The Spread of Disinformation on the Web: An Examination of Memes on Social Networking

Marc Dupuis
Computing and Software Systems
University of Washington
Bothell, Washington, USA
marcjd@uw.edu

Andrew Williams
Computing and Software Systems
University of Washington
Bothell, Washington, USA
andrewwi@uw.edu

***Abstract*—Social media has become a potent vector for political disinformation and propaganda, often spread by malicious actors such as trolls or even foreign intelligence services, as famously occurred during the 2016 United States election. However, what makes social media a particularly potent vector for disinformation is not so much the behavior of malicious actors themselves, but rather, ordinary users, who play a vital role in spreading and magnifying this disinformation. In order to understand and combat the spread of disinformation, we conduct two surveys examining the patterns of user behavior in sharing different types of disinformation. The first survey classifies a variety of image memes based on user reaction and interpretation. The second survey will be evaluating user behavior toward those memes with additional measures in place to assess the personality and trait affect of users. The goal is to help understand how ordinary social media users behave in regards to political propaganda and disinformation.**

***Index Terms*—disinformation, propaganda, fake news, social media, memes, information integrity, cybersecurity**

I. Introduction

One of the pillars of cybersecurity has always been finding ways to protect the *integrity* of information transmitted online. 'Integrity' is usually defined as ensuring that information is not altered between the sender and the receiver, however, more broadly, integrity refers to the trustworthiness of information.

In a time in which more and more people are getting their information about the world through social media [30], how do we ensure the integrity of the information they receive through these platforms? The nature of social media, which allows any user to share almost any information, within the bounds of a loosely-enforced code of conduct, means we can never guarantee that all information on social media will be true or trustworthy.

In this case, perhaps protecting the *integrity* of information does not mean not that the information is necessarily true or factual, but merely that it has not been unduly manipulated. For example, how can people be sure that the information they see on their social media feeds is the result of an organic process of posting and sharing by users, rather than the result of malicious actors or targeted manipulation campaigns?

These sort of purposeful disinformation efforts received a lot of attention during the 2016 U.S. Presidential Election [16] and in its wake [51], however, they have not abated in the time since. Such efforts are already well underway in targeting candidates running in the 2020 Presidential election [23].

In addition to long-term efforts during Presidential campaigns, shorter-term efforts to manipulate social media also spring up in the wake of major news events. For example, in the wake of the release of Robert Mueller's report on Russian interference in the 2016 election, a network of 5,000 Twitter bots was mobilized to promote the hashtag #Russiagate and discredit the Russia investigation [8]. And after the Notre Dame fire in April 2019, conspiracy theories spread quickly on social media through a variety of means [26].

Various technological solutions have been proposed to help combat the spread of false information on social media, such as fact checking [5], crowd signals [46], natural language processing [34], and more. Technological solutions that can automatically flag such efforts are important piece of the puzzle, however, it is also important to consider the role that ordinary users play in the magnification and spread of misinformation.

Misinformation campaigns, as conducted by foreign state actors in the 2016 election or by individuals controlling networks of bots [8], often rely on ordinary users to pick up a false story and make it go viral. To make matters more complicated, such false stories may not be actual stories or articles, but simple image memes or blocks of text. In such a case, analyzing the source or the language patterns in the item may not be a useful way of determining if an item is legitimate or not.

Even if an item is known to be false, or not legitimate in some way, will that prevent ordinary users from spreading it? For this paper, we examined the ways in which ordinary users contribute to the spread of misinformation. Ultimately, in order to ensure the integrity of social media, we must examine not just the sources of disinformation, but the role that ordinary users play in helping such disinformation spread or go viral.

To that end, we have conducted one survey with another survey in development. This first survey was designed to identify a subset of content that represents a wide political and emotional spectrum. The second survey will identify how likely users are to share this content on social media, and to see how user perceptions of their platform, their audience and

their anonymity correlate to the willingness of users to share disinformation, either unwittingly or on purpose.

II. Background

Propaganda, disinformation, and hoaxes have always been pitfalls in reporting the news and current events. However, with the decline of 'traditional' sources of journalism such as newspapers and the rise of social media, it has become possible for malicious actors to manipulate the media in new and disturbing ways.

In Media Manipulation and Disinformation Online, by Alice Marwick and Rebecca Lewis [29], several recent case studies are presented in which malicious actors such as conspiracy theorists or even hate groups use (or abuse) the tools of social media to amplify their message. For example, in one such case in 2015, a white nationalist website, through coordinated action by its users on social media, was able to get a hoax about the creation of 'White Student Unions' picked by media outlets as well-known as USA Today [6].

Political candidates can be targets of media manipulation as well. Networks of bots may be used to target candidates and spread false messages, which are then shared or re-tweeted by regular users, and amplified much more as a result. An early example of this occurred in 2010 when a 'Twitter bomb' targeted Massachusetts Senate candidate Martha Oakley in the days before the 2010 Senate special election [32], in a successful effort to influence real-time search results for the candidates.

These tactics have become more far-reaching and problematic over time. Controversy regarding media influence and 'fake news' swirled, and continues to swirl, around two of the most consequential events of 2016: the Brexit referendum and the U.S. Presidential election [7].

Today, scholars and researchers have pointed to disinformation on the Internet and social media as a growing threat to the functioning of democratic institutions around the world. [3] Social media gives users the ability to isolate themselves from opposing viewpoints, and construct 'cocoons' in which they are less frequently exposed to viewpoints other than their own [45] [14]. Research has also shown that people may react with hostility to factual sources that do not agree with them [42]. If this is true, people may be particularly susceptible to the type of manipulations we discuss here, in which ordinary users either knowingly or unknowingly magnify disinformation and 'fake news' [43].

Note that the polarization of social media discussed in the above papers is not in itself the result of outside tampering or influence campaigns, but instead the natural behavior of ordinary users. Manipulation campaigns seek to *exploit*, not necessarily *create*, this polarization.

Therefore, in order to ensure the integrity of information on social media, we need to understand how ordinary users interact with information, how their behavior can be distinguished from malicious influence campaigns, and how ordinary user behavior– inadvertent or otherwise– may actually aid those campaigns. But before we delve into how social media users interact with information, we need to provide a little more context about the topic of disinformation, or "fake news".

The term 'fake news' has entered the common vernacular since the 2016 election, to describe news that has been manipulated or is intentionally false. It is frequently used as a pejorative in the face of damaging legitimate stories, and can include satirical sites such as The Onion or The Borowitz Report. 'Fake news' is a broad term whose meaning varies from user to user, and often from moment to moment. 'Disinformation' may be a more precise term, although it does not necessarily cover all types of malicious or false information on social media. 'Disinformation' usually refers to information which is *intentionally* incorrect, whereas 'misinformation' refers to information which is *unintentionally* incorrect [28].

This distinction is important, because if users are unintentionally sharing false information, then additional education of users could provide a remedy. Incorporating more fact-checking functionality into social media [5], or using automated natural language processing to flag such items and alert users [34] [7], could provide a useful tool to combat 'fake news'.

However, if users will intentionally share information even if they know it is false, then user education may provide little remedy to the propagation of disinformation or misinformation on social media. Indeed, one survey on the efficacy of fact-checking found that users were likely to care about fact-checking only insofar as it was advantageous to their own group, and fact-checking that went against their chosen group was more likely to draw hostility [42]. Additionally, the simple mechanic of repeated exposure to a story was likely to increase users' perception of its truth, even if the story was flagged as inaccurate. [35]

A paper by Rashkin, et al [38] examines the truth of an item via two axes: on one axis is the Quality of the Information (ranging from Fake to Trustworthy), on the other axis is the intention of the author, or how much they are intending to deceive the reader. *Satire* such as The Onion might have false or low-quality information, but it would also rank low regarding the intention of the author to deceive. On the other hand, an actual *hoax* would rank similarly low in information quality but rank high regarding the intention of the author to deceive. *Propaganda* can occupy a wide range of values on the graph, but usually it is almost always misleading and the author usually intends to deceive the reader to some degree, for the sake of pushing a particular political message.

For the purposes of this paper we are less interested in examining the precise motives of the original authors, and more interested in examining the behavior of ordinary social media users when confronted with disinformation or propaganda-style content. Throughout the paper, we will use 'disinformation' as an umbrella term for 'fake news', propaganda, hoaxes, and other false information, which are common factors in malicious campaigns to manipulate or influence social media.

Suppose for a moment that you were a creator of such disinformation, looking to design content to go viral as part of an effort to manipulate social media. If you were to want

to design content for the sake of going viral, what would that content look like? What content is most likely to get shared and magnified by users across your target social media network? What motivates users to share this content in the first place?

Generally speaking, users have a variety of motivations for sharing news on social media. In a study from 2015, in which 18 participants were interviewed in-depth about their motivations for sharing news [50], the results could be broadly broken down into a desire to *inform* or a desire to *entertain*. Additionally, sub-motivations were found; some users were motivated by maintaining a connection with a group to which they identify– which could also contribute to polarization– whereas other users are motivated by changing minds or sparking debate. Some shared items in order to distinguish themselves, others shared items in order to 'join the crowd', so to speak.

Another 2015 study by Oh, et al [33] found that people's motivations also varied across different social media platforms. They found that Facebook and Twitter users are more motivated by learning than Youtube users, and Facebook users in particular were more motivated by social engagements. However, interviews and in-depth surveys on user motivations may capture after-the-fact justifications, rather than the emotion of the moment when the user decides to like or share something. Users tend to scroll through social media quickly, and sharing decisions may be made more on momentary emotional or psychological arousal than anything else.

Many studies have been done into how different emotions affect the tendency of users to share content (e.g., [13]). A study by Jonah Berger [4] found that more than a particular emotion, it was psychological arousal that increased the social transmission of information. For example, when considering negative emotions, evoking *anxiety* resulted in more willingness to share than evoking *sadness*, and on the positive side, evoking *humor* or amusement resulted in more willingness to share than merely evoking *contentment*.

Other studies, such as a 2013 study by Guadagno, et al [15] found similar results when examining a single piece of content, finding that users who felt sronger affective responses to a video were much more likely to share it.

Ideologically extreme content is more likely to be shared, perhaps precisely because it arouses stronger emotions. A study by [37] found that Twitter users with more extreme ideological positions shared content disproportionately more than moderate users. This means that users may be exposed to a disproportionate ratio of more extreme content as compared to more moderate content.

During the 2016 political campaign, a survey was conducted studying how anger and anxiety influenced users to share information related to the campaign. It found that users who were deeply plugged in to online news were both more angry and more anxious about the opposing Presidential candidates, and that people who felt more anger about the opposing candidate shared information about the election more frequently [17]. Anger and anxiety, once again, high arousal emotions, are seen to play a big role in determining user sharing behavior.

A user's perception of their audience, and their anonymity, also affect their behavior online. In the area of cyberbullying, for example, problematic behavior tends to increase when users believe they are acting anonymously. [1] Other studies have pointed to a relationship between the users' anonymity and their sharing behavior, suggesting that user anonymity does have an effect on what users share online, and users may be more comfortable sharing items of negative valence if they feel safer in their anonymity [27]. Additionally, a 2014 study found that controversial content was 3.2x more likely to be shared anonymously than non-anonymously [52].

Different social media allows users to interact more or less anonymously. Facebook, Twitter, and Instagram all allow different levels of anonymity, and have designed their interfaces to allow the users varying degrees of freedom in the ways in which they can manipulate their identity [31]. This alters the way in which users use social media, and it may also alter their perceptions of their audience.

Users often have an imagined target audience in mind when they post on social media [13], [25]. This audience could be people with personal ties, such as family or friends. It could be people with professional ties, such as co-workers or potential employers, as one might find on LinkedIn. Or it could include communal ties, such as people who share a hobby– or people who support a political candidate, or even members of a hate group.

We have already seen, in the survey from Oh et al [33] that user motivations can be measured differently across social media. But how do perceptions of anonymity and audience factor into user behavior, specifically as it relates to sharing political (dis)information? To answer that question, we decided to look at user behavior as it relates to sharing one particular type of content: image memes.

We chose to focus on image memes because their nature allows for user behavior to be more easily surveyed on a large scale, through an automated survey platform such as Mechanical Turk. Users can quickly view and digest the meaning of an image meme within a couple of seconds, without having to click on an external link or do additional reading. Memes usually convey an uncomplicated idea, often using humor to get their point across [22]. A user can interact with an image meme within a second or two and continue scrolling, which makes them well-adapted to modern social media.

The term 'meme' has existed since Richard Dawkins originally coined it in 1976, as a unit of cultural transmission that serves a similar role in cultural evolution that a gene serves in biological evolution. On the modern web a 'meme' can be defined as an image or video containing a block of text, which can be easily shared on social media. Variations of a meme may consist of lots of different versions of the same image, often featuring a certain stock character, containing different blocks of text, such as 'sheltering suburban Mom' or 'Annoying Facebook girl' [41].

Memes such as Pepe the Alt-Right Frog, are frequently used

by the alt-right in their efforts to influence social media [24]. Political memes may also be images of political figures, such as Donald Trump holding up a signed piece of paper, which has been altered in various ways, or former Speaker of the House Paul Ryan gesturing to a board that has been altered to say various things [40]. Memes are also frequently used during the sorts of misinformation campaigns discussed earlier in this section, as they are easy to create, easy to circulate, and hard to fact-check.

Because memes are highly visual and often contain humor, they are well-suited to being viewed and shared on social media, and can potentially be a powerful way to influence opinion online [22] [19]. Memes were a frequent tool of disinformation campaigns during the 2016 election, such as the 'Draft our Daughters' campaign that was targeted at the Hillary Clinton campaign [16]. In a way, memes provide one of the most potent weapons for intentional disinformation campaigns, as they provide an easy way for almost anyone to create catchy, colorful content that is easily digested, viewed, and shared across multiple platforms

While much work has been done showing how influence campaigns use memes, less work has been done as far as demonstrating their actual efficacy in promoting their viewpoints, perhaps because most memes are meant primarily for amusement or humor. However, there is no doubt that political memes are widely used for propaganda purposes, and more study of their efficacy is needed. They are certainly effective in terms of their ability to spread on social media, but their ability to convince viewers is still uncertain.

One study of feminist memes featuring Ryan Gosling did find that exposure to the test memes increased viewer endorsement of specific feminist beliefs [49]. But regardless of whether memes actually convince viewers of their viewpoint, that may not even be the most important point. If a meme becomes widespread enough on social media affects the national conversation, as trends on social media will almost always get picked up by national media, politicians, and spread beyond social media. The questions that swirled around Hillary Clinton's health during the 2016 campaign were just one example of this, in which memes and posts on social media fed stories in traditional media, which in turn fed social media, and so on [16] [29].

With this in mind, our first goal was to compile a set of image memes for testing, and measure user reaction to ensure that we had a set of memes that contained a variety of political content and provoked a range of reactions. In the second survey, our goal will be to measure how user sharing behaviors changed based on the user's perception of the truthfulness of the meme, as well as their perceived audience and how they evaluated their own anonymity. Measures for personality and trait affect will also be incorporated into this second survey.

III. Methods

Using Google image search, we compiled a set of memes that represented a variety of political content and addressed a variety of issues, which we ultimately whittled down to 12 memes for testing in our first survey.

The memes included a mix of left-leaning and right-leaning memes, as well as memes that made fun of both parties. The memes included two that were circulated on Facebook by Russia during the 2016 election, as well as a mix of general memes on a range of issues. For example, we included memes that promoted both left- and right-leaning messages on issues such as climate change and gun control. We also included a Bernie Sanders meme with both a left-leaning message and a right-leaning message.

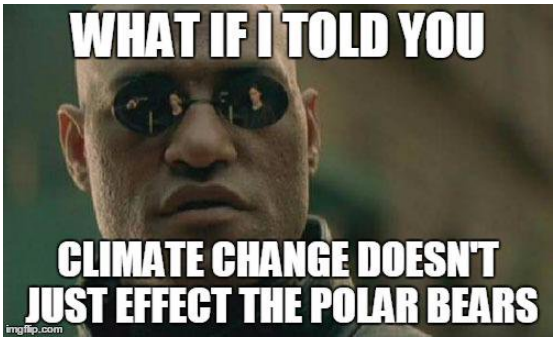

Fig. 1. Example of a left-leaning climate change meme

Fig. 2. Example of a right-leaning climate change meme

In order to obtain some baseline data on how political affiliation, political ideology, and various demographic factors influence one's judgment of a meme, a large-scale survey was employed. Amazon's Mechanical Turk (MTurk) was used to recruit survey participants. MTurk provides researchers with a relatively low-cost and quick turnaround platform for participant recruitment [11], [44]. Participants generally represent a broader cross-section of the population than other methods often employed, such as college sophomores in an

introductory psychology class [39]. IRB approval was on file prior to collecting data and informed consent was obtained. Participants were compensated with $2 for their participation in the study. One quality control questions was used. If participants failed the quality control question, the survey would conclude with an explanation of why it has ended.

We used the Qualtrics survey platform. A total of 203 responses were collected. Participants are asked at the end of the survey how the effort and time required to complete the survey compared to similar work offered through the MTurk platform. Most participants indicated that it was either easier (19.4%) or comparable (61.2%) to other projects with some indicating more effort was required (19.4%). Of note, a pilot study was employed beforehand to check for any issues with the survey, including survey logic and question wording problems, as well as the same question noted above. The compensation was subsequently adjusted from the pilot study to better reflect a comparable amount of time and effort for research participants. Thus, we believe we accomplished this given the above results from this question in the final survey.

In the process of selecting memes for the study, as well as the survey questions themselves (i.e., what were going to measure), we employed the Delphi technique [9], [18], [36]. The Delphi technique is a method that is used to reach consensus on a matter. In the context of this study, we wanted to make sure there was appropriate coverage in what we were assessing for each meme and for the memes themselves. We employed three rounds of the Delphi technique to a small group of participants, which is considered a good number of rounds that effectively balances robustness with fatigue that can set in from too many rounds. Additionally, we set a 75% threshold for consensus. In other words, if 75% of the participants involved in the Delphi technique were in agreement, then consensus was considered achieved.

Ultimately, we decided that we would measure the following items for each meme:

1) Ideological Agenda
2) Political Party being Advanced
3) Propaganda (Classification)
4) Hoax (Classification)
5) Satire (Classification)
6) Truth (Classification)
7) Funny (Is it?)
8) Trying to be Funny (Is it?)
9) True (Is it?)
10) Underlying Message True (Is it?)

The memes were presented in a random order to the participants with the same questions for each meme. The 12 memes used in this study can be found at: http://www.aristotle.cc/Memes.pdf

Results were obtained from those that classified themselves as Democrats (N=86), Republicans (N=58), and Independents (N=56). There was an equal split between males (49.3%) and females (50.2%) that completed the survey with one participant indicating other. Next, we discuss some of the findings from the initial phase of the study.

IV. Discussion

Table I contains the results from the survey. Not surprisingly, individuals responded consistent with their political affiliation. For example, Meme A is a pro gun rights meme. Republicans rated this lower on propaganda than either Democrats or Independents, with the latter rating it between the two others. Republicans also rated it higher with respect to its truth value.

In contrast, Meme B is for gun control. Interestingly, Independents rated this the highest for propaganda, even more so than Republicans. Republicans rated this as lower on truth value than either Democrats or Independents.

Meme G is generally considered politically neutral as it suggests that there is no discernible difference between the political parties. Nonetheless, we still see some notable differences in how it is rated based on political affiliation. Perhaps Democrats believe it is targeting their party to some extent as they rate it higher on propaganda value and lower on truth value than either Republicans or Independents.

Overall, the results are not overly surprising. Additional analysis involving political ideology would also be interesting as it may tease out some differences we may see with Independents. While the results are not surprising, they do help validate the approach we will be taking in the next phase of this research.

V. Looking Ahead

In the first phase of this research, we wanted to determine appropriate baseline numbers for the memes that were employed in this study. This will allow us to control for political affiliation and other factors in the next phase. It also provides us with an opportunity to ensure we assess a broad spectrum of political memes in the next phase of the study.

For the final phase of this study, we will be looking at how different type of people interact with memes and what they ultimately think about the message they are trying to convey. This will include an examination of how this varies based on political affiliation, political ideology, gender, age, personality [2], [20], [21], and trait affect [47], [48].

Several pertinent questions will be addressed through this research, including whether some personality types or those with a certain type of trait affect are more prone to spreading misinformation than others as prior research suggests differences based on these factors in the context of social media, security, and privacy [10], [12]. Misinformation continues to be a significant problem and strikes at the very foundation of cybersecurity through its compromise of information integrity.

In addition to understanding more fully how misinformation is spread, the results will lend themselves to various data analytic tools and techniques, such as machine learning. As evidence suggests that Russia continues to take aim at elections in the United States and elsewhere [23], it is more important now than ever that proactive measures are taken to address the threats to information integrity.

Meme	Responders	Rate the politics of this meme	This meme can be classified as....				Is the meme...			
			Propaganda	Hoax	Satire	Truth	Funny	Trying to be Funny	True	Underlying message true
A	Democrats	5.09	3.08	2.21	2.93	2.80	2.87	3.44	2.93	3.05
A	Republicans	5.52	2.19	2.09	2.41	3.88	2.79	2.95	3.84	3.98
A	Independents	5.38	2.64	1.95	3.09	3.18	2.82	3.29	3.13	3.29
B	Democrats	2.83	2.64	2.02	2.22	3.97	2.27	2.63	3.72	3.84
B	Republicans	3.88	2.83	2.54	2.38	3.14	2.14	2.34	3.16	3.48
B	Independents	2.25	3.21	2.11	2.07	3.36	1.96	2.37	3.18	3.16
C	Democrats	5.44	3.65	2.79	2.56	2.20	2.00	2.79	1.99	2.08
C	Republicans	4.86	2.76	2.29	2.66	3.53	2.41	2.58	3.64	3.62
C	Independents	5.57	3.52	2.62	2.34	2.02	1.77	2.52	1.80	1.98
D	Democrats	2.69	2.44	2.02	1.81	3.65	1.90	2.00	3.51	3.77
D	Republicans	4.07	2.48	2.49	2.14	3.69	2.16	2.19	3.69	3.72
D	Independents	2.52	2.86	2.29	1.75	3.23	1.86	1.88	3.18	3.25
E	Democrats	5.69	4.03	3.10	2.10	1.87	1.86	2.26	1.69	1.79
E	Republicans	4.60	2.86	2.80	2.28	2.98	2.14	2.19	3.19	3.33
E	Independents	5.59	3.95	2.89	2.15	2.07	1.66	2.11	2.09	2.04
F	Democrats	2.78	2.70	2.02	2.88	3.48	2.58	3.01	3.38	3.51
F	Republicans	4.38	3.12	2.52	2.12	3.19	2.26	2.60	3.41	3.45
F	Independents	2.68	3.16	1.89	2.80	2.86	2.43	2.82	2.89	3.00
G	Democrats	4.15	3.13	2.16	3.44	2.81	3.15	4.09	2.85	2.92
G	Republicans	4.97	2.56	2.29	3.53	3.07	3.32	3.52	3.28	3.34
G	Independents	4.14	2.75	2.18	3.64	3.21	3.43	4.02	3.43	3.57
H	Democrats	3.83	2.56	2.10	3.79	2.87	3.15	4.13	3.05	3.13
H	Republicans	4.83	2.52	2.40	3.61	3.03	3.45	3.71	3.24	3.29
H	Independents	4.02	2.50	1.96	3.95	3.00	3.55	4.09	3.23	3.29
I	Democrats	2.71	2.16	1.81	2.26	4.07	2.46	2.81	3.97	4.09
I	Republicans	3.41	2.91	2.93	2.40	2.91	2.30	2.64	3.07	3.14
I	Independents	2.45	2.43	1.80	2.25	3.87	2.54	3.13	3.68	3.87
J	Democrats	5.53	2.90	2.20	3.62	2.19	2.97	4.09	2.17	2.29
J	Republicans	5.54	2.40	2.17	3.26	2.98	3.59	3.60	3.38	3.52
J	Independents	5.57	2.82	1.93	3.96	2.29	3.13	3.95	2.52	2.73
K	Democrats	2.49	2.51	2.08	3.65	3.31	3.59	4.07	3.51	3.77
K	Republicans	4.33	2.91	2.45	2.96	3.09	2.66	3.28	3.22	3.41
K	Independents	2.64	2.89	1.95	3.70	2.79	3.21	3.93	2.95	3.07
L	Democrats	5.05	3.53	2.34	2.93	2.52	2.57	3.81	2.52	2.62
L	Republicans	4.72	2.95	2.55	2.81	3.42	2.81	3.33	3.66	3.69
L	Independents	4.79	3.57	2.39	3.05	2.64	2.63	3.68	2.80	2.77

TABLE I
MEANS OF THE RESPONSES FROM THE LIKERT QUESTIONS ASKED IN THE FIRST SURVEY.

VI. CONCLUSION

This study takes aim at better understanding the antecedents of the spread of misinformation. We do this through the lens of it being an attack on cybersecurity through its compromise on information integrity. Through the development of a classification and rating scheme via the Delphi technique, and subsequent data collection and validation, we are able to then take the next step and assess how various factors relate to the spread of misinformation.

REFERENCES

[1] Christopher P. Barlett. Anonymously hurting others online: The effect of anonymity on cyberbullying frequency. *Psychology of Popular Media Culture*, 4(2):70–79, 2015.

[2] Vernica Benet-Martnez and Oliver P. John. Los cinco grandes across cultures and ethnic groups: Multitrait-multimethod analyses of the big five in spanish and english. *Journal of personality and social psychology*, 75(3):729, 1998.

[3] W Lance Bennett and Steven Livingston. The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, 33(2):122–139, April 2018.

[4] Jonah Berger. Arousal Increases Social Transmission of Information. *Psychological Science*, 22(7):891–893, July 2011.

[5] Petter Bae Brandtzaeg, Asbjrn Flstad, and Mara ngeles Chaparro Domnguez. How Journalists and Social Media Users Perceive Online Fact-Checking and Verification Services. *Journalism Practice*, 12(9):1109–1129, October 2018.

[6] Walbert Castillo. 'Illini White Student Union challenges 'Black Lives Matter', November 2015.

[7] Murphy Choy and Mark Chong. Seeing Through Misinformation: A Framework for Identifying Fake Online News. *ARXIV*, page 14, 2018.

[8] Ben Collins. After Mueller report, Twitter bots pushed Russiagate hoax narrative, April 2019.

[9] C Duffield. The delphi technique. *The Australian journal of advanced nursing: a quarterly publication of the Royal Australian Nursing Federation*, 6(2), 1988.

[10] Marc Dupuis and Robert Crossler. The compromise of one's personal information: Trait affect as an antecedent in explaining the behavior of individuals. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*. IEEE, 2019.

[11] Marc Dupuis, Barbara Endicott-Popovsky, and Robert Crossler. An analysis of the use of amazon's mechanical turk for survey research in the cloud. In *International Conference on Cloud Security Management*, Oct 2013.

[12] Marc Dupuis and Samreen Khadeer. Curiosity killed the organization: A psychological comparison between malicious and non-malicious insiders and the insider threat. In *Proceedings of the 5th Annual Conference on Research in Information Technology*, page 3540. ACM Press, 2016.

[13] Marc Dupuis, Samreen Khadeer, and Joyce Huang. "i got the job!": An exploratory study examining the psychological factors related to status updates on facebook. *Computers in Human Behavior*, 73:132140, 2017.

[14] Kiran Garimella, Gianmarco De Francisci Morales, Aristides Gionis, and Michael Mathioudakis. Political Discourse on Social Media: Echo Chambers, Gatekeepers, and the Price of Bipartisanship. In *Proceedings of the 2018 World Wide Web Conference*, WWW '18, pages 913–922, Republic and Canton of Geneva, Switzerland, 2018. International World Wide Web Conferences Steering Committee. event-place: Lyon, France.

[15] Rosanna E. Guadagno, Daniel M. Rempala, Shannon Murphy, and Bradley M. Okdie. What makes a video go viral? An analysis of emotional contagion and Internet memes. *Computers in Human Behavior*, 29(6):2312–2319, November 2013.

[16] Douglas Haddow. Meme warfare: how the power of mass replication has poisoned the US election. *The Guardian*, November 2016.

[17] A. Hasell and Brian E. Weeks. Partisan Provocation: The Role of Partisan News Use and Emotional Responses in Political Information Sharing in Social Media. *Human Communication Research*, 42(4):641–661, October 2016.

[18] Felicity Hasson, Sinead Keeney, and Hugh McKenna. Research guidelines for the delphi survey technique. *Journal of Advanced Nursing*, 32(4):10081015, 2000.

[19] Heidi Huntington. MENACING MEMES? AFFECT AND EFFECTS OF POLITICAL INTERNET MEMES. *AoIR Selected Papers of Internet Research*, 5(0), 2015.

[20] Oliver P. John, Eileen M. Donahue, and Robert L. Kentle. The big five inventoryversions 4a and 54. *Berkeley: University of California, Berkeley, Institute of Personality and Social Research*, 1991.

[21] Oliver P. John, Laura P. Naumann, and Christopher J. Soto. Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research*, 3:114158, 2008.

[22] Ofra Klein. Manipulative Memes : How Internet Memes Can Distort the Truth Connected Life Conference, June 2018.

[23] Natasha Korecki. Sustained and ongoing disinformation assault targets Dem presidential candidates, February 2019.

[24] Nicolle Lamerichs, Dennis Nguyen, Mari Carmen Puerta Melguizo, Radmila Radojevic, and Anna Lange-Bhmer. Elite male bodies: The circulation of alt-Right memes and the framing of politicians on Social Media. *Participations*, 15(1):27, 2018.

[25] Eden Litt and Eszter Hargittai. The Imagined Audience on Social Network Sites. *Social Media + Society*, 2(1):205630511663348, January 2016.

[26] Jane Lytvynenko and Craig Silverman. Here Are The Hoaxes And Misinformation About The Notre Dame Fire, April 2019.

[27] Xiao Ma, Jeff Hancock, and Mor Naaman. Anonymity, Intimacy and Self-Disclosure in Social Media. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 3857–3869, New York, NY, USA, 2016. ACM. event-place: San Jose, California, USA.

[28] Alice Marwick. Why Do People Share Fake News? A Sociotechnical Model of Media Effects. *2 GEO. L. TECH. REV. 474*, July 2018.

[29] Alice Marwick and Rebecca Lewis. Media Manipulation and Disinformation Online. *Data and Society Research Institute*, page 106, 2017.

[30] Robert Mason and Marc Dupuis. Cultural values, information sources, and perceptions of security. In *iConference 2014 Proceedings*, Mar 2014.

[31] Dar Meshi, Diana I. Tamir, and Hauke R. Heekeren. The Emerging Neuroscience of Social Media. *Trends in Cognitive Sciences*, 19(12):771–782, December 2015.

[32] Panagiotis T. Metaxas and Eni Mustafaraj. Social Media and the Elections. *Science*, Vol. 338:472–473, October 2012.

[33] Sanghee Oh and Sue Yeon Syn. Motivations for sharing information and social support in social media: A comparative analysis of Facebook, Twitter, Delicious, YouTube, and Flickr. *Journal of the Association for Information Science and Technology*, 66(10):2045–2060, 2015.

[34] Ray Oshikawa, Jing Qian, and William Yang Wang. A Survey on Natural Language Processing for Fake News Detection. *arXiv:1811.00770 [cs]*, November 2018. arXiv: 1811.00770.

[35] Gordon Pennycook, Tyrone Cannon, and David G. Rand. Prior Exposure Increases Perceived Accuracy of Fake News. SSRN Scholarly Paper ID 2958246, Social Science Research Network, Rochester, NY, May 2018.

[36] Catherine Powell. The delphi technique: Myths and realities. *Journal of Advanced Nursing*, 41(4):376382, 2003.

[37] Daniel Preoiuc-Pietro, Ye Liu, Daniel Hopkins, and Lyle Ungar. Beyond Binary Labels: Political Ideology Prediction of Twitter Users. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 729–740, Vancouver, Canada, 2017. Association for Computational Linguistics.

[38] Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova, and Yejin Choi. Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-Checking. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2931–2937, Copenhagen, Denmark, 2017. Association for Computational Linguistics.

[39] David O. Sears. College sophomores in the laboratory: Influences of a narrow data base on social psychologys view of human nature. *Journal of Personality and Social Psychology*, 51(3):515, 1986.

[40] Jens Seiffert-Brockmann, Trevor Diehl, and Leonhard Dobusch. Memes as games: The evolution of a digital discourse online. *New Media & Society*, 20(8):2862–2879, August 2018.

[41] Limor Shifman. The Cultural Logic of Photo-Based Meme Genres. *Journal of Visual Culture*, 13(3):340–358, December 2014.

[42] Jieun Shin and Kjerstin Thorson. Partisan Selective Sharing: The Biased Diffusion of Fact-Checking Messages on Social Media. *Journal of Communication*, 67(2):233–255, April 2017.

[43] Craig Silverman. This Analysis Shows How Viral Fake Election News Stories Outperformed Real News On Facebook.

[44] Zachary R. Steelman, Bryan I. Hammer, and Moez Limayem. Data collection in the digital age: Innovative alternatives to student samples. *MIS Quarterly*, 38(2):355378, 2014.

[45] Cass R. Sunstein. *#Republic: Divided Democracy in the Age of Social Media*. Princeton University Press, April 2018. Google-Books-ID: nVBLDwAAQBAJ.

[46] Sebastian Tschiatschek, Adish Singla, Manuel Gomez Rodriguez, Arpit Merchant, and Andreas Krause. Fake News Detection in Social Networks via Crowd Signals. In *Companion Proceedings of the The Web Conference 2018*, WWW '18, pages 517–524, Republic and Canton of Geneva, Switzerland, 2018. International World Wide Web Conferences Steering Committee. event-place: Lyon, France.

[47] David Watson and Lee Anna Clark. The panas-x: Manual for the positive and negative affect schedule - expanded form. 1994.

[48] David Watson, Lee Anna Clark, and Auke Tellegen. Development and validation of brief measures of positive and negative affect: The panas scales. *Journal of Personality and Social Psychology*, 54(6):10631070, Jun 1988.

[49] Linzi E A Williamson, Sarah L Sangster, and Karen L Lawson. HEY GIRL: THE EFFECT OF RYAN GOSLING FEMINIST MEMES ON FEMINIST IDENTIFICATION AND ENDORSEMENT OF FEMINIST BELIEFS. *Researchgate.net*, page 1, 2014.

[50] L. Y.C. Wong and Jacquelyn Burkell. Motivations for Sharing News on Social Media. In *Proceedings of the 8th International Conference on Social Media & Society*, #SMSociety17, pages 57:1–57:5, New York, NY, USA, 2017. ACM. event-place: Toronto, ON, Canada.

[51] Samuel C. Woolley and Douglas R. Guilbeault. Computational Propaganda in the United States of America: Manufacturing Consensus Online. *Project on Computational Propaganda*, 2017.

[52] Kaiping Zhang and Ren F. Kizilcec. Anonymity in Social Media: Effects of Content Controversiality and Social Endorsement on Sharing Behavior. In *Eighth International AAAI Conference on Weblogs and Social Media*, May 2014.