (Miller, 2019). Miller's (2019) topic

models suggest that the corpus was more anti-Hillary than pro-Trump.

4.2 Emotion in IRA Tweets

Previous research on emotionality and sentiment of IRA tweets suggests that it exists, but not

to an enormous extent. In the 2017 dataset, Farkos and Bastos (2018) manually coded IRA

tweets for high emotionality and found only 10% of the tweets comprising of highly

emotional statements. In another study on the 2017 dataset, researchers sourced Amazon's

Mechanical Turk and had individuals rate IRA-linked tweets along a dimension of emotional

intensity (neutral, low, medium, high, and very high) (Spangher et al., 2018). The researchers

found nearly 42% of tweets to be labelled 'neutral' and approximately 20% to be labelled high or very high. However, when looking at the tweets that received more traffic on Twitter, content with medium and high emotional valence which received the most traffic (Spangher et al., 2018). The reason for such low emotionality may be that some IRA Twitter accounts were strung to garner credibility and engagement. Researchers call them 'sleeper' accounts as they are politically inactive, and once these accounts are needed, they may become more emotionally intense and political, already having the followers and credibility as a genuine user. Still, one fifth of the corpus generating high or very emotional content is still decent.

To the best of my knowledge, only one other study has specifically analysed the emotion of IRA tweets. Miller (2019) conducted a temporal emotion analysis of the IRA-linked tweets and found that the IRA increasingly used aggressive language on social media during and immediately preceding the 2016 election. During the months before the election, the overall number of aggressive words in the corpus increased in quantity (Miller, 2019). Further, the number of anger and fear-associated words in the corpus was highly concentrated around the window of the election (Miller, 2019). From Miller's (2019) study, we know that anger and fear exist in IRA tweets. However, Miller (2019) did not report on any other emotions, or the degree of their existence within the IRA tweet corpus, just their change over time. Pushing forward Miller's (2019) research, the present study will analyse the presence of each emotion

that exists in Plutchik's wheel of emotions (fear, anger, sadness, surprise, trust, disgust, joy, and anticipation). As such, we will attempt to answer the research question:

RQ1: What emotions exist in Russian disinformation?

RQ2: Are certain discrete emotions more prevalent than others in Russian disinformation (IRA) on Twitter?

Studies from framing theory disproportionately focus on fear and anger. Although we have no specific predictions with regards to which emotions may be most common, having analysed the presence of specific emotions, it will be interesting to note whether specific themes emerged within discrete emotions. This would take the research a bit further, and could provide insight into the framing of political issues by the IRA. This leads us to our second research question:

RQ2: Are certain topics/themes discussed via specific discrete emotions?

Under the emotions as frames model, an appeal to specific emotions within specific topics/themes would suggest that Russian disinformation actors engage in a strategy tied to emotional appeals.

Chapter 5 Research Design and Methodology

This study addresses the presence of emotion in Russian disinformation through a mixed

methods design. More specifically, we used an explanatory sequential mixed methods design

in which tweets were assigned to emotion categories through a quantitative phase.

Subsequently, themes within the emotion categories were produced by a qualitative content

analysis. This method was particularly useful for the present study because it allowed us to

contextualise the emotion categories determined by the quantitative analysis.

In this chapter, we address the research design, provide an overview of the data material, and

discuss both the quantitative and qualitative research methodologies. Lastly, we present the

limitations of our analysis, how we attempted to alleviate those limitations, and identify the

study's strengths.

5.1 Research Design

The aim of the present study is to explore the presence of emotion in Russian disinformation

on Twitter through three related research questions:

RQ1: What emotions exist in Russian disinformation (IRA) on Twitter?

RQ2: Are certain discrete emotions more prevalent than others in Russian disinformation

(IRA) on Twitter?

RQ3: Do specific themes or topics emerge from discrete emotion categories in Russian

disinformation (IRA) on Twitter?

To address the study's three research questions, we used an explanatory sequential mixed

methods design. This research design is characterised by an initial quantitative phase,

followed by a qualitative phase, and the two methods are integrated during an interpretation

of the results (Creswell & Clark, 2018: 270-284; Kroll & Neri, 2009). We characterised the

emotions of tweets via the quantitative phase (answering the first two research questions) and

followed up with a qualitative analysis to explain and contextualise the quantitative results

(answering the third research question) (Figure 3). Before outlining the methodology of the

two phases, we will briefly give an overview of the data material.
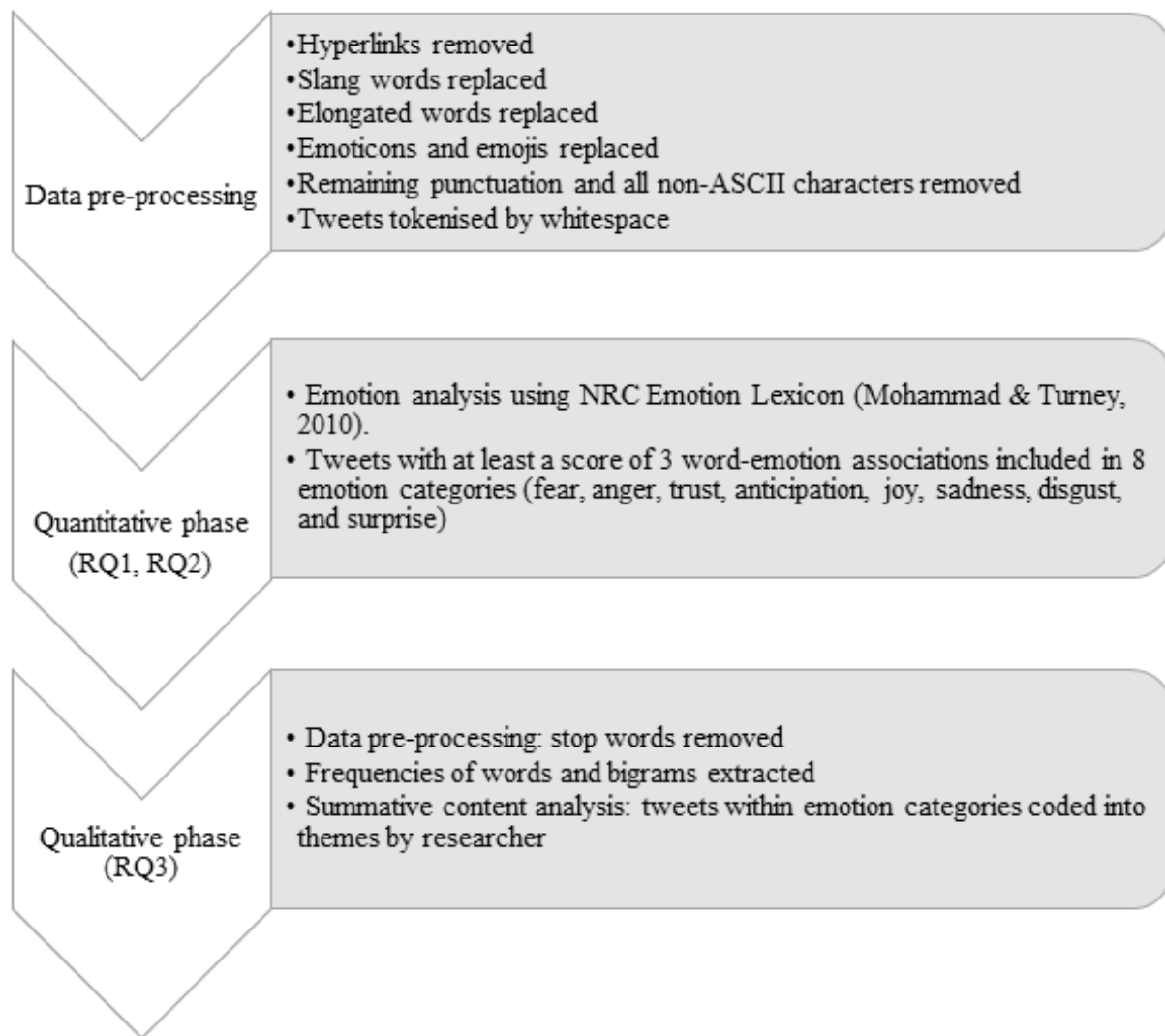
Figure 3. Present study research design

5.1.1 Overview of the data material

The present study used a dataset of tweets from IRA-made Twitter accounts, as identified by

Twitter in October 2018 (Twitter Elections Integrity, 2019). The dataset is considered to be

the most comprehensive empirical record of Russian disinformation troll activity on social

media to date (Kim et al., 2019). Unfortunately, Twitter has not provided the collection and

selection methodology of the IRA-labelled tweets. During testimony to US Congress, Twitter's Acting General Counsel could only reveal that IRA-linked accounts were identified through information from 'third party sources' (Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism, 2017). Therefore, the lack of clarity to the identification of the IRA tweets is an unavoidable limitation of the present study. Given Twitter's technical expertise, reputation, and the use of the dataset in previous academic works (Zannettou et al., 2018; Kim et al., 2019; Miller, 2019), we assess the dataset is reliable for the present study.

The dataset itself consists of 9 million tweets from Twitter's Elections Integrity (2019) campaign. Out of over 9 million tweets, only IRA tweets posted in English were analysed, relying on the metadata accompanying each tweet in the dataset to filter out non-English tweets. Duplicate tweets and retweets were removed as we assessed the content of unique tweets, and repetitions of tweets may have skewed the results. This pre-processing measure resulted in 1,599,367 (1.6 million) tweets.

5.2 Quantitative Phase: Detection of Discrete Emotions

To detect discrete emotions of the IRA tweets we used a variant of sentiment analysis, called emotion analysis. Generally, there are two approaches to automatically detecting emotion

through text; a machine learning and a lexicon-based approach (D'Andrea et al., 2015; Taboada et al., 2011). The present study employed a lexicon-based approach in which a previously made emotion lexicon was used to compare the words in tweets. Emotion lexicons are lists of words with prior associations to discrete emotions (in our case, Plutchik's 8 basic emotions). The idea behind the lexicon-based approach is quite simple; words in the research text are compared to words with pre-assigned emotions from a lexicon. Since the creation of a lexicon is the central part of the lexicon-based approach and vastly influences the analysis, its selection is crucial to our investigation.

The lexicon we adopted is the National Research Council of Canada's (NRC) Emotion Lexicon developed by Mohammad & Turney (2010). The NRC Emotion Lexicon consists of 14,182 words and their accompanying emotion associations (anger, trust, fear, joy, sadness, surprise, disgust, and anticipation) (Mohammad & Turney, 2010; 2013). The lexicon applies Plutchik's 8 basic emotions because unlike Ekman's six emotions, Plutchik's emotions are not composed of mostly negative emotions (Mohammad & Turney, 2013). To obtain the text-emotion associations, researchers paid participants via Amazon's Mechanical Turk to associate a word with a discrete emotion (Mohammad & Turney, 2010). It was possible for words to be associated with more than one emotion. For example, the word 'advice' was only associated with the emotion 'trust' whereas the word 'avalanche' was associated with the

emotions 'fear,' 'sadness,' and 'surprise.' Figure 4 illustrates the portion of the total number

of words associated with each emotion and Table 2 shows the exact number of emotion-

related words in the dataset. As the numbers are not even, we took into account the

percentage of emotion-related words in the NRC lexicon when analysing the frequency of

emotions in the IRA tweet dataset. For a detailed description of the methods used to obtain

the word-emotion associations, please see Mohammad and Turney (2010; 2013).

Figure 4. NRC Emotion Lexicon visual representation

Table 2. Exact number of words associated to each emotion in the NRC Emotion Lexicon

	Fear	Anger	Trust	Sadness	Disgust	Anticipation	Joy
Number of words	1483	1250	1234	1195	1060	842	691
Percent of the total lexicon	10.5%	8.8%	8.7%	8.4%	7.5%	5.9%	4.9%

5.2.1 Data Pre-processing

We adopted several procedures to the tweet dataset before carrying out emotion analysis (Appendix A). First, each tweet was broken into 'tokens' comprising of words and other elements (such as hashtags and URL links), delineated by whitespace. Text on social media is a special challenge to tokenise because it includes many special elements (e.g. mentions and retweets), misspellings and abbreviations. As such, these considerations were dealt with accordingly in the methodology. For the following data pre-processing measures and subsequent emotion analysis we used R (3.6.0) and RStudio (1.2.1335) for Mac OS X.

First, hyperlinks within tweets were removed as they do not contain any significant sentiment. Next, we used the R package 'textclean' developed by Tyler Rinker to replace abbreviations of slang words with their subsequent meanings (e.g. 'omg' replaced with 'oh my god'). The function, 'replace_internet_slang', draws from a dataset of 175 slang terms and their meanings from Possel (2018). Next, we used the 'replace_word_elongation' function to fix words with elongated letters (e.g. 'whyyyyy' to 'why'). As this is a slightly more complex function and is an augmented form of Armstrong & Fogarty's (2007) algorithm, the full explanation of it is accessible in Rinker (2018). As emoticons and emojis are frequently used to show emotion, they were not removed. Here, emoticons and emojis

were replaced with their word equivalent with functions from the 'textclean' package,

'replace_emoji' and 'replace_emoticon' (e.g. :(was replaced with 'frown') (Rinker, 2018).

Finally, the remaining punctuation and all non-ASCII characters were removed.

5.2.2 Emotion Analysis

Each token within a tweet was compared to each word in the NRC Emotion Lexicon

(Appendix A). If a word within a tweet matched an emotion (from the NRC lexicon), that

resulted in a tweet score of '1' for the specific emotion. To increase the validity of the

emotion analysis, only tweets with a score of 3 or more matching tokens were labelled under

a specific emotion (See Table 3 for an example).

Table 3. Example of the emotion tweet labelling procedure

	anger	antip*	disgust	fear	joy	sad	surpr
'arizona court reinstates death sentence for 1993 killing'	4	3	2	4	0	3	

Within this tweet, 4 words were associated with anger ('court', 'death', 'sentence', 'killing'), 3 with anticipation
and so on. As the tweet needs at least 3 word-emotion associations to be labelled under an emotion category,
labelled under the emotions anger, anticipation, fear, and sadness.
*anticipation

To answer the first research question, we assessed whether tweets contained at least a score of 3 word-emotion associations for each emotion category. To answer the second research question, we standardised the numbers of tweets within each emotion category as not all emotions are represented equally in the NRC Lexicon. For this reason, the number of tweets within each emotion was divided by the total number of words within the emotion category in the NRC Emotion Lexicon (Figure 4).

5.3 Qualitative phase: Summative content analysis

Having established the emotion categories, we conducted a summative content analysis of each category to answer the third research question:

RQ3:Do specific themes or topics emerge from discrete emotion categories in Russian disinformation (IRA) on Twitter?

Compared to the previous quantitative phase, this phase was less structured and more exploratory in nature. Summative content analysis began with searches for the most commonly occurring words (and in the present study, bigrams) (Hsieh & Shannon, 2005). The purpose of detecting bigrams (pairs of consecutive words), was to get a better picture of the context of the most frequent words.

Before we looked up the frequencies, we pre-processed the data once more (Miner et al., 2012: 46). Tweets were again tokenised by using whitespace as token delimiters (Miner et al., 2012: 47). Second, within each emotion category stop-words (words which carry a connecting function in the text such as prepositions and articles and do not generally have sentiment applied to them) were removed (Kolchyna et al., 2015). Some of the most common stop words included 'is', 'at', 'the', and 'on'. Next, the top ten most frequent words and bigrams were extracted from each emotion category.

To generate common themes within the emotion categories we utilised text mining techniques in RStudio to extract the most commonly used words and bigrams. Here we made the bag-of-words assumption in which the order of the words in tweets did not matter (Miner et al., 2012: 45). This one of the main strengths of text mining because we were able to use all of the words in each emotion category. Tweets within emotion categories were filtered through frequent words and bigrams and then coded by the researcher into themes. Themes emerged from coding multiple tweets and they summarised the bulk of each emotion category. This step allowed the researcher to address incidents of sarcasm, irony, and other subjective phenomena.

5.4 Limitations

We divide the present study's limitations into two categories; conceptual and methodological.

First, we address the conceptual issue of operationalising Russian disinformation. As

previously mentioned, Twitter's dataset lacks transparency in how the IRA tweets were

distinguished from the tweets of genuine users (Twitter Elections Integrity, 2019). Due to the

clandestine nature of Russian disinformation, this dataset is the most accessible way to

approach our research aims. Previous academic scholars have also had to accept this

limitation (Zannettou et al., 2018; Kim et al., 2019; Miller, 2019), and given Twitter's

expertise and reputation we deemed the dataset reliable for research.

Second, we address another operationalisation issue: the presence of subjective emotion in

text. Emotion is inherently subjective and as we approach this study from a constructivist

epistemological position, emotion varies not only among individuals, but also across social

contexts. This is an ongoing issue within computational linguistics and beyond the limits of

this thesis to provide a solution. We accept this as a limitation, but argue that the presence of

an extremely large dataset increases the validity of emotion analysis. With a tweet corpus of

1.5 million tweets, tweet outliers were diluted in the aggregate, and the emotion analysis

methodology is sufficiently moderate.

Moreover, emotion analysis of tweets brings new limitations and challenges. Tweets are limited in length and tend to have many misspellings, slang terms, and shortened forms of words (Kiritchenko et al., 2014). This is why preprocessing is so important as the conversion of raw, messy data to a structured format is the most time-consuming step. Text within a specific domain may change due to the surrounding contextual factors. Additionally, online discussions often contain irony and sarcastic sentences which, when automatically coded, lose their subjectivity. That being said, the present study incorporated a mixed methods design to reduce these limitations by including qualitative methodology; the summative content analysis (Hsieh & Shannon, 2005). By including a researcher-coded approach to the tweets, instances of irony and sarcasm were managed appropriately.

A final limitation was the methodology's lack of negations handling due to limits in researcher coding abilities. Negations handling is important as it converts the sentiment of text from positive to negative using special words (e.g. not, don't). The whole sentiment of text may be changed with the addition of a negation. However, given the 3 word-emotion association minimum for tweet inclusion in an emotion category, this should have offset the lack of negations handling.

Chapter 6 Results

This chapter presents the results from both the quantitative and qualitative methodologies.

Results demonstrate the presence of all 8 of Plutchik's emotions in the IRA tweets dataset,

with fear and anger categories topping the list by a large margin. Summative content analysis

showcases the different themes among the 8 emotion categories, with several themes

overlapping across emotions. This chapter is organised first by the type of methodology

(quantitative and qualitative) and within the qualitative results section, by emotion category.

6.1 Emotion Analysis Results

The present study found evidence of all of Plutchik's 8 emotions in the IRA tweet dataset.

According to our operationalisation of an emotionally framed tweet (at least 3 word-emotion

associations), approximately 5% of the total IRA tweet corpus was emotionally framed

(Table 4).

Table 4. Total number of tweets in each emotion category

	Total number of tweets	Percent of total tweet corpus
Fear	31,318	2%
Anger	14,548	1%
Trust	9,961	1%
Sadness	8,070	1%
Anticipation	7,700	<1%
Joy	5,352	<1%
Disgust	2,921	<1%
Surprise	1,991	<1%

Fear outranked all other emotions in our analysis, even when emotion categories were standardised. Second in place was anger, yet with approximately half the amount of tweets as the fear category. The subsequent categories (in order) were anticipation, trust, joy, sadness, disgust, and surprise (Table 5).

Table 5. Emotion tweets ranked

	Total tweets/number of word-emotion associations in NRC lexicon
Fear	21.1
Anger	11.6

Anticipation	9.1
Trust	8.1
Joy	7.8
Sadness	6.8
Disgust	2.8
Surprise	2.2

6.2 Summative content analysis results (by emotion category)

6.2.1 Fear Category Analysis

Figure 5. Fear category word cloud

Crime news. Tweets within this theme consisted of local and breaking news overly focused on shootings involved with the police. Due to the amount of 'police' word frequencies, we also looked at the likely word associations with the word police. Within the fear category, the word 'police' was mostly associated with words 'officer', 'say', 'suspect', 'shot', 'shooting', and 'arrest.' Tweets were worded as news bites written in the third person and most likely accompanied URLs linking to news stories. Other crimes associated in this category were armed robberies, car crashes, and terror attacks (bomb threats).

Violence towards police/ Police shooting. Under the bigram 'police officer', tweets with the most fear emotion word associations discussed news about the shooting or killing of police officers (i.e. harm inflicted onto police officers). Within the 'police say' bigram, tweets contained information mainly around the harm of civilians, as reported by police.

Weight loss. Tweets in this theme conveyed how users 'hate feeling like [they] need to lose weight' and desires for losing weight. Tweets mentioned tips on losing weight, problems with losing weight, and a general wish to lose weight. Within this theme, tweets were apolitical, very informal, and written in the first person.

Terror attack. Many tweets reported breaking news stories regarding terror attacks. Tweets

were again worded as news bites in the third person and consisted of police warnings about

terror threats, punishments for terrorism, and current developments during a terror attack. If

the type of terror attack was described it mainly included jihadi and white nationalist

terrorism. Tweets were not confined to the US and included news about terror attacks in

Europe (London, Barcelona, Berlin, Manchester, and Paris), Canada, Turkey, Syria, and

Pakistan. Interestingly, there were several tweets about Russia and their aim to 'fight against

terrorism' in Syria alongside the Syrian government. Tweets also mentioned the

#columbianchemical hashtag 128 times, which was a reported terror attack hoax in St. Mary

Parish, Louisiana. Users under this hashtag wrote how they 'fear for their life' and spread

news about the possible terror attack.

Table 6. Frequencies of words and bigrams in fear category

Word	Frequency	Bigram	Frequency
police	11218	fatal shooting	878
news	6906	police say	762
shooting	4677	world news	714
suspect	4432	man accused	653
man	4076	police officer	652

death	2595	lose weight	622
shot	2515	local news	570
accused	2110	armed robbery	521
fatal	2085	man shot	461
court	1739	police shooting	425

Table 7. Samples of tweets in fear category and their corresponding themes

Tweets	Coded Themes
'man accused of killing 2 palm springs police officers in ambush faces death penalty'	Violence towards police, Death penalty, Crime news
'police say fuck this guy and shoot homeless mentally ill man 14 times killing him'	Police shooting, Crime news
'this is the moment the parliament terror attack suspect is wheeled into an ambulance after being shot by police damn amish terrorist'	Terror attack, Police shooting
'church terror alert worshippers told to be vigilant as uk police warn of christmas attack'	Terror attack
'putin we will continue supporting the syrian government in the fight against terrorism'	Terror attack, Russia

'problem is i have a nice face im scared
if i lose weight ill look like a rat
buyiasotea'

Weight loss

6.2.2 Anger Category Analysis

Figure 6. Anger category frequent words cloud

Crime news. Similar to the fear category, reported news of crimes was a popular theme for anger-related tweets. As 'crime' was the most frequently mentioned word, we looked at words frequently co-occurring with 'crime.' These included (in order) 'man', 'cleveland', 'blotter', 'police', 'akron', and 'hate'. Both Cleveland and Akron are cities in Ohio. Noting that crime was usually placed at the end of a tweet, it was a very likely choice for a hashtag.

Politics. The word politics did not appear as a word used within tweets themselves, but was added on at the end, indicating that it was a very likely hashtag. The majority of tweets were news stories and likely contained URL links. Within this theme, tweets remarked on a variety of divisive social issues in the US. Gun violence was a major topic, and contained tweets which advocated for both the implementation of gun laws and the right to bear arms. For example, one tweet stated that the 'loss of gun rights [is] more devastating than bullet wounds.' Another common issue was abortion as tweets reported on a 'faithbased birth control mandate' and 'restictive texas abortion laws.' Further, immigration was common as tweets reported on the deportation of undocumented immigrants, and 'immigrants in america held hostage.' Other issues included LGBT proposals, terrorism, violent crime, domestic

abuse, and the Fukushima Daiichi Accident. Tweets represented both liberal and conservative views.

Death penalty. Under the 'supreme court' bigram, most of the tweets concerned upholding sex offender registration laws and news regarding the death penalty. News stories about the death penalty were overall neutral (not pro or anti-death penalty) and news were widespread across the US - states included Ohio, Texas, Florida, Kansas, Missouri, and Illinois. News regarding the death penalty was also very prevalent in the politics theme.

Terror attack. This theme was very similar to the terror attack theme in the fear category. Tweets within this theme reported breaking news about terror attacks in the US, Europe, and Canada and mainly included jihadi and white nationalist terror attacks. One common hashtag among these tweets was #islamkills.

Table 8. Frequencies of words and bigrams in anger category

Word	Frequency	Bigram	Frequency
crime	2649	man accused	523
man	2490	pleads guilty	427
news	2412	fatal shooting	422
politics	2127	supreme court	331
shooting	2073	death penalty	313

death	1789	hate crime	208
police	1578	world news	235
accused	1546	accused killing	235
court	1315	terrorist attack	220
gun	1061	armed robbery	219

Table 9. Samples of tweets in anger category and their corresponding themes

Tweets	Coded Themes
'supreme court rejects texas death row inmates appeal crime news'	Death penalty
'cleveland man accused in shooting death of 5monthold aavielle wakefield appears in court video holmes faces an aggravated crime'	Crime news
'carson doesnt support exceptions for abortions in rape incest politics'	Politics
'lisamurkowski lisa all the terrorist attacks shock me it was a terrorist attack wasnt it columbianchemicals neverforget'	Terror attack

6.2.3 Trust Category Analysis

Figure 7. Trust frequent word cloud

Police appreciation. There was a common theme under the 'police officer' bigram of news about heroic acts of police officers and praise for their service (Table XX). Within the police theme, the majority of tweets describe harm inflicted upon police officers rather than harm inflicted by police. Included in this theme were tweets about general respect for the law and law enforcement appreciation (including ICE). Illinois, Texas, and Ohio were the most mentioned states.

Patriotism. The 'god bless' bigram consisted of the most thematically congruent tweets.

Tweets praised the American flag, the military, and the police. In many tweets, users call

each other 'patriots', celebrate the 4th of July (US day of independence), and report proud

feelings towards the US. Although there were a few critical tweets about Trump, the majority

were pro-Trump and many thanked President Trump for his support of the military and

police.

Table 10. Frequencies of words and bigrams in trust category

Word	Frequency	Bigram	Frequency
police	1668	police officer	367
news	1476	white house	231
top	806	attorney general	213
president	781	god bless	226
trump	719	world news	152
white	619	president trump	146
school	616	local news	125
officer	607	high school	120
county	575	united states	89
good	546	police say	88

Table 11. Samples of tweets in trust category and their corresponding themes

Tweets	Coded Themes
'god bless our soldiers they are the elite of our nation and the guarantee of our freedom'	Patriotism, Military appreciation
'god bless the police lets pray well never hear about new assassinations on our law enforcement officers'	Police appreciation
'good morning patriots fly our wonderful flag high and proud god bless america'	Patriotism
'i am proud of police officers who do their job for such little pay and such little appreciation'	Police appreciation

6.2.4 Sadness Category Analysis

Figure 8. Sadness category frequent word cloud

Death penalty. Within the 'death penalty' bigram, the majority of tweets mention cases with

charges of the death penalty, and a couple of tweets mention the case of the Boston marathon

bomber and a possible death penalty. The majority of the tweets appear to be news items and

likely contained links to news sources.

Weight loss. Users in this theme stressed how much they aspire to lose weight, are hateful

towards their bodies and are communicating tips on losing weight. Within this theme, the

tweets were less formal and personal. Unlike the death penalty theme, the tweets were mostly

written in first-person.

Violence towards black men. Within this theme, there were lots of tweets about a police

killing of a mentally ill black man in Brooklyn. In addition, the majority of tweets

communicated news about police violence towards black men and unjust treatment in prison.

Also within this theme were news about instances of racial discrimination and possible hate

crime charges. The majority of tweets were news items and in the third person.

Table 12. Frequencies of words and bigrams/trigrams in sadness category

Word	Frequency	Bigram/Trigram	Frequency
news	1264	lose weight	792
death	1168	pleads guilty	258
weight	1093	weight loss	247
man	933	death penalty	217
lose	886	found guilty	129
black	807	black man	127
case	787	fatal crash	117
guilty	687	new orleans	93
crash	637	need lose weight	90
murder	617	local news	88

Table 13. Sample of tweets in sadness category and their corresponding themes

Tweets	Coded Themes
'on the 50th anniversary of king assassination nyc police shot and killed a mentally ill black man in crown heights brooklyn he had a shower head in his hand'	Violence towards black men
'all i want in life is to lose weight and gain money yet instead here i am gaining weight and losing money'	Weight loss
'accused shooter in death penalty trial was ready to kill news'	Death penalty

6.2.5 Anticipation Category Analysis

Figure 9. Anticipation frequent word cloud

Patriotism. Under the 'god bless' bigram, the majority of tweets espouse patriotic views;

suggesting to praise the American flag and appreciate the police and military and 'people

who protect this country'. Alongside this, the majority of tweets were pro-Trump, suggesting

that he 'owns every narrative' and praising his public appearances. Users call each other

'patriots' and 'fellow americans', wish each other 'happy sundays', and praise the American

flag. Within this theme, there were nods to Christianity, wishing users a 'blessed sunday' and

notions of praying for police and for the military.

White supremacy. Tweets within this category were mostly against white supremacy and

mostly anti-Trump. Several tweets stated that President Trump and the White House spread

and legitimated white supremacy and others argued that the Democrats invented white

supremacy. The majority of tweets were informal and opinionated and commonly responded

to current events (e.g. Charlottesville Unite the Right rally and reactions to Trump). De blasio

was frequent mentioned within this theme as tweets stated that he called '63 million trump

supporters a white supremacist movement.'

Table 14. Frequencies of words and bigrams in anticipation category

Word	Frequency	Bigram	Frequency
news	1060	god bless	283
time	911	white house	221
top	793	make money	160
good	781	supreme court	144
money	636	world news	126

white	605	happy birthday	105
god	568	good morning	90
court	533	white supremacist	83
watch	504	lose weight	82
death	481	de blasio	76

Table 15. Samples of tweets in anticipation category and their corresponding themes

Tweets	Coded Themes
'bannon is out but trump white house spreads white supremacy and hate time to purge racially divisive forces inside the wh rootoutracism'	White supremacy
'we wont let some hateful extremists blacken the name of our protectors we stand with our cops god bless america'	Pro-police
'it feels great to finally have a president that respects the military god bless president trump maga'	Pro-Trump, Patriotism
'i am ready to fight for america god bless america instotus patriots patriot conservative politics'	Patriotism

6.2.6 Joy Category Analysis

Figure 10. Joy category frequent word cloud

Inspirational messages. The word 'love' was related to messages about romance, parenting,

religious inspirational messages, and general inspirational messages. Under the 'god bless'

bigram there were messages about American as the best country in the world. 'Lose weight' renders positive and inspirational messages of weight loss and 'make money' about making money.

Crime news. Under 'found safe' there were mainly news stories of children and who were found safe after being missing. The 'oh my god' trigram rendered several tweets inconsistent with the joy category as many contained negations, sarcasm and irony. For example, there were several tweets which concerned the possible poisoning of store-bought turkeys during Thanksgiving and blaming Koch Farms.

Patriotism. Under the 'white house' bigram, tweets were mostly feel-good messages about Trump, Obama, and previous political leaders. The 'god bless america' trigram renders patriotic posts celebrating the 4th of July, praising the American flag and military, and calls other users 'patriots'.

Table 16. Frequencies of words and bigrams/trigrams in joy category

Word	Frequency	Bigram/Trigram	Frequency
love	806	god bless	343
god	697	make money	157
good	683	lose weight	152

happy	527	happy birthday	123
news	476	happy thanksgiving	95
money	459	good morning	87
bless	362	oh my god	81
entertainment	340	white house	76
laughing	335	found safe	71
found	314	god bless america	69

Table 17. Samples of tweets in joy category and their corresponding themes

Tweets	Coded Themes
'every time you smile at someone it is an action of love a gift to that person a beautiful thing mother teresa'	Inspirational message
'good morning patriots let the sun always shine upon our wonderful land god bless america'	Inspirational message, Patriotism
'happy thanksgiving thanks god ive got allergy to a turkey kochfarms cant poison me'	Crime news
'missing ohio toddler found safe with mother reports say'	Crime news

6.2.7 Disgust Category Analysis

Figure 11. Disgust frequent word cloud

Weight loss. Within the 'lose weight' bigram, the majority of tweets highlighted user insecurities about losing weight. Tweets mentioned tips on losing weight, problems with losing weight, and a general wish to lose weight. Within this theme, tweets were apolitical, very informal, and written in the first person.

Anti-McCain. All tweets surrounding John McCain were negative. Many tweets used foul language to insult him and called him a 'traitor', a 'rat', 'jealous', and a 'liar.' The majority of tweets did not read like news stories, but rather as opinions of John McCain. When a few tweets within this category mentioned Trump, they were overtly pro-Trump.

Murder news. Within the 'death' and 'murder' word frequencies, tweets commonly reported crimes involving rape, shootings, murder, and death. Several tweets reported violent crimes perpetrated by 'illegal aliens.' Tweets were written as news bites and likely contained URLs to the reported news stories.

Special counsel/Anti-Mueller. Tweets which mentioned the special counsel called it an 'abuse of power' or 'gross abuse.' Robert Mueller was frequently insulted as a 'witchhunter', 'pathetic', a 'dirty cop', and a 'swamp creature.' Within this theme, tweets were all negative and opposed to the special counsel and Robert Mueller. When Russia was mentioned, users called the investigation a 'fake russia lie' concocted by the US Democrats.

Table 18. Frequencies of words and bigrams in disgust category

Word	Frequency	Bigram	Frequency
weight	785	lose weight	737
lose	777	john mccain	104
death	367	need lose	103
trump	309	gross abuse	82
fat	261	special counsel	82
hate	220	traitor john mccain	82
news	208	counsel gross abuse	80
john	184	lawyer rips special	77
just	176	swamp creature	75
murder	170	deep state	72

Table 19. Samples of tweets in disgust category and their corresponding themes

Tweets	Coded Themes
'i have a great idea you wanna lose weight get your fat ass up and go to the damn gym deanmachine'	Weight loss
'illegal alien beats 4 year old child to death in denver judge calls it the most violent child murder she has seen'	Murder news, Immigration
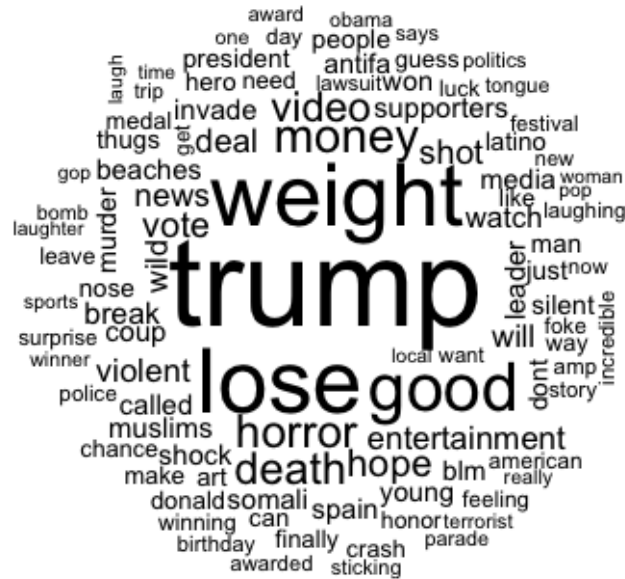
| 'traitor mccain goes on face the nation 2 trash trump over bogus russian hackers john mccain still believes that his opinion really matters' | Anti-Mccain |
| 'dobbs nasty swamp creature mueller assembling clinton team in trump coup' | Special counsel, Anti-Mueller |

6.2.8 Surprise Category Analysis

Weight loss. Users tweeted about wanting to lose weight and tips on losing weight. In contrast with tweets in other themes the content was very informal and written in the first person.

Pro-/Anti- Trump. Under the 'trump supporters' bigram, tweets communicated hate crime hoaxes and violence towards Trump supporters. Within this category, there was a mix of both pro-Trump and anti-Trump tweets. Pro-trump tweets expressed how Trump is a businessman able to earn money while anti-Trump tweets communicated news stories about Trump's inability to serve the US as President. There were also a lot of tweets about Missouri senator Maria Chappelle-Nadal saying that she hopes Trump will be assassinated.

Antifa. Many tweets communicated news about Antifa (anti-fascism group) and violence perpetrated by that group, especially towards Trump supporters. Several tweets conveyed sarcasm and irony about Antifa as a 'peaceful' and non-violent' group, yet tweets stated that Antifa's website calls upon protestors to 'stab trump supporters.'

Table 20. Frequencies of words and bigrams/trigrams in surprise category

Word	Frequency	Bigram/Trigram	Frequency
trump	655	lose weight	410
lose	454	trump supporters	79

weight	438	trump won	73
good	318	coup trump	71
money	201	blm leader called	69
death	178	violent coup	69
horror	151	shock video	65
hope	139	media silent	64
video	128	antifa thugs	63
vote	125	donald trump	54

Table 21. Sample of tweets in surprise category and their corresponding themes

Tweets	Coded Themes
'another hate hoax muslim teen who claimed 3 drunk trump supporters called her a terrorist amp attacked her on the subway lied about it all'	Trump, Hate hoax
'today is sun 25 feb 2018 400 days into the worst presidency in the history of the just another citizen of the world i too call for the resignation of donald trump a corrupt erratic dishonest vindictive unstable man unfit to serve'	Anti-Trump

'rt joerodr36477397 will the ag jeffsessions finally go after the antifa terrorist groups sponsored by george soros or after trump support'	Antifa
'rt 614swat thanks donald trump business is the best its been in years im not only sick of winning im sick of all this money trum'	Pro-Trump

6.3 Chapter Conclusion

The results of the present study demonstrate the presence of all 8 of Plutchik's emotions in the IRA tweets dataset; fear, anger, trust, sadness, anticipation, joy, disgust, and surprise.

Fear was the most present emotion by a large margin; the next most frequent emotion, anger, was present in half as many tweets as fear. Through summative content analysis we showcased themes within the 8 different emotion categories. Fear, the most present category, contained themes of violence towards police, terror attacks and general reports about crime. Anger was most similar to fear, but produced more content on the state of the death penalty and didn't include the popular theme 'Weight loss.' Trust, most similar to anticipation contained tweets which conveyed American patriotism and appreciation for the police and

the military. While anticipation also contained the 'Patriotism' theme, this emotion category

was more political as it also contained pro-Trump tweets. Although the joy category included

patriotic tweets, it was one of the most apolitical emotion categories. The sadness category

included themes of violence towards black men (only present in the sadness category), death

penalty news, and 'Weight loss'. Disgust, arguably the most political emotion category was

fixed on anti-McCain sentiments and opposing the White House Special Counsel. Lastly, the

surprise category focused on pro and anti-Trump sentiments, hate hoaxes, weight loss, and

Antifa.

In the next chapter we will discuss the results of both the quantitative and qualitative analyses

and connect them with the larger disinformation literature.

Table 22. Summary of themes in each emotion category

	Themes
Fear	Crime news, Violence towards police, Police shooting, Terror attack, Weight loss
Anger	Crime news, Death penalty, Terror attack, Politics
Trust	Police appreciation, Patriotism, Military appreciation
Sadness	Death penalty, Weight loss, Violence towards black men
Anticipation	Patriotism, White supremacy, Pro-police, Pro-Trump
Joy	Inspirational messages, Crime news, Patriotism

| Disgust | Weight loss, Anti-McCain, Murder news, Special counsel, Anti-Mueller, Immigration |
| Surprise | Weight loss, Pro-Trump, Anti-Trump, Antifa, Hate hoax |

Chapter 7 Discussion of Results

The present study aimed to answer three interrelated research questions:

RQ1: What emotions exist in Russian disinformation (IRA) on Twitter?

We found presence of all 8 of Plutchik's emotions in the IRA tweet corpus; fear, anger, trust, anticipation, sadness, joy, disgust, and surprise. Within the tweet corpus, approximately 5% of tweets belonged to a discrete emotional frame. This finding runs contrary to previous studies which found significantly higher degrees of emotion. Farkos and Bastos (2018) demonstrated that 10% of IRA tweets comprised of highly emotional statements. Spangher et al (2018) noted an even higher number of emotionality. After coding IRA tweets researchers found 42% of them 'neutral' and approximately 20% as 'high' or 'very high' in emotionality. This finding suggests that Russian disinformation may not be as highly emotionalised as previously thought, at least within IRA tweets.

Another interpretation of this finding may be explained by differences in methodology. Farkos and Bastos (2018) and Spangher et al (2018) manually coded tweets for emotionality whereas the present study adopted quantitative emotion analysis. This lends us to ask the question of whether qualitative analyses overstate the levels of emotion or whether the present methodology was too stringent in assigning tweets to emotion categories.

RQ2: Are certain discrete emotions more prevalent than others in Russian disinformation (IRA) on Twitter?

Within the IRA tweet dataset, fear was by far the most prevalent emotion. The fear emotion category contained an abundance of crime-related news, and tweets related to developing or past terror attacks. This corroborates the findings of Farkas and Bastos (2018) who discovered IRA tweet bias towards issues around public security, (particularly crime), fatal accidents and natural disasters. Furthermore, the fear category possessed similar themes to the second most present emotion in the IRA dataset, anger. Both fear and anger contained

themes of terror attacks and crime-related news. We noted that anger is a key emotion linked to political action, and unlike fear, the anger category possessed news related to politically divisive social issues in the US (Roseman, 2018).

RQ3: Do specific themes or topics emerge from discrete emotion categories in Russian disinformation (IRA) on Twitter?

Yes, several distinct themes emerged in each emotion category. Within the fear category we witnessed themes of violence towards police, terror attacks and general reports about crime. Anger, while most similar to fear, additionally contained political content regarding immigration, abortion, and gun violence. Trust, most similar to anticipation contained tweets which conveyed American patriotism and appreciation for the police and the military. While anticipation also contained the 'Patriotism' theme, this emotion category was more political as it also contained pro-Trump tweets. Although the joy category included patriotic tweets, it was one of the most apolitical emotion categories. The sadness category included themes of violence towards black men (only present in the sadness category), death penalty news, and 'Weight loss'. Disgust, arguably the most political emotion category was fixed on anti-McCain sentiments and opposing the White House Special Counsel. Lastly, the surprise category focused on pro and anti-Trump sentiments, hate hoaxes, weight loss, and Antifa. In addition to the themes

Polarising news stories. Within many of the emotion categories, there were several political and highly disseminated news stories. Within the surprise category, IRA tweets focused on Missouri senator Chappelle-Nadal, who appeared in a genuine news story in August of 2017 (Wong & Bever, 2017; Oppenheim, 2017). Chappelle-Nadal, an outspoken critic of racial

discrimination in the US, criticised the shooting of Michael Brown in Ferguson in 2014 (Oppenheim, 2017). In response to a comment on her Facebook page, she wrote '*I hope Trump is assassinated*!' (Wong & Bever, 2017). This sparked outrage among the conservative voices in the US and IRA accounts took advantage and propagated this story. Further in the surprise category, IRA accounts frequently disseminated news story about Antifa (short for anti-fascism) violence, especially towards Trump supporters. Many of the tweets were posted around the time of the Unite the Right rally in Charlottesville in August 2017, a white nationalist demonstration which included both far-right groups and counter-protesters including Antifa (Stockman, 2017; Cammeron, 2017). The dissemination of these stories around the same time as the 'Unite the Right' rally provides a microcosm of the way IRA accounts attach themselves to genuine news and aggravate existing racial and political tensions.

Weight loss category. The weight loss theme was one of the most surprising themes given it is apolitical and strongly spread in several emotion categories. The weight loss theme occurred in the fear, sadness, disgust, and surprise emotion categories and tweets were mostly negative in sentiment. One explanation for the substantial presence of this theme is a strategy within the IRA to appear more human, relatable, and authentic as a proxy for credibility (Linvill et al., 2019).

Political emotion categories. Miller's (2019) analysis of IRA tweets contained a significant amount of discussion around political issues. We noted earlier that this did not occur in the fear category, however anger, anticipation, and disgust contained the most politicised tweets. Anger, containing very similar themes to the fear category distinctively contained political

news and opinions. The most frequent topics included gun violence, gun rights, abortion, immigration, LGBT rights, terrorism, and domestic abuse. Miller (2019) also noted refugee and border issues, former President Obama, Russia/Putin, and black lives matter movement/black power movement. Both the present study, Miller (2019), and Spangher et al (2018) found evidence of the IRA exacerbating both liberal and conservative issues.

Disgust, while also very political, was one of the most singularly coherent categories and contained anti-Mueller and anti-McCain sentiment. Arizona senator John McCain has been a target of Russian disinformation for years and Robert Mueller is an understandable target as he is leading the US Special Counsel Russia investigation. What struck out most however was the tone of the tweets describing the two political leaders as 'dirty', 'rats', and 'traitors.' The disgust category was by a large margin the most polarising and contained the most extreme sentiments. Under the anticipation category tweets spread informal and opinionated content around current events (e.g. Charlottesville Unite the Right rally).

Patriotism. One theme that did not occur in the disinformation literature is patriotism. Within the trust and anticipation emotion categories, IRA tweets spread informal messages to praise the American flag, and appreciate the police and military, and 'people who protect this country.' One very related theme to patriotism that we also observed was Christianity which corroborates what Miller (2019) found in his topic model of IRA tweets.

7.1 Criticism

A valid criticism of the current study may be that over the entire tweet corpus, the majority of IRA tweets are banal content which is not reflective of an overall disinformation strategy. IRA accounts may tweet not to influence, but rather for other seemingly innocuous strategies, like maintaining credibility or trying to increase their follower count. Furthermore, previous studies of IRA content have found that news stories, oftentimes recycled for tweet content are irrelevant content for a massive disinformation manipulation campaign (Linvill et al., 2019). To this criticism, we argue that even if news stories are being repurposed, previous studies have shown that they are statistically different in their language use (Boyd et al., 2018). Boyd et al (2018: 7) found 'evidence that the IRA actors were not only composing their own tweets, but were doing so in a carefully constructed, intentionally deceptive manner.' Therefore, tweets were directly composed by IRA accounts as their style and composition was unique and consistently differentiable from the general Twitter population (Boyd et al., 2018). This provides justification that even if the tweets are banal news stories, IRA actors still directly choose the stories to cover and disseminate into the Twitter environment.

While this study explored the presence of emotion in Russian disinformation, it did not compare the disinformation sample to a sample of genuine Twitter users. As such, this renders it difficult to make conclusions as we don't have a standard sample to compare to. We do not know what the average levels of each emotion are on Twitter which makes it difficult to draw final conclusions.

7.2 Conclusion and Implications

This study provides a first look at the emotions present in Russian disinformation on Twitter. Within a Twitter dataset from IRA-linked accounts, we observed that the most present emotions were fear and anger, followed by anticipation, trust, joy, sadness, disgust, and surprise. While fear and anger have been widely cited as effective emotional appeals within public health communication contexts (Tannenbaum et al., 2018), to our knowledge these appeals have not been studied within the context of disinformation. This research provides researchers a starting point for analysis - especially qualitative researchers - to look into fear and anger in Russian disinformation. As we also completed a summative content analysis of the tweets in each emotion category, the study brings new themes to light into the discussion of Russian disinformation (e.g. violence against black men, patriotism, and weight loss).

Bibliography

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

Adriaans, P., 2010. A critical analysis of Floridi's theory of semantic information. Knowledge, Technology & Policy, 23(1-2), pp.41-56.

Alaphilippe, A., Ceccarelli, C., Charlet, L., Mycielski, M. (2018). Disinformation detection system: 2018 Italian elections Case report. EU Disinfo Lab. Retrieved from: https://disinfoportal.org/wp-content/uploads/ReportPDF/Disinformation-Detection-System-2018-Italian-Elections.pdf.

Allen, T.S. and Moore, A.J., 2018. Victory without Casualties: Russia's Information Operations. Parameters, 48(1), pp.59-71.

American Psychological Association. (2015, October 22). Fear-based appeals effective at changing attitudes, behaviors after all: Effects are most apparent in women and for one-time behaviors, says study. ScienceDaily. Retrieved July 24, 2019 from www.sciencedaily.com/releases/2015/10/151022094459.htm

Andrew, C. 2000. The Mitrokhin Archive. The RUSI Journal, 145(1), 52-56.

Arif, A., Stewart, L.G. and Starbird, K., 2018. Acting the part: Examining information operations within# blacklivesmatter discourse. Proceedings of the ACM on Human-Computer Interaction, 2(CSCW), p.20.

Armstrong, D. B., Fogarty, G. J., & Dingsdag, D. (2007). Scales measuring characteristics of small business information systems. Proceedings of the 2011 Conference on Empirical Methods in Natu-ral Language Processing (pp. 562-570). Edinburgh, Scotland. Retrieved from http://www.aclweb.org/anthology/D11-1052

Aro, J., 2016. The cyberspace war: propaganda and trolling as warfare tools. European View, 15(1), pp.121-132.

Aro, J., 2016. The cyberspace war: propaganda and trolling as warfare tools. European View, 15(1), pp.121-132.

Badawy, A., Addawood, A., Lerman, K. and Ferrara, E., 2018. Characterizing the 2016 Russian IRA Influence Campaign. arXiv preprint arXiv:1812.01997.

Badawy, A., Lerman, K. and Ferrara, E., 2018. Who Falls for Online Political Manipulation?. arXiv preprint arXiv:1808.03281.

Bakamo. (2017). French Election 2017: Social Media Landscape. Retrieved from: https://static1.squarespace.com/static/58495e3329687f8bfbb3f25c/t/590904cb15cf7da7dc3b15d8/1493763322834/Patterns+of+Disinformation+in+the+2017+French+Presidential+Election+-+Report+2+-+Bakamo.pdf.

Bakir, V. and McStay, A., 2018. Fake news and the economy of emotions: Problems, causes, solutions. Digital Journalism, 6(2), pp.154-175.

Bellamy, C. (2001). What is Information Warfare?. In: R. Mathews and J. Treddenick, ed., Managing the Revolution in Military Affairs, 1st ed. New York: Palgrave.

Bennett, W.L. and Livingston, S., 2018. The disinformation order: Disruptive communication and the decline of democratic institutions. European journal of communication, 33(2), pp.122-139.

Bentzen, 2018. Foreign influence operations in the EU. Accesed 03/03/19 at www.europarl.europa.eu/RegData/etudes/BRIE/.../EPRS_BRI(2018)625123_EN.pdf

Berger, J. and Milkman, K.L., 2012. What makes online content viral?. Journal of marketing research, 49(2), pp.192-205.

Berger, J. and Milkman, K.L., 2012. What makes online content viral?. Journal of marketing research, 49(2), pp.192-205.

Besemeres, J., 2016. A Difficult Neighbourhood: Essays on Russia and East-Central Europe since World War II. ANU Press. (CHAPTER TITLE: Russian disinformation and Western misconceptions)

Bessi, A. and Ferrara, E., 2016. Social bots distort the 2016 US Presidential election online discussion. First Monday, 21(11).

Blaikie, N., 2007. Approaches to social enquiry: Advancing knowledge. Polity.

Bleiker, R. and Hutchison, E., 2008. Fear no more: emotions and world politics. Review of international studies, 34(S1), pp.115-135.

Bogdanov and Gorbunov. 2009. 'On the Character of Armed Confrontation in the Twenty-First Century," Voennaya Mysl' (Military Thought), No. 3 2009, pp. 2, 11.

Boghardt, T. (2009). Soviet Bloc Intelligence and Its AIDS Disinformation Campaign. Studies in intelligence, 53(4), 1-24.

Boyd, R.L., Spangher, A., Fourney, A., Nushi, B., Ranade, G., Pennebaker, J. and Horvitz, E., 2018. Characterizing the Internet Research Agency's Social Media Operations During the 2016 US Presidential Election using Linguistic Analyses.

Brader, T., 2005. Striking a responsive chord: How political ads motivate and persuade voters by appealing to emotions. American Journal of Political Science, 49(2), pp.388-405.

Brader, T., 2006. Campaigning for hearts and minds: How emotional appeals in political ads work. University of Chicago Press.

Bradley, M.M. and Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry, 25(1), pp.49-59.

Brady, W.J., Wills, J.A., Jost, J.T., Tucker, J.A. and Van Bavel, J.J., 2017. Emotion shapes the diffusion of moralized content in social networks. Proceedings of the National Academy of Sciences, 114(28), pp.7313-7318.

Briscoe, E.J., Appling, D.S. and Hayes, H., 2014, January. Cues to deception in social media communications. In 2014 47th Hawaii international conference on system sciences (pp. 1435-1443). IEEE.

Broniatowski, D.A., Jamison, A.M., Qi, S., AlKulaib, L., Chen, T., Benton, A., Quinn, S.C. and Dredze, M., 2018. Weaponized health communication: Twitter bots and Russian trolls amplify the vaccine debate. American journal of public health, 108(10), pp.1378-1384.

Burch, L.M., Frederick, E.L. and Pegoraro, A., 2015. Kissing in the carnage: An examination of framing on Twitter during the Vancouver riots. Journal of Broadcasting & Electronic Media, 59(3), pp.399-415.

Cacciatore, M.A., Scheufele, D.A. and Iyengar, S., 2016. The end of framing as we know it... and the future of media effects. Mass Communication and Society, 19(1), pp.7-23.

Calamus, 2017. Putin Says 'Patriotic Hackers' May Have Targeted U.S. Election. Accessed 07/02/2019 https://www.theatlantic.com/news/archive/2017/06/putin-russia-us-election/528825/

Calvo, R.A. and Mac Kim, S., 2013. Emotions in text: dimensional and categorical models. Computational Intelligence, 29(3), pp.527-543.

Cammeron, 2017https://www.bbc.co.uk/news/world-us-canada-40930831

Casero-Ripollés, A., 2012. Beyond newspapers: News consumption among young people in the digital era. Comunicar: Revista Científica de Comunicación y Educación, 20(39), pp.151-158.

Chadwick, A.E., 2015. Toward a theory of persuasive hope: Effects of cognitive appraisals, hope appeals, and hope in the context of climate change. Health communication, 30(6), pp.598-611.

Chekinov and Bogdanov. 2010 "Asymmetric Actions in Support of the Military Security of Russia," Voennaya Mysl' (Military Thought), No. 3 , p. 16.

Chekinov and Bogdanov. 2011. 'The Strategy of the Indirect Approach: Its Impact on Modern Warfare,' Voennaya Mysl' (Military Thought), No. 6 2011, p. 4.

Chekinov and Bogdanov. 2012. "Initial Periods of War and Their Impact on a Country's Preparations for a Future War," Voennaya Mysl' (Military Thought), No. 11 2012, pp. 14-27.

Chekinov and Bogdanov. 2013. 'The Nature and Content of a New–Generation War," Voennaya Mysl' (Military Thought), No. 10 2013, pp. 13-25.

Chekinov and Bogdanov. 2015 'The Art of War at the Beginning of the 21 stCentury: Problems and Opinions," Voennaya Mysl' (Military Thought), Vol. 1 2015, p. 34.

Chekinov and Bogdanov. 2015. 'A Forecast of Future Wars: Meditation on What They Will Look Like," Voennaya Mysl' (Military Thought), No. 10, 2015, pp. 41-49.

Chilek, L.A., 2017. Tugging at the heartstrings?: examining discrete emotion in nonprofit Twitter messages and its effect on pass along behavior (Doctoral dissertation).

Christensen, M.B., 2018. The Ethics of Social Media Policy: National Principles of Justice, Security, Privacy and Freedom Governing Online Social Platforms in Russia, China and The United States (Doctoral dissertation, UCLA).

Čižik, T., 2018. Cultural Identity as Tool of Russian Information Warfare. in Science & Military 1/2018.

Colombetti, G., 2009. From affect programs to dynamical discrete emotions. Philosophical Psychology, 22(4), pp.407-425.

Colonel S. G. Chekinov (Res.) and Lieutenant General S. A. Bogdanov (Ret.), "The Nature and Content of a New-Generation War," Military Thought 4 (2013).

Crawford, N.C., 2000. The passion of world politics: Propositions on emotion and emotional relationships. International Security, 24(4), pp.116-156.

Creswell, J.W. and Clark, V.L.P., 2018. Designing and conducting mixed methods research. Sage publications.

Cunningham, S.B., 2001. Philosophizing Propaganda.

D'Andrea, A., Ferri, F., Grifoni, P. and Guzzo, T., 2015. Approaches, tools and applications for sentiment analysis implementation. International Journal of Computer Applications, 125(3).

Damasio, A.R., 2002. Descartes' Error: Emotion, Reason and the Human Brain. Bulletin of the American Meteorological Society, 83(5), p.742.

Darczewska, J., 2014. The anatomy of Russian information warfare. The Crimean operation, a case study. Ośrodek Studiów Wschodnich im. Marka Karpia.

Darczewska, J., 2015. The devil is in the details. Information warfare in the light of Russia's military doctrine. Ośrodek Studiów Wschodnich im. Marka Karpia.

de los Santos, T.M. and Nabi, R.L., 2019. Emotionally Charged: Exploring the Role of Emotion in Online News Information Seeking and Processing. Journal of Broadcasting & Electronic Media, 63(1), pp.39-58.

Dehaene, S., 2014. Consciousness and the brain: Deciphering how the brain codes our thoughts. Penguin.

Della Sala, V., 2010. Political myth, mythology and the European Union. JCMS: Journal of Common Market Studies, 48(1), pp.1-19.

Demos, 2018 https://demos.co.uk/press-release/new-demos-analysis-finds-russian-influence-operations-on-twitter-targeted-at-uk-were-most-visible-when-discussing-islam/

Demos, 2018. New Demos analysis finds Russian influence operations on Twitter targeted at UK were most visible when discussing Islam. Accessed 04/02/2019 at https://demos.co.uk/press-release/new-demos-analysis-finds-russian-influence-operations-on-twitter-targeted-at-uk-were-most-visible-when-discussing-islam/

Dickerson, J.P., Kagan, V. and Subrahmanian, V.S., 2014, August. Using sentiment to detect bots on twitter: Are humans more opinionated than bots?. In Proceedings of the 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (pp. 620-627). IEEE Press.

Digital Forensic Lab, (atlantic council), 2018 https://medium.com/dfrlab/trolltracker-twitter-troll-farm-archives-8d5dd61c486b

DrConservaMom, 2018. If Liberal elites imagined 10,000 low-income housing "refugees" marching up THEIR streets 2 invade THEIR mansions in hopeful search of better homes with cooler stuff, they might understand the threat a migrant #Mob poses 2 America's borders, resources& sovereignty #StopTheCaravan. @@ConservaMomUSA. 23 October 2018. Available from: https://twitter.com/conservamomusa/status/1054866991146774528 [Accessed 30 April 2019].

Dretske, F.I., 1983. Précis of Knowledge and the Flow of Information. Behavioral and Brain Sciences, 6(1), pp.55-63.

Dugin's 'Post modern geopolitics' =

Dutch intelligence Service (2017). Annual report AIVD 2017. Retrieved from: https://www.aivd.nl/documenten/jaarverslagen/2018/03/06/jaarverslag-aivd-2017

Dutt, R., Deb, A. and Ferrara, E., 2018, December. "Senator, We Sell Ads": Analysis of the 2016 Russian Facebook Ads Campaign. In International Conference on Intelligent Information Technologies (pp. 151-168). Springer, Singapore.

Edson Tandoc Jr,.C., Lim, Z.W. and Ling, R., 2018. Defining "fake news" A typology of scholarly definitions. Digital Journalism, 6(2), pp.137-153.

Ekman, P. and Friesen, W.V., 1971. Constants across cultures in the face and emotion. Journal of personality and social psychology, 17(2), p.124.

Ekman, P., 1992. An argument for basic emotions. Cognition & emotion, 6(3-4), pp.169-200.

Ekman, P., 1993. Facial expression and emotion. American psychologist, 48(4), p.384.

EU vs Disinfo, 2018. Russia's Double Media Standards. Accessed 04/05/19 at:

https://euvsdisinfo.eu/russias-double-media-standards/

European Commision, 2018. Questions and Answers – The EU steps up action against disinformation. Accessed 07/20/19 at http://europa.eu/rapid/press-release_MEMO-18-6648_en.htm

European Commission, 2019 - action plan against disinformation

European Parliamentary Research Service, 2019. Automated tackling of disinformation. Accessed 06/04/2019 at www.europarl.europa.eu/RegData/etudes/STUD/.../EPRS_STU(2019)624278_EN.pdf

Everett, R.M., Nurse, J.R. and Erola, A., 2016, April. The anatomy of online deception: What makes automated text convincing?. In Proceedings of the 31st Annual ACM symposium on applied computing (pp. 1115-1120). ACM.

Fallis, D., 2009. A conceptual analysis of disinformation.

Fallis, D., 2011. Floridi on Disinformation. Etica & Politica, 13(2), pp.201-214.

Fallis, D., 2015. What is disinformation?. library trends, 63(3), pp.401-426.

Fan, R., Zhao, J., Chen, Y. and Xu, K., 2014. Anger is more influential than joy: Sentiment correlation in Weibo. PloS one, 9(10), p.e110184.

Farkas, J. and Bastos, M., 2018, July. IRA Propaganda on Twitter: Stoking Antagonism and Tweeting Local News. In Proceedings of the 9th International Conference on Social Media and Society (pp. 281-285). ACM.

Ferrara, E., 2017. Disinformation and social bot operations in the run up to the 2017 French presidential election.

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

Fessler, D. M. (2010). Madmen: An evolutionary perspective on anger and men's violent responses to transgression. In M. Potegal, G. Stemmler, & C. Spielberger (Eds.), International handbook of anger (pp. 361–381). New York, NY: Springer.

Fetzer, J.H., 2004. Disinformation: The use of false information. Minds and Machines, 14(2), pp.231-240.

Floridi, L. (2005). Semantic conceptions of information. In Stanford Encyclopedia of Philoso-phy. Retrieved May 1, 2019, from http://plato.stanford.edu/entries/information -semantic

Floridi, L. (2011). The philosophy of information. New York: Oxford University Press.

Floridi, L., 2003, July. Informational realism. In Selected papers from conference on Computers and philosophy-Volume 37 (pp. 7-12). Australian Computer Society, Inc..

Floridi, L., 2012. Steps Forward in the Philosophy of Information. Etica & Politica, 14(1), pp.304-310.

Fontaine, J.R., Scherer, K.R. and Soriano, C. eds., 2013. Components of emotional meaning: A sourcebook. Oxford University Press.

Freedman, L. (2006). The transformation of strategic Affairs. London: International Institute for Strategic Studies.

Frijda, N.H., 1986. The emotions. Cambridge University Press.

Galeotti, 2018. https://foreignpolicy.com/2018/03/05/im-sorry-for-creating-the-gerasimov-doctrine/

Galeotti, M., 2016. Hybrid, ambiguous, and non-linear? How new is Russia's 'new way of war'?. Small Wars & Insurgencies, 27(2), pp.282-301.

Gastil, J., Braman, D., Kahan, D.M. and Slovic, P., 2005. The'Wildavsky Heuristic': The cultural orientation of mass political opinion. Yale Law School, Public Law Working Paper, (107).

Gerasimov, "New Forms of Confrontation Employed by Western Countries will be Considered when Developing Russia's Defense Plan," Armeiskiy Sbornik (Army Journal), No. 3 2015.

Gerasimov, "The Role of the General Staff in the Organization of the Country's Defense in Accordance with the New Statue on the General Staff, Approved by the President of the Russian Federation," Vestnik Akademii Voennykh Nauk (Bulletin of the Academy of Military Science), No. 1 2014, pp. 14-22.

Gerasimov, "The Syrian Experience. Hybrid Warfare Requires High-Tech Weapons and Scientific Substantiation," Military Industrial Courier Online, 9-15 March 2016.

Gerasimov, "The Value of Science is in Foresight: New Challenges Demand Rethinking the Forms and Methods of Carrying out Combat Operations," Voyenno-Promyshlennyy Kuryer Online (Military-Industrial Courier Online), 26 February 2013.

Giles, K., 2011, June. "Information Troops"-A Russian Cyber Command?. In 2011 3rd International Conference on Cyber Conflict (pp. 1-16). IEEE.

Giles, K., 2016. Handbook of Russian information warfare. NATO Defence College Research Division.

Giles, Sherr, Seaboyer. 2018. Russian Reflexive Control. Royal Military College of Canada. Accessed 20/07/2019. Available at:

https://www.researchgate.net/publication/328562833_Russian_Reflexive_Control

Greenzeiger, M. F., Phulari, R., & Sanghavi, M. K. (2015). U.S. Patent No. 8,965,828. Washington, D.C.: U.S. Patent and Trademark Office.

Grigas, A., 2016. Beyond Crimea: the new Russian empire. Yale University Press.

Groenendyk, E. W., & Banks, A. J. (2014). Emotional rescue: How affect helps partisans overcome collective action problems. Political Psychology, 35, 359–378.

Guardian, 2019. 'Mueller's testimony on Trump and Russia: the biggest takeaways.' Accessed 26/07/2019 at https://www.theguardian.com/us-news/2019/jul/24/robert-mueller-testimony-key-takeaways-exoneration-indictment

Hager, A., 2019. Do Online Ads Influence Vote Choice?. Political Communication, pp.1-18.

Hansen, L.K., Arvidsson, A., Nielsen, F.Å., Colleoni, E. and Etter, M., 2011. Good friends, bad news-affect and virality in twitter. In Future information technology (pp. 34-43). Springer, Berlin, Heidelberg.

Harlow, S. and Johnson, T.J., 2011. Overthrowing the protest paradigm? How The New York Times, Global Voices and Twitter covered the Egyptian Revolution. International Journal of Communication.

Hellman, M. and Wagnsson, C., 2017. How can European states respond to Russian information warfare? An analytical framework. European Security, 26(2), pp.153-170.

Helmond, A., 2015. The platformization of the web: Making web data platform ready. Social Media+ Society, 1(2), p.2056305115603080.

Hjørland, B., 2007. Information: Objective or subjective/situational?. Journal of the American society for information science and technology, 58(10), pp.1448-1456.

Hjorth, F. and Adler-Nissen, R., 2019. Ideological Asymmetry in the Reach of Pro-Russian Digital Disinformation to United States Audiences. Journal of Communication, 69(2), pp.168-192.

Hoskins, A. and Shchelin, P., 2018. Information war in the Russian media ecology: the case of the Panama Papers. Continuum, 32(2), pp.250-266.

Howard, P.N. and Kollanyi, B., 2016. Bots,# StrongerIn, and# Brexit: computational propaganda during the UK-EU referendum. Available at SSRN 2798311.

Hsieh, H.F. and Shannon, S.E., 2005. Three approaches to qualitative content analysis. Qualitative health research, 15(9), pp.1277-1288.

Hudlicka, E., 2011. Guidelines for designing computational models of emotions. International Journal of Synthetic Emotions (IJSE), 2(1), pp.26-79.

Iasiello, E.J., 2017. Russia's Improved Information Operations: From Georgia to Crimea. Parameters, 47(2).

Ilakkuvan, V., Turner, M.M., Cantrell, J., Hair, E. and Vallone, D., 2017. The Relationship Between Advertising-Induced Anger and Self-efficacy on Persuasive Outcomes. Family & community health, 40(1), pp.72-80.

Ingo Feinerer and Kurt Hornik. 2018. tm: Text Mining Package. R package version 0.7-6. https://CRAN.R-project.org/package=tm

Institute for Strategic Dialogue. (2017). Make Germany Great Again: . Kremlin, Alt-Right and International Influences in the 2017 German Elections. Retreived from: https://www.isdglobal.org/wp.../12/Make-Germany-Great-Again-ENG-061217.pdf

Izard, C.E., 1992. Basic emotions, relations among emotions, and emotion-cognition relations.

Izard, C.E., 2009. Emotion theory and research: Highlights, unanswered questions, and emerging issues. Annual review of psychology, 60, pp.1-25.

Johnson, K., Jin, D. and Goldwasser, D., 2017, May. Modeling of Political Discourse Framing on Twitter. In Eleventh International AAAI Conference on Web and Social Media.

Kagan, J., 1978. On emotion and its development. In The development of affect (pp. 11-41). Springer, Boston, MA.

Kahneman, D. and Egan, P., 2011. Thinking, fast and slow (Vol. 1). New York: Farrar, Straus and Giroux.

Karlova, N.A. and Fisher, K.E., 2013. Plz RT": A social diffusion model of misinformation and disinformation for understanding human information behaviour. Information Research, 18(1), pp.1-17.

Karlova, N.A. and Lee, J.H., 2011. Notes from the underground city of disinformation: A conceptual investigation. Proceedings of the American Society for Information Science and Technology, 48(1), pp.1-9.

Karlsen, G.H., 2019. Divide and rule: ten lessons about Russian political influence activities in Europe. Palgrave Communications, 5(1), p.19.

Kartapolov, A.V. "Lessons of Military Conflicts and Prospects for the Development of Means and Methods of Conducting Them, Direct and Indirect Actions in Contemporary International Conflicts," Vestnik Akademii Voennykh Nauk (Bulletin of the Academy of Military Science), No. 2 2015, pp. 26-36.

Keltner, D., & Haidt, J. (2001). Social functions of emotions. In T. Mayne & G. Bonanno (Eds.), Emotions: Current issues and future directions (pp. 192–213). New York, NY: Guilford Press.

Kertzer, J.D. and Tingley, D., 2018. Political psychology in international relations: beyond the paradigms. Annual Review of Political Science, 21, pp.319-339.

Kilgo, D.K., Harlow, S., García-Perdomo, V. and Salaverría, R., 2018. A new sensation? An international exploration of sensationalism and social media recommendations in online news publications. Journalism, 19(11), pp.1497-1516.

Kim, D., Graham, T., Wan, Z. and Rizoiu, M.A., 2019. Tracking the Digital Traces of Russian Trolls: Distinguishing the Roles and Strategy of Trolls On Twitter. arXiv preprint arXiv:1901.05228.

Kim, H.J. and Cameron, G.T., 2011. Emotions matter in crisis: The role of anger and sadness in the publics' response to crisis news framing and corporate crisis response. Communication Research, 38(6), pp.826-855.

Kiritchenko, S., Zhu, X. and Mohammad, S.M., 2014. Sentiment analysis of short informal texts. Journal of Artificial Intelligence Research, 50, pp.723-762.

Kitzie, V.L., Mohammadi, E. and Karami, A., 2018. "Life never matters in the DEMOCRATS MIND": Examining strategies of retweeted social bots during a mass shooting event. Proceedings of the Association for Information Science and Technology, 55(1), pp.254-263.

Kok, G., Peters, G.J.Y., Kessels, L.T., Ten Hoor, G.A. and Ruiter, R.A., 2018. Ignoring theory and misinterpreting evidence: the false belief in fear appeals. Health psychology review, 12(2), pp.111-125.

Kolchyna, O., Souza, T.T., Treleaven12, P.C. and Aste12, T., 2015. Methodology for Twitter Sentiment Analysis. arXiv preprint arXiv:1507.00955.

Kragh, M. and Åsberg, S., 2017. Russia's strategy for influence through public diplomacy and active measures: the Swedish case. Journal of Strategic Studies, 40(6), pp.773-816.

Kroll, T. and Neri, M., 2009. Designs for mixed methods research. Mixed methods research for nursing and the health sciences, 31.

Kühne, R. and Schemer, C., 2015. The emotional effects of news frames on information processing and opinion formation. Communication Research, 42(3), pp.387-407.

Kühne, R., 2014, January. Political news, emotions, and opinion formation: Toward a model of emotional framing effects. In Annual conference of the international communication association (ICA), Phoenix, AZ.

Kušen, E. and Strembeck, M., 2018. Why so emotional? An analysis of emotional bot-generated content on Twitter. In Proc. of the 3rd International Conference on Complexity, Future Information Systems and Risk (COMPLEXIS). SciTePress, Setubal.

Lecheler, S., Bos, L. and Vliegenthart, R., 2015. The mediating role of emotions: News framing effects on opinions about immigration. Journalism & Mass Communication Quarterly, 92(4), pp.812-838.

Lecheler, S., Schuck, A.R. and de Vreese, C.H., 2013. Dealing with feelings: Positive and negative discrete emotions as mediators of news framing effects.

Lecheler, S., Schuck, A.R. and de Vreese, C.H., 2013. Dealing with feelings: Positive and negative discrete emotions as mediators of news framing effects.

Lench, H.C. and Carpenter, Z.K., 2018. What Do Emotions Do for Us?. In The Function of Emotions (pp. 1-7). Springer, Cham.

Lerner, J.S. and Keltner, D., 2001. Fear, anger, and risk. Journal of personality and social psychology, 81(1), p.146.

Lerner, J.S., Han, S. and Keltner, D., 2007. Feelings and consumer decision making: Extending the appraisal-tendency framework. Journal of consumer psychology, 17(3), pp.181-187.

Lerner, J.S., Li, Y., Valdesolo, P. and Kassam, K.S., 2015. Emotion and decision making. Annual review of psychology, 66, pp.799-823.

Linvill, D. and Warren, P., 2018. Troll factories: The internet research agency and state-sponsored agenda building. pwarren. people. clemson. edu. Linvill Warren TrollFactory. pdf.

Linvill, D.L., Boatwright, B.C., Grant, W.J. and Warren, P.L., 2019. "THE RUSSIANS ARE HACKING MY BRAIN!" Investigating Russia's Internet Research Agency Twitter Tactics during the 2016 United States Presidential Campaign. Computers in Human Behavior.

Lodge, M. and Taber, C.S., 2005. The automaticity of affect for political leaders, groups, and issues: An experimental test of the hot cognition hypothesis. Political Psychology, 26(3), pp.455-482.

Lodge, M. and Taber, C.S., 2007. The rationalizing voter: Unconscious thought in political information processing. Available at SSRN 1077972.

Lohr, S., 2015. Data-ism. Oneworld Publications.

M. A. Gareev, "Anticipate Changes in the Nature of War: Every Era Has Its Own Kind of Military Conflict, and its Own Constraints, and its Own Special Biases," Voyenno-Promyshlennyy Kuryer Online (Military-Industrial Courier Online), 5 June 2013

M.A. Gareev, "Lessons and Conclusions Drawn From the Experience of the Great Patriotic War for Building Up and Training the Armed Forces," Voennaya Mysl' (Military Thought), No. 5 2010, p. 20.

Mair. 2013. Ruling the Void: The Hollowing of Western Democracy. New York: Verso Books.

Manor, I., 2019. The Specter of Echo Chambers—Public Diplomacy in the Age of Disinformation. In The Digitalization of Public Diplomacy (pp. 135-176). Palgrave Macmillan, Cham.

Matsumoto, D., Hwang, H. C., & Frank, M. G. (2014). Emotions expressed in speeches by leaders of ideologically motivated groups predict aggression. Behavioral Sciences of Terrorism and Political Aggression, 6(1), 1–18.

Matthis, I., 2000. Sketch for a metapsychology of affect. International Journal of Psycho-Analysis, 81(2), pp.215-227.

Mattsson, P.A., 2015. Russian Military Thinking–A New Generation of Warfare. Journal on Baltic Security, 1(1), pp.61-70.

McDermott, R.N., 2016. Does Russia have a Gerasimov doctrine?. Parameters, 46(1), p.97.

McDonald, M. 2013. Constructivisms. In: P. Williams, ed., Security Studies: an introduction, 2nd ed. London: Routeledge, pp.63-76.

Mejias, U.A. and Vokuev, N.E., 2017. Disinformation and the media: the case of Russia and Ukraine. Media, Culture & Society, 39(7), pp.1027-1042.

Meraz, S. and Papacharissi, Z., 2013. Networked gatekeeping and networked framing on# Egypt. The international journal of press/politics, 18(2), pp.138-166.

Mercer, J. 'Approaching Emotion in International Polities', paper presented at theInternational Studies Association Conference, San Diego, California, 25 April 1996, p. 1.

Merriam-webster, 2019 https://www.merriam-webster.com/dictionary/propaganda

Mesquita, B. and Leu, J., 2007. The cultural psychology of emotion. Handbook of cultural psychology, pp.734-759.

Miller, D.T., 2019. Topics and emotions in Russian Twitter propaganda. First Monday, 24(5).

Miner, G., Elder IV, J., Fast, A., Hill, T., Nisbet, R. and Delen, D., 2012. Practical text mining and statistical analysis for non-structured text data applications. Academic Press.

Miskimmon, A., O'Loughlin, B., & Roselle, L. (2013). Strategic narratives. Communication power and the New World Order. Abingdon: Routledge.

Mohammad, S.M. and Turney, P.D., 2010, June. Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon. In Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text (pp. 26-34). Association for Computational Linguistics.

Mohammad, S.M. and Turney, P.D., 2013. Crowdsourcing a word–emotion association lexicon. Computational Intelligence, 29(3), pp.436-465.

Morgan, S., 2018. Fake news, disinformation, manipulation and online tactics to undermine democracy. Journal of Cyber Policy, 3(1), pp.39-43.

Munezero, M.D., Montero, C.S., Sutinen, E. and Pajunen, J., 2014. Are they different? Affect, feeling, emotion, sentiment, and opinion detection in text. IEEE transactions on affective computing, 5(2), pp.101-111.

Munnich, E.L., Foster, M.I. and Keane, M.T., 2019. Editors' Introduction and Review: An Appraisal of Surprise: Tracing the Threads That Stitch It Together. Topics in cognitive science, 11(1), pp.37-49.

Nabi, R.L. and Moyer-Gusé, E., 2013. The psychology underlying media-based persuasion. In The Oxford handbook of media psychology.

Nabi, R.L. and Prestin, A., 2016. Unrealistic hope and unnecessary fear: Exploring how sensationalistic news stories influence health behavior motivation. Health communication, 31(9), pp.1115-1126.

Nabi, R.L., 2002. The theoretical versus the lay meaning of disgust: Implications for emotion research. Cognition & Emotion, 16(5), pp.695-703.

Nabi, R.L., 2003. Exploring the framing effects of emotion: Do discrete emotions differentially influence information accessibility, information seeking, and policy preference?. Communication Research, 30(2), pp.224-247.

Narayanan, V., Howard, P.N., Kollanyi, B. and Elswah, M., 2017. Russian involvement and junk news during Brexit. comprop. oii. ox. ac. uk/wp-content/uploads/sites/93/2017/12/Russia-and-B rexit-v27. pdf.

NATO StratCom Centre of Excellence (2018). FAKE NEWS A ROADMAP EXECUTIVE SUMMARY

Nimmo, 2015 https://www.stopfake.org/en/anatomy-of-an-info-war-how-russia-s-propaganda-machine-works-and-how-to-counter-it/

Nisbet, E.C. and Kamenchuk, O., 2019. The Psychology of State-Sponsored Disinformation Campaigns and Implications for Public Diplomacy. The Hague Journal of Diplomacy, 14(1-2), pp.65-82.

Oppenheim, 2017 https://www.independent.co.uk/news/world/americas/maria-chappelle-nadal-donald-trump-missouri-state-senator-assassination-us-president-secret-service-a7900151.html

Ostrovsky, A., 2016. The invention of Russia: the rise of Putin and the age of fake news. Penguin.

Panagopoulos, C., 2010. Affect, social pressure and prosocial motivation: Field experimental evidence of the mobilizing effects of pride, shame and publicizing voting behavior. Political Behavior, 32(3), pp.369-386.

Pariser, E. (2011). The filter bubble: What the internet is hiding from you. New York: Penguin.

Paul, C. and Matthews, M., 2016. The Russian "firehose of falsehood" propaganda model. Rand Corporation, pp.2-7.

Pierre Nord, 'L'Intoxication'. Quotation from Vladimir Volkoff, 'La désinformation – arme de guerre', Paris 1986.

Plutchik, R., 1980. A general psychoevolutionary theory of emotion. In Theories of emotion (pp. 3-33). Academic press.

Plutchik, R., 1984. Emotions: A general psychoevolutionary theory. Approaches to emotion, 1984, pp.197-219.

Possel, H. (n.d.). English language smart words. Retrieved from http://www.smart-words.org

Quercia, D., Ellis, J., Capra, L. and Crowcroft, J., 2011, October. In the mood for being influential on twitter. In 2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing (pp. 307-314). IEEE.

Quick, B.L., Bates, B.R. and Quinlan, M.R., 2009. The utility of anger in promoting clean indoor air policies. Health communication, 24(6), pp.548-561.

Ravi, K. and Ravi, V., 2015. A survey on opinion mining and sentiment analysis: tasks, approaches and applications. Knowledge-Based Systems, 89, pp.14-46.

Reisenzein, R., Horstmann, G. and Schützwohl, A., 2019. The cognitive-evolutionary model of surprise: A review of the evidence. Topics in cognitive science, 11(1), pp.50-74.

Reisenzein, R., Meyer, W.U. and Schützwohl, A., 1996. Reactions to surprising events: A paradigm for emotion research. In Proceedings of the 9th conference of the International Society for Research on Emotions (pp. 292-296). ISRE Toronto.

Richards, Barry. 2007. Emotional Governance: Politics, Media and Terror. Houndsmill: Palgrave MacMillan.

Ridout, T.N. and Searles, K., 2011. It's my campaign I'll cry if I want to: How and when campaigns use emotional appeals. Political Psychology, 32(3), pp.439-458.

Rinker, T. W. (2018). textclean: Text Cleaning Tools version 0.9.3. Buffalo, New York. https://github.com/trinker/textclean

Rozin et al 2008 Ch 47 Disgust in Handbook of Emotions 2008

Rumer, 2019. https://carnegieendowment.org/2019/06/05/primakov-not-gerasimov-doctrine-in-action-pub-79254

Russell, J.A., 1980. A circumplex model of affect. Journal of personality and social psychology, 39(6), p.1161.

Russell, J.A., 2003. Core affect and the psychological construction of emotion. Psychological review, 110(1), p.145.

Rychiak, R. and Pacepa, I.M., 2013. Disinformation: Former Spy Chief Reveals Secret Strategies for Undermining Freedom, Attacking Religion, and Promoting Terrorism. Washington, Dc: WND books.

Sander, D. and Scherer, K. eds., 2014. Oxford companion to emotion and the affective sciences. OUP Oxford.

Scherer, K.R., 2000. Psychological models of emotion. The neuropsychology of emotion, 137(3), pp.137-162.

Scherer, K.R., 2005. What are emotions? And how can they be measured?. Social science information, 44(4), pp.695-729.

Scherer, K.R., 2009. Emotions are emergent processes: they require a dynamic computational architecture. Philosophical Transactions of the Royal Society B: Biological Sciences, 364(1535), pp.3459-3474.

Scheufele, D.A. and Krause, N.M., 2019. Science audiences, misinformation, and fake news. Proceedings of the National Academy of Sciences, p.201805871.

Schmitt, O., 2018. When are strategic narratives effective? The shaping of political discourse through the interaction between political myths and strategic narratives. Contemporary Security Policy, 39(4), pp.487-511.

Security Information Service of the Czech Republic, Annual Report of the Security Information Service (BIS) (2008), 5.

Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism (2017). Extremist Content and Russian Disinformation Online: Working with Tech to Find Solutions, Testimony of Sean J. Edgett

Sherman, G.D. and Haidt, J., 2011. Cuteness and disgust: the humanizing and dehumanizing effects of emotion. Emotion Review, 3(3), pp.245-251.

Shouse, E., 2005. Feeling, emotion, affect. M/c journal, 8(6), p.26.

Sivek, S. C., 2018. Both facts and feelings: Emotion and news literacy. Journal of Media Literacy Education, 10(2), pp.123-138.

Skurka, C., Niederdeppe, J., Romero-Canyas, R. and Acup, D., 2018. Pathways of influence in emotional appeals: Benefits and tradeoffs of using fear or humor to promote climate change-related intentions and risk perceptions. Journal of Communication, 68(1), pp.169-193.

Slaughter, A.M., 2011. International relations, principal theories. Max Planck encyclopedia of public international Law, 129.

Slovic, P., Finucane, M.L., Peters, E. and MacGregor, D.G., 2007. The affect heuristic. European journal of operational research, 177(3), pp.1333-1352.

Smith, C.A. and Ellsworth, P.C., 1985. Patterns of cognitive appraisal in emotion. Journal of personality and social psychology, 48(4), p.813.

Søe, S.O., 2018. Algorithmic detection of misinformation and disinformation: Gricean perspectives. Journal of Documentation, 74(2), pp.309-332.

Spangher, A., Ranade, G., Nushi, B., Fourney, A. and Horvitz, E., 2018. Analysis of Strategy and Spread of Russia-sponsored Content in the US in 2017. arXiv preprint arXiv:1810.10033.

Stahl, B.C., 2006. On the Difference or Equality of Information, Misinformation, and Disinformation: A Critical Research Perspective. Informing Science, 9.

Stelzenmüller, C. (2017). The Impact of Russian Interference on Germany's 2017 Elections. Testimony before the U.S. Senate Select Committee on Intelligence. Retrieved from: https://www.brookings.edu/testimonies/the-impact-of-russian-interference-on-germanys-2017-elections/

Stewart, L.G., Arif, A. and Starbird, K., 2018, February. Examining trolls and polarization with a retweet network. In Proc. ACM WSDM, Workshop on Misinformation and Misbehavior Mining on the Web.

Stieglitz, S. and Dang-Xuan, L., 2013. Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. Journal of management information systems, 29(4), pp.217-248.

Suler, John. 2016. Psychology of the Digital Age: Humans Become Electric. New York: Cambridge University Press.

Taber, C.S. and Young, E., 2013. Political information processing. In The Oxford handbook of political psychology.

Taboada, M., Brooke, J., Tofiloski, M., Voll, K. and Stede, M., 2011. Lexicon-based methods for sentiment analysis. Computational linguistics, 37(2), pp.267-307.

Tannenbaum, M.B., Hepler, J., Zimmerman, R.S., Saul, L., Jacobs, S., Wilson, K. and Albarracín, D., 2015. Appealing to fear: A meta-analysis of fear appeal effectiveness and theories. Psychological bulletin, 141(6), p.1178.

Tayor, 2016 https://www.washingtonpost.com/news/worldviews/wp/2016/11/26/before-fake-news-there-was-soviet-disinformation/?utm_term=.1e827d47b8ad

Theohary, C. (2018). Information Warfare: Issues for Congress. [online] Congressional Research Service. Available at: https://fas.org/sgp/crs/natsec/R45142.pdf [Accessed 17 Jul. 2019].

Thomas, T.L., 2016. Thinking Like A Russian Officer: Basic Factors And Contemporary Thinking On The Nature of War. Foreign Military Studies Office.

Thornton, R., 2015. The changing nature of modern warfare: responding to Russian information warfare. The RUSI Journal, 160(4), pp.40-48.

Tolstoy, A. and McCaffray, E., 2015. Mind games: Alexander dugin and Russia's war of ideas. World Affairs, pp.25-30.

Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relation-ship to internal regulatory variables. In M. Lewis, J. Haviland, & L. F. Barrett (Eds.), Handbook of emotions (3rd ed., pp. 114–137). New York, NY: Guilford Press.

Twitter Elections Integrity, 2019 https://about.twitter.com/en_us/values/elections-integrity.html#data

Twitter, 2018 https://blog.twitter.com/official/en_us/topics/company/2018/2016-election-update.html

U.S. National Intelligence Council, 2017. "Assessing Russian activities and intentions in recent US elections," Office of the Director of National Intelligence, National Intelligence Council, Intelligence Community Assessment, ICA 2017-01 D, at https://www.dni.gov/files/documents/ICA_2017_01.pdf, accessed 11 April 2019.

U.S. Senate Committee on the Judiciary, Subcommittee on Crime and Terrorism, 2017. "Testimony of Sean J. Edgett, Acting General Counsel, Twitter, Inc." (31 October), at https://www.judiciary.senate.gov/imo/media/doc/10-31-

UK government, 2018- disinformation and fake news final report

Unattributed report, "How Does One Develop a Modern Army?" Krasnaya Zvezda Online, 11 March 2016.

United States Information Agency. Soviet Active Measures in the "Post-Cold War" Era 1988-1991

United States. Department of State. (1987). Soviet influence activities: a report on active measures and propaganda, 1986-87 (Vol. 9627). US Department of State.

US v. Internet Research Agency. United States District Court for the District of Columbia. 2018. Case 1:18-cr-00032-DLF - USA v. IRA et al. (Feb. 2018). Retrieved April 17, 2018 from https://www.justice.gov/file/1035477/download

Van Boven, L. and Ashworth, L., 2007. Looking forward, looking back: Anticipation is more evocative than retrospection. Journal of Experimental Psychology: General, 136(2), p.289.

Van Dijck, J., 2014. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. Surveillance & Society, 12(2), pp.197-208.

Van Kleef, G.A., Van Doorn, E.A., Heerdink, M.W. and Koning, L.F., 2011. Emotion is for influence. European Review of Social Psychology, 22(1), pp.114-163.

Van Stekelenburg, J., & Klandermans, B. (2013). The social psychology of protest. Current Sociology, 61, 886–905.

Vladimirov, A. Основы _общей _теории _войны (General Theory of War) Москва:Московский финансово промышленный университет "Синергия". 2013, p. 121.

Wakefield, 2016. Social media 'outstrips TV' as news source for young people. Accessed 04/02/2019 at https://www.bbc.co.uk/news/uk-36528256

Wakefield, 2016. Social media 'outstrips TV' as news source for young people. Accessed 04/02/2019 at https://www.bbc.co.uk/news/uk-36528256

Walker, C., 2016. The Hijacking of" Soft Power". Journal of Democracy, 27(1), pp.49-63.

Wardle, C., 2018. The Need for Smarter Definitions and Practical, Timely Empirical Research on Information Disorder. Digital Journalism, 6(8), pp.951-963.

Warriner, A.B., Kuperman, V. and Brysbaert, M., 2013. Norms of valence, arousal, and dominance for 13,915 English lemmas. Behavior research methods, 45(4), pp.1191-1207.

Wierzbicka, A., 1999. Emotions across languages and cultures: Diversity and universals. Cambridge University Press.

Witte, K. and Allen, M., 2000. A meta-analysis of fear appeals: Implications for effective public health campaigns. Health education & behavior, 27(5), pp.591-615.

Wong & Bever, 2017. ''I hope Trump is assassinated': A Missouri lawmaker faces mounting calls to resign after Facebook comment'. Washinton Post. Accessed 07/10/2019 at https://www.washingtonpost.com/news/the-fix/wp/2017/08/18/i-hope-trump-is-assassinated-a-missouri-lawmakers-facebook-comment-leads-to-calls-for-her-resignation/?utm_term=.27a55c3e6d16

Woolley, S.C., 2016. Automating power: Social bot interference in global politics. First Monday, 21(4).

Xia, Y., Lukito, J., Zhang, Y., Wells, C., Kim, S.J. and Tong, C., 2019. Disinformation, performed: self-presentation of a Russian IRA account on Twitter. Information, Communication & Society, pp.1-19.

Yin, L., Roscher, F., Bonneau, R., Nagler, J. and Tucker, J.A., 2018. Your Friendly Neighborhood Troll: The Internet Research Agency's Use of Local and Fake News in the 2016 US Presidential Campaign. SMaPP Data Report. New York: Social Media and Political Participation Lab, New York University.

Zajonc, R.B., 1980. Feeling and thinking: Preferences need no inferences. American psychologist, 35(2), p.151.

Zannettou, S., Caulfield, T., De Cristofaro, E., Sirivianos, M., Stringhini, G. and Blackburn, J., 2018. Disinformation warfare: Understanding state-sponsored trolls on Twitter and their influence on the Web. arXiv preprint arXiv:1801.09288.

Zannettou, S., Caulfield, T., Setzer, W., Sirivianos, M., Stringhini, G. and Blackburn, J., 2018. Who let the trolls out? towards understanding state-sponsored trolls. arXiv preprint arXiv:1811.03130.

Appendix A.

```r
# Load in relevant libraries
library(syuzhet)
library(dplyr)
library(tm)
library(textclean)

# Load in data
data_agn <- read.csv('~/Desktop/new_ira/50000 XE split up/Big E xae.csv', header = TRUE)

#Data preprocessing

#turn data_agn into corpus
clean_corpus<- Corpus(VectorSource(data_agn$tweet_text))
clean_corpus <- tm_map(clean_corpus, tolower)

# remove URLs
removeURL <- function(x) gsub("http[^[:space:]]*", "", x)
clean_corpus <- tm_map(clean_corpus, content_transformer(removeURL))

#turn corpus back into dataframe
clean_df <- data.frame(text = sapply(clean_corpus, as.character), stringsAsFactors = F)

clean_df<- apply(clean_df, 2, function(x) gsub(" :) ", "smile", x))

#replace slang words
clean_df<- replace_internet_slang(clean_df, slang = paste0("\\b",
                                 lexicon::hash_internet_slang[[1]], "\\b"),
                    replacement = lexicon::hash_internet_slang[[2]], ignore.case = TRUE,)

#word elongation
clean_df<- replace_word_elongation(clean_df)

#emoticon replacement
clean_df<- replace_emoticon(clean_df, emoticon_dt = lexicon::hash_emoticons,)

#back to normal data frame
cleaned_df <- as.data.frame(t(clean_df))
cleaned_df <- as.data.frame(t(as.matrix(cleaned_df)))

#remove punctuation
cleaned_corpus<- Corpus(VectorSource(cleaned_df$V1))
cleaned_corpus <- tm_map(cleaned_corpus, removePunctuation)

#back to data frame for sentiment analysis
cleaned_df <- data.frame(text = sapply(cleaned_corpus, as.character), stringsAsFactors = F)

#remove non-ascii characters
cleaned_df$text <- gsub("[^\x20-\x7E]", "", cleaned_df$text)

# Extract nrc sentiment from each tweet
nrc_list <- c()
for (tweet in cleaned_df$text) { # Loop over each tweet to get nrc sentiment
  nrc <- get_nrc_sentiment(tweet, language = 'english')
```

```
  nrc_list <- c(nrc_list, nrc)
}
```

```
# Put nrc_list into a data frame
df_nrc <- data.frame(matrix(unlist(nrc_list), nrow = length(nrc_list), byrow = TRUE))
df_nrc$index <- 1:nrow(df_nrc)
df_nrc$type <-
rep(c('anger','anticipation','disgust','fear','joy','sadness','surprise','trust','negative','positive'),
times = (length(nrc_list) / 10))
names(df_nrc)[names(df_nrc) == 'matrix.unlist.nrc_list...nrow...length.nrc_list...byrow...TRUE.']
<- 'value'

# Separate data per 10 units by nrc quality
anger <- c()
for (r in df_nrc$index) {
 if (r %% 10 == 1) {
   anger <- append(anger, df_nrc[r, 1])
 }
}
anticipation <- c()
for (r in df_nrc$index) {
 if (r %% 10 == 2) {
   anticipation <- append(anticipation, df_nrc[r, 1])
 }
}
disgust <- c()
for (r in df_nrc$index) {
 if (r %% 10 == 3) {
   disgust <- append(disgust, df_nrc[r, 1])
 }
}
fear <- c()
for (r in df_nrc$index) {
 if (r %% 10 == 4) {
   fear <- append(fear, df_nrc[r, 1])
 }
}
joy <- c()
for (r in df_nrc$index) {
 if (r %% 10 == 5) {
   joy <- append(joy, df_nrc[r, 1])
 }
}
sadness <- c()
for (r in df_nrc$index) {
 if (r %% 10 == 6) {
   sadness <- append(sadness, df_nrc[r, 1])
 }
}
surprise <- c()
for (r in df_nrc$index) {
 if (r %% 10 == 7) {
   surprise <- append(surprise, df_nrc[r, 1])
 }
}
trust <- c()
```

```
for (r in df_nrc$index) {
  if (r %% 10 == 8) {
    trust <- append(trust, df_nrc[r, 1])
  }
}
negative <- c()
for (r in df_nrc$index) {
  if (r %% 10 == 9) {
    negative <- append(negative, df_nrc[r, 1])
  }
}
positive <- c()
for (r in df_nrc$index) {
  if (r %% 10 == 0) {
    positive <- append(positive, df_nrc[r, 1])
  }
}

# Create new data frame
tweets <- as.vector(cleaned_df$text)
data_clean <- data.frame(tweets, anger, anticipation, disgust, fear, joy, sadness, surprise, trust,
negative, positive)

#Save file
setwd("~/Desktop/new_ira/Sentiment")
write.csv(data_clean, 'XE_e_sentiment50,000.csv')
```

```
negative <- c()
for (r in df_nrc$index) {
 if (r %% 10 == 9) {
   negative <- append(negative, df_nrc[r, 1])
 }
}
positive <- c()
for (r in df_nrc$index) {
 if (r %% 10 == 0) {
   positive <- append(positive, df_nrc[r, 1])
 }
}

# Create new data frame
tweets <- as.vector(cleaned_df$text)
data_clean <- data.frame(tweets, anger, anticipation, disgust, fear, joy, sadness, surprise, trust,
negative, positive)

#Save file
setwd("~/Desktop/new_ira/Sentiment")
write.csv(data_clean, 'XE_e_sentiment50,000.csv')
```