JOURNAL OF DIGITAL SOCIAL RESEARCH

More human than human
vol 4, no 1, 2022

JOURNAL OF DIGITAL SOCIAL RESEARCH

VOL. 4 : NO. 1 : 2022

JOURNAL METADATA

ISSN
2003–1998

DOI
http://doi.org/10.33621/jdsr

WEB
www.jdsr.io

ABOUT JDSR

JDSR is a interdisciplinary, online, open-access journal, focusing on the interaction between digital technologies and society. JDSR is published by DIGSUM, the Centre for Digital Social Research at Umeå University, Sweden.

CONTACT US

E-MAIL
editor@jdsr.se

JOURNAL OF DIGITAL SOCIAL RESEARCH

VOL. 4 : NO. 1 : 2022

JOURNAL OF DIGITAL SOCIAL RESEARCH

WWW.JDSR.IO ISSN: 2003-1998

VOL. 4, NO. 1, 2022, 1–22

ETHICAL USE OF INFORMANT INTERNET DATA: SCHOLARY CONCERNS AND CONFLICTS

Michael Nycyk[a]

ABSTRACT

This article explores the scholarly concerns and conflicts debated by authors in the field of the use of informant internet data in research. The importance lies in informant protection and how to minimize harm to them, a long-standing cornerstone of research practice. It is also a public domain issue as increased calls for data privacy grew because of reported data breach scandals. Although not a new problem, academic researchers and university ethics boards struggle with concerns over data use and are in conflict about managing the problem. This article uses thematic analysis to identify, analyze and interpret patterns of concerns and conflicts over internet data use. Data was obtained from academic publications on these issues. Three themes from this data are discussed with examples demonstrating the types of, and complexity of, scholarly concerns and conflicts. These themes are: the problems of informant data use risks, gaining mass informed consent and the challenges ethics boards face, especially conflicts with researchers over internet data use on projects. This article contributes insights into a widely, and continuously, debated area which is constantly evolving as privacy laws and public awareness place pressure on researchers and ethics boards to address protecting informant public internet data.

Keywords: ethics boards, informed consent, informant protection, internet data ethics, researcher concerns and conflicts, thematic analysis

[a] Queensland University of Technology, Brisbane, Australia.

1 INTRODUCTION

In 2014, Adam Kramer and his research colleagues' emotional contagion study was published in an open access journal (Kramer, Guillory and Hancock 2014). Their study used Facebook user's personal data without consent. Scholars, and the public, were vocal in condemning the study. Flick (2016) argued that it took advantage of peoples' autonomy to decide how their internet data should be used. Hallinan, Brubaker and Fiesler (2020) stated that the study violated social media users' expectations of data protection and had negative ramifications on the public's perceptions of the use of their data for intrusive research. User data is especially regarded by commercial organizations as being for monetization, not taking into account that it is the users who produce and create content and give their data in good faith to social media platforms and the internet (Johanssen, 2021). Koenig (2018) likens the harvesting and use of user internet data as a form of serfdom, where data is given to an overlord in exchange for using that part of the internet or a social media platform.

Social media and other large datasets, also called big data, have the potential to contribute to understanding and solving global research problems. The logistics of obtaining informed consent from thousands of people to use their data is often not viable. People are concerned that they do not know how their data will be used or if it will be shared with others (Hand 2018). Issues such as the Cambridge Analytica Data Breach Scandal, and Kramer, Guillory and Hancock's (2014) misuse of Facebook data have alerted the public to this issue, but also challenged researchers and universities on how to manage this significant problem. Internet and data privacy scholars argue more critical questions about who gets access to what data, how the analysis is undertaken and to what it will be used for, need further crucial scrutiny and continuous scholarly debate (Boyd and Crawford 2012).

This article contributes insights into the issue of researcher use of informant's internet data and the concerns and conflicts this causes. It is not a new problem as scholars have used peoples' online data for research since the public internet became available. Scholarly reflection on this problem, and how to solve it, has emerged as a large field with a growing body of literature. It is not necessarily a difficult to control or unmanageable dilemma. Large public internet datasets offer insights not seen before in academic history. Identifying and managing concerns and conflicts is an important problem to examine. Data use ethical issues are growing as privacy laws and new technological developments arise, further challenging how to balance using informant data without identifying or harming informants.

This article uses thematic analysis methods (Clarke, Braun and Hayfield 2015; Braun and Clarke 2006) to identify and analyze themes that represent the concerns and conflicts of informant internet data use. In identifying these themes, it gives an appreciation of the problem of using large public datasets without informant's consent that concerns researchers. Taking an exploratory and inductive research approach to this problem, this article's contribution is to examine and

explain the problems of public data use that arise, based on examining scholarly writing from researchers. The article closes with a discussion of, and implications of, the themes' findings and also give a brief evaluation of using thematic analysis in this study.

2 CONCERNS AND CONFLICTS OVER PUBLIC INTERNET DATA USE FOR RESEARCH

An important argument illustrating the concerns and conflicts over informant public internet data use is that the researcher and university must be bound by traditional ethical research practices. The protection of informants and getting informed consent to use their data is a primary concern of ethical research practice. The arrival of the internet encouraged a 'the data is public' type attitude among researchers, as they used text and image data from the internet without asking those that produced them. Metcalf and Crawford (2016) claim an emerging ethics divide has arisen with discontinuities between researcher practices and ethical practices. It is proven that business and government use peoples' personal data for analysis. Academia, by contrast, has been challenged by a number of ethical issues with using large datasets, including privacy, identification and surveillance by outsiders' issues.

How to regulate and adhere to ethical research practices using informant internet data is increasingly difficult. Using a metaphor, data to researchers is oxygen as a growing share of our lives are dedicated to its release (Svensson and Guillen 2020; Jarzombek, 2016). Koenig (2018) believes individual data is one's own property that one should have total control over. Practically, consent cannot be obtained from every person who creates and posts data (Sinclair, 2017). Consent is often not obtained because large dataset creates a valuable digital image of human reality wanted by researchers for their research. As Zwitter (2014) argues, even in a sea of unidentified data, this lack of consent has a possible potential to cause harm through the potential identification of the creator of the data.

It is argued that internet data is free for taking because although internet users expect privacy for their public data, they freely and frequently post it in the public internet domain (Bay 2018; Fiesler and Nicholas 2018). This is factual, but because of mass-media reporting of data misuse, researchers and their universities are being forced to rethink established ethical principles of informed consent, privacy and anonymity (Sugiura, Wiles and Pope 2017). Despite difficulty in practicing this, views have been expressed by researchers that the public should be the arbiters of deciding if their data should be used in research without consent. Ainscough et al. (2018, p. 505) and McKee (2013) argue that deciding what is public and private data, even if that data is posted publicly online, can only really be defined by the subjects themselves "as they are the authors of their own words."

An ethical use of data example is the use of microblogging platform Twitter's data, which is problematic despite Twitter allowing researchers to use its members' data without informed consent. This problem leads to confusion and ambiguity

about what data and details of that data should or should not be used, as Ahmed, Bath and Demartini (2017) state:

> A justification often provided by those working in an academic context with Twitter data with regard to the ethical and legal implications of using data without informed consent is that the reuse of data is permitted by Twitter's Terms and service as well as within the privacy policy. However, it is important to note that the act of scraping tweets or downloading tweets from Twitter's Advance Search will contravene Twitter's Terms and Conditions, therefore voiding any protection these policies are likely to offer (Ahmed, Bath and Demartini 2017, p. 7).

However, although studies using informant public data continue unabated, lessons have not always been learned from over a decade of data ethics research (Markham and Buchanan, 2012) as practices such as identified by Ahmed, Bath and Demartini (2017) continue.

The fast-paced growth of using online datasets is also not keeping pace with ethical boundaries. Although studies using informant public data continue unabated, lessons have not always been learned from decades of data ethics research. For example, problems with consent were previously recognized in studies of virtual communities, newsgroups, Internet Relay Chat and Usenet. Researchers lurked in these communities, sometimes known, but often anonymously, not informing people they were collecting data. If the researcher became known or informants later found out their posts were collected, the researcher was seen as an intruder potentially damaging the confidential community relationships built up over time (Eysenbach and Till 2001).

Early studies using internet data obtained from newsgroups and other sites would use data without informed consent. Ethical concerns of this practice began to be debated. Potential participant harm was no longer localized as the internet grew and data could be obtained from anyone, anywhere (Ess 2006). Researchers would employ pseudonyms to protect informants. The problem can be that using text examples verbatim, such as a tweet, can possibly be traced back to the informant. This may not be problematic to the person, but harm is possible, hence concerns about ethical data use became consistently expressed by researchers.

Informant protection from harm is still a central concern in all research. Data collected from the internet may produce a sense of detachment for the researcher. This can foster an attitude that using any internet data without informed consent is acceptable. That is not to imply researchers do not care about their informants as they are obliged, legally and morally, to protect the informants they study. To illustrate this, a study by Samuel, Derrick and van Leeuwen (2019) identified attitudes towards the use of social media data in research concluding:

> This article argues that this governance of ethical behaviour by individual researchers perpetuates a negative cycle of academic practice that is dependent on subjective judgements by researchers themselves, rather than governed by more

> formalised academic institutions such as the research ethics committee and funding council guidelines (Samuel, Derrick and van Leeuwen 2019, p 317).

Their further finding was that in their sample of 324 United Kingdom research articles, 234 articles did not mention ethical approval or discussions of ethical implications of using social media data (Samuel, Derrick and van Leeuwen 2019). The suggestion from this is that there is a moral belief that more transparency in conducting studies using informant data should be a mandatory requirement.

Managing and solving these conflicts and concerns have been made in academia, but also in some country's legal systems to protect informants' internet data. Samuel and Buchanan (2020, p. 8) proposed a framework of ethical issues that needed to be addressed before doing an internet-mediated research study. Their call for best practice procedures involves the researcher having high ethical standards and respect for individuals and communities. Governments have also responded to public concerns for data privacy. For example, The General Data Protection Regulation (EU) (GDPR), part of European Law, clearly sets out what personal data is. Although applicable to corporations and governments, the regulations states researchers must be careful with an individual's identification:

> If you cannot directly identify an individual from that information, then you need to consider whether the individual is still identifiable. You should take into account the information you are processing together with all the means reasonably likely to be used by either you or any other person to identify that individual. (Information Commissioner's Office 2018, p. 10).

This illustrates that some countries are taking seriously the protection of peoples' public and private data, urging researchers to take seriously the identification of people who researchers take data from without informed consent.

3 CONFLICTS BETWEEN RESEARCHERS AND ETHICS BOARDS

Conflicting beliefs about informed consent and data use, and how to protect informants' data, still occur as standards set by external laws and organizations are often still unclear. The literature suggests ethics boards are still lacking in understanding public data use. However, they and other organizations supporting internet researchers' projects have been proactive in developing ethical conduct policies. In the literature ethics boards are blamed for their lack of clarity on this issue. Conflicts between internet researchers and ethics boards, who are either over cautious or unclear on how to approve internet data collection without harming informants, do occur.

In one study, Stevens, O'Donnell and Williams (2015) wanted to find out about personal experiences of those living with chronic illness. The data in their ethics application was disclosed as public Facebook posts. Commenting on Stevens, O'Donnell and Williams project, Ainscough et al. (2018) stated that the initial

rejection of the Stevens led project, which was later overturned, was because of the panel's uncertainty whether informed consent from informants was required. This was an example of the uncertainty that exists in approving online data research projects, what is public and private, and if informed consent is needed or not (Ainscough et al. 2018).

Several authors have commented on the contested and hostile relationships between researchers and ethics boards. Examples include: deceiving ethics boards to rush through approvals, collecting data without approval from the board and not reporting changes for fear of having restrictions placed on their research (Vitak et al. 2017; Feeley 2007; Keith-Spiegel, Koocher and Tabachnick 2006). Internet researchers may succumb to the temptation, or necessity in large datasets, to collect data without informed consent. However, ethical boards in approving projects must understand the difficult logistics of obtaining consent and consider the level of risk to participants. Researchers and ethics boards need to be empathetic towards each other and work co-operatively.

The Association of Internet Researchers (AoIR) is an academic association that has been proactive in attempting to develop worldwide guidelines in internet data use ethics. They claim that ethics boards are now becoming overly cautious approving research projects after public data use scandals were reported. Their strategy to remedy this is to develop guidelines to mediate ethical relationships between boards and researchers. franzke et al. (2020) state their aim to address this:

> This points to one of the founding justifications for and central subsequent uses of the AoIR IRE guidelines namely, the importance of being able to help inform ethical review boards of the distinctive characteristics of internet research, as well as to provide researchers with resources, beginning with the guidelines themselves, to help them in the processes of negotiating the process of seeking approval for their research with such boards (franzke et al. 2020, p. 13).

Managing internet research data use ethical questions requires mutual agreement and constant communication between researchers, research students' academic advisors and university ethics boards. One question identified in the literature that hinders this process is if social media data is counted as a human subject research or if it is merely text research (Samuel and Buchanan 2020). The issues though of the misuse of such data, combined with peoples' concerns over this, move this issue into the realm of not being detached text only data, but being part of a person's data ownership where informed consent, however impractical with large datasets, should be a part of any research project.

Academic human ethics committees are struggling with understanding potential ethical issues. Researchers may become frustrated at justifying why such research warrants taking peoples' social media data for analysis. To investigate this further, a thematic analysis of the scholarly concerns and conflicts in this field is offered that contributes to this growing, ongoing debate. The question guiding the

analysis is, what are the priority scholarly concerns and conflicts that the ethical, or not, use of internet data are of concern to researchers and university ethics boards?

4 METHODS

An inductive thematic analysis method was used to examine these scholarly concerns and conflicts. An overarching goal is to identify and search for themes appearing as patterns in data (Braun and Clarke 2006). Thematic analysis aims to identify or examine the underlying ideas, assumptions, conceptualizations and ideologies that are theorized as shaping or informing the semantic content of the data (Braun and Clarke 2006). Being inductive and exploratory, no existing theory is used in the data analysis. The method, though generically theoretically flexible, is not atheoretical (Braun and Clarke 2019; Clarke and Braun 2018). This means an inductive approach can generate theories, or explanations, that give insights as to what is going on in an examined problem.

Data collecting was done through searching for peer-reviewed academic articles and journal editorials using academic research databases and Google Scholar. Search term examples included: 'social media data', 'privacy', 'big data', 'participant risk' and 'ethics boards'. A total of 50 publications, between years 2001 to 2019, were chosen, 4 of which were refereed conference papers (see Appendix for publications chosen and their classification). What is distinguished is, the total of 50 publications is a small dataset. Inductive research can have reduced datasets for smaller projects such as this (Clarke, Braun and Hayfield 2015). The text lifted from the publications that was analyzed was though substantive with many sentences coded for possible theme development. It was assessed that the data was reliable because the articles were all peer-reviewed, increasing the validity and truthfulness of the authors' text taken from the publications.

The analysis was conducted according to Clarke, Braun and Hayfield's (2015) methods as:

1. Data familiarization requiring constant reading of the text
2. Generating an initial set of codes across the dataset
3. Searching for themes and sub themes from the codes
4. Reviewing themes including merging them for accuracy
5. Final theme naming and definitions of themes produced

Although appearing formulaic, thematic analysis is rigorous and time-consuming. Themes are important to search for and justify because they describe a phenomenon (Ferday and Muir-Cochrane 2006; Daly, Kellehear and Gliksman 1997). During the process, recognizing and encoding a moment in the text was done as this describes types of observations that can interpret the phenomenon (Ferday and Muir-Cochrane 2006; Boyatzis 1998). Another advantage is, despite the straightforward research problem presented in this article, thematic analysis allows

for unexpected results to be found that may provide insights previously unknown or unclear (Nowell et al. 2017; Côté and Turgeon, 2005). To assist with validity, a constant reviewing of codes and themes was done. The time to review themes also assisted in ensuring the themes represented explanations of scholarly concerns and conflicts about internet data.

5 THEMES AND ANALYSIS OF SCHOLARLY CONCERNS AND CONFLICTS

The themes from the analysis demonstrated the conflicts that have been present since the early period of gathering data off the internet for research. Being a public sphere, people have been concerned if their internet data is made public or their posts are identified and linked to them. The dataset provided instances of growing concerns and conflicts as internet technologies evolved, especially with the arrival of social media. The concerns from the public are represented by commentators like Koenig (2018) who still believe internet data is the property of the person who creates it. The themes show this is often not respected by researchers who may hold the attitude of it is public data there for the taking. Yet a constant theme, being far larger than this study's sample, suggests by force or willingness, researchers want to do no harm to their informants who provide data.

Three themes were created from the inductive study showing the types of scholarly concerns and conflicts. Risk, consent and practices, and the researcher's and ethics board's perception of how to manage them, were consistent patterns found in the text and make up the three themes. It needs to be noted that although not a factor in the analysis, the texts in the academic publications came from many countries. Even though they had different laws and policies, the consistency of this problem transcends geographic location.

Table 1 shows the themes that thematic analysis study yielded to explain researcher concerns and conflicts:

Table 1. Themes of scholarly concerns and conflicts in using informant internet data defined

Theme Name	Definition
Informant's Data Use Risks	Captures and identifies the individual and structural risks that are a consequence of using internet and social media data for research.
Research Consent and Practice	Gaining consent from informants in some way despite the massive datasets and maximizing informant protection.
Ethics Boards' Practices	Challenges that boards that approve researcher studies must contend with but may also clash with researchers over.

The three themes showed three sets of interest groups, researchers, informants and ethics boards, often in conflict with each other on how to manage informant internet data protection. Informant protection and harm minimization remains a pillar of research practice, but clearly using and managing large internet datasets that may have identifying factors is still not resolved. The study's themes are now discussed with example text to support them.

6 SCHOLARLY CONCERNS AND CONFLICTS THEMES

To support the themes, evidence illustrating each of them is given with text examples. The quotes displayed are from journal and conference articles that exist in some form in a public or semi-private domain (such as a journal issue behind a pay wall). Therefore, these quotes attributed to the author/s do not need consent to be used, but are referenced for academic integrity.

6.1 Informant's Data Use Risks

Authors were overall very concerned with informant risks and potential negative consequences that may occur from identification of the informant. This was especially evident with authors advising caution when using informant data extracts, for example tweets, in scholarly and other publications. A second risk concern was the sharing of internet datasets with other researchers or organizations, as well as concerns about data storage where problems such as cyber criminals hacking databases might occur.

Buchanan (2017) illustrates this problem seen across the dataset and also in the theme of data use risks. Ethical dilemmas that pose significant risks to informant identification are appearing regularly as research datasets grow in volume:

> Ethics and methods are interdependent, and the rise of mass data mining across social media and the Internet has presented ethical dilemmas surrounding privacy, rights and autonomy, and such social justice issues as discrimination (Buchanan 2017, p. 2).

Ethics boards have a reputational and moral responsibility to ensure informant research risks are reduced and managed in line with high university standards. Consistent standards to protect informant internet data still seem to be in their infancy. In the articles' data, calls for standards that can be flexible to change in time as privacy laws, new global problems requiring examining of large datasets and technological advances occur was considered important. Kumar and Nanda (2019) illustrate this common problem:

> Development of ethical standards for the social media is not just a theoretical exercise, but a carefully planned strategy assuming responsibility and active participation of both the individuals and organizations (Kumar and Nanda 2019, p. 67).

As a generalization, it was not clear in the data if ethics boards were solely to blame for not keeping up with the use of large internet datasets and potential issues for informants these may cause. It is inaccurate to conclude that many university ethics boards are not addressing and managing this issue correctly. What was identified is that increased informant data risks can have potential seriously consequences to informant privacy if a person's name or other personal details are shown in a dataset or in a publicly available publication. It is this potential, however unlikely, that must be addressed when setting out boundaries for researchers in their projects to protect all informants, especially where no informed consent is given.

Jouhki et al. (2016) also stated a similar concern seen across the sampled literature. Even though researchers may mask informant details, people are concerned at possible identification, especially with sophisticated internet tools able to find even small items of information about someone:

> Especially when informed consent cannot be obtained in human-subject research, the benefits of the study should outweigh the harm of any invasion of privacy (Jouhki et al. 2016, p. 77).

Therefore, this theme demonstrates that protecting informant's internet data is crucial to reducing risk. Despite researcher's eagerness to use large internet datasets, and the acceptance that informed consent is almost impossible to obtain, care needs to be taken to mask personal details. The researcher and ethics boards must agree to all safeguards to protect informants and reduce potential harm. But researchers must not be dismissive that just because peoples' details can be masked in large datasets that it is acceptable to take and use the public data because it is there.

6.2 Research Consent and Practice

This theme correlates with the other two themes because it involves a cornerstone of any research; getting an informant's consent to use their data. An informant should always feel free to withdraw that data at will. Again, with large datasets that is often not possible. Researchers often decide to use the data without consent, but also use it for other reasons, such as storing and sharing that data in the future, showing the difficulty of always practicing ethical consent procedures, as Nunan and Yenicioglu (2013) illustrate this conflict:

> On the one hand they promise that ownership of content remains in the hands of the individuals who posted it. On the other, they grant themselves permission to use the content for a wide range of purposes, purposes which are largely unspecified (Nunan and Yenicioglu 2013, p. 797).

Another part of this theme was deciding what details and data examples to publish in academic and other research publications coming out of a study using informant internet data. Researchers have a conflict that it was not just consent to use data, but also what to publish in terms of informant's details. Across the sample a consistent pattern was to strongly recommend researchers to think carefully about

what data examples to publish. Researchers do have a strong moral obligation to be careful with what informant details they publish. Fiesler and Proferes (2018) recommend:

> Therefore, we suggest that publication of user identity should only occur when the benefits of doing so clearly outweigh the potential harms, or with user permission (Fiesler and Proferes 2018, p. 797).

This finding did not imply though that not putting real names or other informant details in publications was unacceptable. In some studies, an informant may actually request their real name and other details be published, but the ethics board must be made aware the informant allows this to happen and makes the university exempt from any reputational or other damage.

An expected and reoccurring pattern in the data was that the authors reported how and why it was important to respect the informant's personal internet data, especially if no consent was obtained to use it. It is important to be considerate to not to do harm when using data to which there is no consent given. That part of the theme was recognized in many of the publications. Gupa (2017) states that respecting participants' data means, regardless of the size of the dataset, the participants still hold the right to not give consent for data use:

> Consent is considered to be knowledgeable when individuals know the purpose of the information collected about them and their right to give, withhold, or withdraw the consent anytime they wish (Gupa 2017, p. 7).

How this can be practiced remains a problem that researchers and ethics boards need to consistently work on. Again, the concern arises that it is unlikely informants will be contacted to have their internet data used due to the dataset's volume. Yet they still retain the right to ask to have their data, even if it is part of an anonymous large dataset, to be not used.

6.3 Ethics Boards' Practices

In the data, the authors were critical of the role of university ethics boards in the boards not understanding how serious an ethical issue it has become to protect participants' internet data. Boundaries of what constitutes unethical use of the data are unclear. Gustafson and Woodworth (2014) commented on this conflict in their study of using large amounts of patient internet medical data:

> ...differing assumptions about the social boundaries of public and private communication that occurs in virtual communities; and ... the ethical dilemmas that emerge from these assumptions (Gustafson and Woodworth 2014, p. 5).

Though these conflicts are being addressed, ethics boards are also criticized by researchers as not keeping pace with fast developments and emerging new ethical issues. Ethics boards may approve research projects without considering a website

or social media platforms' terms and condition's changes or passing of new data protection privacy laws. Kinder-Kurlanda et al. (2017) argue the following:

> Sharing legally and ethically also means to follow the changes and updates in terms of services and policies and to participate in negotiations about data sharing for the sake of reproducibility with platform providers (Kinder-Kurlanda et al. 2017, p. 11).

It can be questioned from this theme why should ethics boards be totally responsible for this? Researchers do have an obligation to disclose potential risks. Ethics boards though need to be aware of informant and study risks as using internet data grows. That is the key solution in addressing scholarly concerns and conflicts that is being addressed by ethics boards and organizations like the AoIR.

This theme also urged greater co-operation between researchers and ethics boards should be a priority. There is not a suggestion that all ethics boards are not doing this. The issue lies in keeping informants' privacy and free from harm requires much work to maintain. In the sample, researchers wrote of how crucial this is. Shilton and Sayles (2016) for example urge greater and constant co-operation between researchers and ethics boards at all stages of the research:

> ...this work suggests that ethics review boards (or alternative institutional structures) might best be positioned as consultants to research design, rather than post-hoc enforcement mechanisms (Shilton and Sayles 2016, p. 9).

Examining the three themes, the concerns and conflicts of researchers must be consistently addressed in all academic institutions worldwide, so informant internet data can be ethically used. Themes, and the examples given in this section, suggest this has been not always handled with care and responsibility by all involved. Concurrently, it is understandable though the rush to use internet data for research has not kept up with protections for informants. The question drawn from this analysis is, if these concerns and conflicts have been identified, what will occur at a global level in academic institutions to urgently remedy these issues?

7 DISCUSSION AND CONCLUSIONS

This analysis gave insights into the scholarly concerns and conflicts when informant internet data is used, the harm and risks its use brings, as well as advising that researchers and ethics boards must pay greater attention to this problem. What is clear is informant internet data will now be a major part of the research landscape. Informed consent cannot be obtained in large datasets. Yet despite this, ways of protecting informant's data, especially through cooperation and communication between researchers and ethics boards, is vital to managing informant harm minimization. Unethical conduct in using data without consent can result in researcher and university reputational harm, withholding of government research funding and journal article retractions.

The study identifies, but does not place sole responsibility on, the role of ethics boards in understanding the consequences of informant internet data use. Ethics boards may meet infrequently, but mechanisms and structures should be in place to monitor the use of internet data over the life of a research project. It also needs to be decided the level of acceptability of informant protection such as, what should appear in public and academic publications or if certain factors, such as sexuality, ethnic background, or medical conditions, warrant extra care in data collection. Laws to protect participant data or changes to a platform's terms in giving permission to use data will also constantly change. Theme three demonstrated that the urgency to remedy this issue is crucial if new laws limit researchers obtaining online data for research. That can only take place with cooperation between researchers and ethics boards.

A comment on the use of thematic analysis is warranted here to reflect on its usefulness in identifying and analyzing scholarly concerns and conflicts. The small set of 50 scholarly articles could have produced minimal results. However, the amount of text that informed the themes was large. Clarke, Braun and Hayfield (2015) though state that for a smaller project this amount of dataset can still generate codes and themes to answer a research question. What has arisen from doing thematic analysis though is that the patterns of concerns and conflicts were turned into themes that explain the problems occurring in the field of data ethics. It is feasible to suggest another researcher may yield different themes and explanations. Yet insights presented in this article did provide one way of examining what is a serious and ever-changing issue, where the protection of informants remains the important pillar that research studies rest on.

Protecting informants is as crucial as it ever was; more so now that privacy has been eroded by increasing levels of personal data published online. People will increasingly expect their data to be theirs to do as they wish with and decide how it will be used. They want to feel their data is safe. Relationships between researchers and ethics boards must not be adversarial. Although talking about informant protection from a commercial, as well as academic, position, Johanssen (2021, p. 101) is accurate in stating beneath the surface, this feeling of security is broken and users are denied mastery over their data and their destiny. This debate will continue necessitating an urgent need for cooperation in protecting informants but continue vital global and local research that the richness of detail of internet data can provide.

FUNDING STATEMENT AND ACKNOWLEDGMENTS

There was no funding received for this study. I would like to acknowledge my former university and the Department of Internet Studies, Curtin, Perth, Australia for giving me the idea to conduct this study into the issues of using public internet data.

REFERENCES

Ahmed, W., Bath, P., & Demartini, G. (2017). Chapter 4 using Twitter as a data source: An overview of ethical, legal, and methodological challenges. In Woodfield, K. (Ed.), *The ethics of online research. advances in research ethics and integrity volume 2* (pp. 79-107). Emerald Publishing Limited.

Ainscough, E., Smith, S. N., Greenwell, K., & Hoare, D. J. (2018). Findings and ethical considerations from a thematic analysis of threads within tinnitus online support groups. *American Journal of Audiology*, *27*(3), 503-512. https://doi.org/10.1044/2018_AJA-IMIA3-18-0013

Bay, M. (2018). Social media ethics: A Rawlsian approach to hypertargeting and psychometrics in political and commercial campaigns. *ACM Transactions on Social Computing*, *1*(4), 1-14. https://doi.org/10.1145/3281450

Boyatzis, R. (1998). Transforming qualitative information: Thematic analysis and code development. Sage.

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication and Society*, *15*(5), 662-679. https://doi.org/10.1080/1369118X.2012.678878

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589-597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Buchanan, E. (2017). Considering the ethics of big data research: A case of Twitter and ISIS/ISIL. *PLoS ONE*, *12*(12), 1-6. https://doi.org/10.1371/journal.pone.0187155

Clarke, V., & Braun, V. (2018). Using thematic analysis in counselling and psychotherapy research: A critical reflection. *Counselling and Psychotherapy Research*, *18*(2), 107-110. https://doi.org/10.1002/capr.12165

Clarke, V., Braun, V., & Hayfield, N. (2015). *Thematic analysis*. In J. Smith (Ed.), Qualitative psychology: A practical guide to research methods (3rd ed., pp. 222-248). Sage Publications.

Côté, L., & Turgeon, J. (2005). Appraising qualitative research articles in medicine and medical education. *Medical Teacher*, *27*(1), 71-75. https://doi.org/10.1080/01421590400016308

Daly, J., Kellehear, A., & Gliksman, M. (1997). *The public health researcher: A methodological approach*. Oxford University Press.

Ess, C. (2006). Internet research ethics. In Joinson, A., McKenna, K., Postmes, T., and Reips, U (Eds.), *Oxford handbook of psychology*. (pp. 487-503). Oxford University Press.

Eysenbach, G., & Till, J. (2001). Ethical issues in qualitative research on internet communities. *BMJ Clinical Research*, *323*(7321), 1103-1105.

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC59687/. https://doi.org/10.1136/bmj.323.7321.1103

Feeley, M. (2007). Legality, social research, and the challenge of institutional review boards. *Law and Society Review*, *41*, 757-776. https://doi.org/10.1111/j.1540-5893.2007.00322.x

Ferday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, *5*(1), 80-92. https://doi.org/10.1177/160940690600500107

Fiesler, C., & Proferes, N. (2018). "Participant perceptions" of Twitter research ethics. *Social Media + Society*, *4*(1), 1-14. https://doi.org/10.1177/2056305118763366

Flick, C. (2016). Informed consent and the Facebook emotional manipulation study. *Research Ethics*, *12*(1), 14-28. https://doi.org/10.1177/1747016115599568

Franzke, A., Bechmann, A., Zimmer, M., Ess, C., & the Association of Internet Researchers. (2020). *Internet research: Ethical guidelines 3.0*. https://aoir.org/reports/ethics3.pdf

Gupa, S. (2017). Ethical issues in designing internet-based research: Recommendations for good practice. *Journal of Research Practice*, *13*(2), 1-14. http://jrp.icaap.org/index.php/jrp/article/view/576/476

Gustafson, D., & Woodworth, C. (2014). Methodological and ethical issues in research using social media: A metamethod of Human Papillomavirus Vaccine Studies. *BMC Medical Research Methodology*, *14*(127), 1-11. https://doi.org/10.1186/1471-2288-14-127

Hallinan, B., Brubaker, J., & Fiesler, C. (2020). Unexpected expectations: Public reaction to the Facebook emotional contagion study. *New Media and Society*, *22*(6), 1076-1094. https://doi.org/10.1177/1461444819876944

Hand, D. (2018). Aspects of data ethics in a changing world: Where are we now? *Big Data*, *6*(3), 176-190. https://doi.org/176-190. 10.1089/big.2018.0083

Information Commissioner's Office. (2018). *Guide to the general data protection regulation (GDPR)*. https://ico.org.uk/media/for-organisations/guide-to-the-general-data-protection-regulation-gdpr-1-0.pdf

Jarzombek, M. (2016). *Digital Stockholm Syndrome in the post-ontological age*. University of Minnesota Press.

Johanssen, J. (2021). Data perversion: A psychoanalytic perspective on datafication. *Journal of Digital Social Research*, *3*(1), 88-105. https://www.jdsr.io/articles/2021/3/26/data-perversion-a-psychoanalytic-perspective-on-datafication. https://doi.org/10.33621/jdsr.v3il.57

Jouhki, J., Lauk, E., Penttinen, M., Sormanen, N., & Uskali, T. (2016). Facebook's emotional contagion experiment as a challenge to research ethics. *Media and Communication*, *4*(4), 75-85. https://doi.org/10.17645/mac.v4i4.579

Keith-Spiegel, P., Koocher, G., & Tabachnick, B. (2006). What scientists want from their research ethics committee. *Journal of Empirical Research on Human Research Ethics*, *1*, 67-81. https://doi.org/10.1525/jer.2006.1.1.67

Kinder-Kurlanda, K., Weller, K., Zenk-Möltgen, W., Pfeffer, J., & Morstatter, F. (2017). Archiving information from geotagged tweets to promote reproducibility and comparability in social media research. *Big Data & Society*, *4*(2), 1-14. https://doi.org/10.1177/2053951717736336

Koenig, G. (2018). Introduction: The Softian bargain. In I. Landreau, G. Peliks, N. Binctin, V. Pez-Pérard & L. Léger (Eds), *My data are mine: Why we should have ownership rights on our personal data* (10-17). Paris: GenerationLibre. https://www.generationlibre.eu/wp-content/uploads/2018/01/Rapport-Data-2018-EN-v2.pdf

Kramer, A., Guillory, J., & Hancock, J. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *PNAS*, *111*(24), 8788-8790. https://doi.org/10.1073/pnas.1320040111

Kumar, V., & Nanda, P. (2019). Social media to social media analytics: Ethical challenges. *International Journal of Technoethics (IJT)*, *10*(2), 57-70. https://doi.org/10.4018/IJT.2019070104

McKee, R. (2013). Ethical issues in using social media for health and health care research. *Health Policy*, *110*(2-3), 298-301. https://doi.org/10.1016/j.healthpol.2013.02.006

Markham, A., & Buchanan, E. (2012). Ethical decision-making and internet research: Recommendations from the AoIR Ethics Working Committee (Version 2.0). https://www.aoir.org/reports/ethics2.pdf

Metcalf, J., & Crawford, K. (2016). Where are human subjects in Big Data research? The emerging ethics divide. *Big Data & Society*, *3*(1), 1-14. https://doi.org/10.1177/2053951716650211

Nowell, L., Norris, J., White, D., & Moules, N. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*(1), 1-13. https://doi.org/10.1177/1609406917733847

Nunan, D., & Yenicioglu, V. (2013). Informed, uninformed and participative consent in social media research. *International Journal of Market Research*, *55*(6), 791-808. https://doi.org/10.2501/IJMR-2013-067

Samuel, G., & Buchanan, E. (2020). Guest editorial: Ethical issues in social media research. *Journal of Empirical Research on Human Research Ethics*, *15*(1-2), 3-11. https://doi.org/10.1177/1556264619901215

Samuel, G., Derrick, G., & van Leeuwen, T. (2019). The ethics ecosystem: Personal ethics, network governance and regulating actors governing the use of social media research data. *Minerva*, *57*, 317-343. https://doi.org/10.1007/s11024-019-09368-3

Sinclair, M. (2017). Using social media for research. *Evidence Based Midwifery*, *15*(2), 39. Available at: https://www.rcm.org.uk/media/2787/evidence-based-midwifery-june-2017.pdf

Stevens, G., O'Donnell, V., & Williams, L. (2015). Public domain or private data? Developing an ethical approach to social media research in an inter-disciplinary project. *Educational Research and Evaluation*, *21*(2), 154-167. https://doi.org/10.1080/13803611.2015.1024010

Sugiura, L., Wiles, R., & Pope, C. (2017). Ethical challenges in online research: Public/private perceptions. *Research Ethics*, *13*(3-4), 184-199. https://doi.org/10.1177/1747016116650720

Svensson, J., & Guillen, O. P. (2020) What is data and what can it be used for? Key questions in the age of burgeoning data essentialism. *Journal of Digital Social Research*, *2*(3), 65-83. https://jdsr.se/ojs/index.php/jdsr/article/view/40/27. https://doi.org/10.33621/jdsr.v2i3.40

Vitak, J., Proferes, N., Shilton, K., & Ashktorab, Z. (2017). Ethics regulation in social computing research: Examining the role of institutional review boards. *Journal of Empirical Research on Human Research Ethics*, *12*(5), 372-382. https://doi.org/10.1177/1556264617725200

Williams, M. L., Burnap, P., & Sloan, L. (2017). Towards an ethical framework for publishing Twitter data in social research: Taking into account users' views, online context and algorithmic estimation. *Sociology*, *51*(6), 1149-1168. https://doi.org/10.1177/0038038517708140

Zwitter, A. (2014). Big data ethics. *Big Data & Society*, *1*(2), 1-6. https://doi.org/10.1177/2053951714559253

APPENDIX

Dataset List: Journal Articles, Conference Papers and Journal Introductions

The following 50 academic articles were used in the dataset. The majority are peer-reviewed studies in academic journals. Also included were some conference papers and introduction editorial sections to journals. They are grouped according to type. Some publications listed in this appendix also appear in the reference list as they were used in data examples and in the body of this article.

Journal Articles - Empirical Studies and Editorials

Adams, S., Veghel, D., & Dekker, L. (2015). Developing a research agenda on ethical issues related to using social media in healthcare. *Cambridge Quarterly of Healthcare Ethics*, *24*, 293-302. https://doi.org/10.1017/S0963180114000619

Andryani, R., Negara, E., & Triadi, D. (2019). Social media analytics: Data utilization of social media for research. *Journal of Information Systems and Informatics*, *1*(2), https://doi.org/10.33557/journalisi.v1i2.23

Bay, M. (2018). Social media ethics: A Rawlsian approach to hypertargeting and psychometrics in political and commercial campaigns. *ACM Transactions on Social Computing*, *1*(4), 1-14. https://doi.org/10.1145/3281450

Berry, D. (2004). Internet research: privacy, ethics and alienation: An open source approach. *Internet Research*, *14*(4), 323-332. https://doi.org/10.1108/10662240410555333

Boyd, D., & Crawford, K. (2012). Critical questions for big data. *Information, Communication and Society*, *15*(5), 662-679. https://doi.org/10.1080/1369118X.2012.678878

Buchanan, E. (2017). Considering the ethics of big data research: A case of Twitter and ISIS/ISIL. *PLoS ONE*, *12*(12), 1-6. https://doi.org/10.1371/journal.pone.0187155

Cesare, N., Lee, H., McCormack, T., Spiro, E., & Zagheni, E. (2018). Promises and pitfalls of using digital traces for demographic research. *Demography*, *55*(5): 1979-1999. https://doi.org/10.1007/s13524-018-0715-2

Clark, K., Duckham, M., Guillemin, M., Hunter, A., McVernon, J., O'Keefe, C., Pitkin, C., Prawer, S., Sinnott, R., Warr, D., & Waycott, J. (2019). Advancing the ethical use of digital data in human research: challenges and strategies to promote ethical practice. *Ethics and Information Technology*, 21, 59-73. https://doi.org/10.1007/s10676-018-9490-4

Côté, I. (2013). Fieldwork in the era of social media: Opportunities and challenges. *Assessment in Political Science*, 46(3), 615-619. https://doi.org/10.1017/S1049096513000565

Crawford, K., & Finn, M. (2015). The limits of crisis data: analytical and ethical challenges of using social and mobile data to understand disasters. *GeoJournal*, *80*, 491-502. https://doi.org/10.1007/s10708-014-9597-z

Eysenbach, G., & Till, J. (2001). Ethical issues in qualitative research on internet communities. *BMJ Clinical Research*, *323*(7321), 1103-1105. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC59687/. https://doi.org/10.1136/bmj.323.7321.1103

Fiesler, C. (2019). Ethical considerations for research involving (speculative) public data. *Proceedings of the ACM on Human-Computer Interaction*, *3*(249), 1-13. https://doi.org/10.1145/3370271

Fiesler, C., & Nicholas, N. (2018). "Participant perceptions" of Twitter research ethics. *Social Media + Society*, *4*(1), 1-14. https://doi.org/10.1177/2056305118763366

Flick, C. (2016). Informed consent and the Facebook emotional manipulation study. *Research Ethics*, *12*(1), 14-28. https://doi.org/10.1177/1747016115599568

Golder, S., Ahmed, S., Norman, G., & Booth, A. (2017). Attitudes toward the ethics of research using social media: A systematic review. *Journal of Medical Internet Research*, *19*(6), e195. https://www.jmir.org/2017/6/e195/. https://doi.org/10.2196/jmir.7082

Gupa, S. (2017). Ethical issues in designing internet-based research: Recommendations for good practice. *Journal of Research Practice*, *13*(2), 1-14. http://jrp.icaap.org/index.php/jrp/article/view/576/476

Gustafson, D., & Woodworth, C. (2014). Methodological and ethical issues in research using social media: A metamethod of Human Papillomavirus vaccine studies. *BMC Medical Research Methodology*, *14*(127), 1-11. https://doi.org/10.1186/1471-2288-14-127

Hesse, A., Glenna, L., Hinrichs, C., Chiles, R., & Sachs, C. (2019). Qualitative research ethics in the big data era. *American Behavioral Scientist*, *63*(5), 560-583. https://doi.org/10.1177/0002764218805806

Jackson, D., Waine, M., & Hutchinson, M. (2015). Blogs as a way to elicit feedback on research and engage stakeholders. *Nurse Researcher*, *22*(3), 41-47. https://doi.org/10.7748/nr.22.3.41.e1300

Jouhki, J., Lauk, E., Penttinen, M., Sormanen, N., & Uskali, T. (2016). Facebook's emotional contagion experiment as a challenge to research ethics. *Media and Communication*, *4*(4), 75-85. https://doi.org/10.17645/mac.v4i4.579

Kennedy, H., & Moss, G. (2015). Known or knowing publics? Social media data mining and the question of public agency. *Big Data & Society*, *2*(2), 1-11. https://doi.org/10.1177/2053951715611145

Kennedy, H., Elgesem, D., & Miguel, C. (2015). On fairness: User perspectives on social media data mining. *Convergence: The International Journal of*

Research into New Media Technologies, *23*(3), 270-288. https://doi.org/10.1177/1354856515592507

Kennedy, H., & Moss, G. (2015). Known or knowing publics? Social media data mining and the question of public agency. *Big Data & Society*, *2*(2), 1-11. https://doi.org/10.1177/2053951715611145

Kinder-Kurlanda, K., Weller, K., Zenk-Möltgen, W., Pfeffer, J., & Morstatter, F. (2017). Archiving information from geotagged tweets to promote reproducibility and comparability in social media research. *Big Data & Society*, *4*(2), 1-14. https://doi.org/10.1177/2053951717736336

Krotoski, A. (2012). Data-driven research: open data opportunities for growing knowledge, and ethical issues that arise. *Insights*, *24*(1), 28-32. Available at: http://eprints.lse.ac.uk/42923/

Kumar, V., & Nanda, P. (2019). Social media to social media analytics: Ethical challenges. *International Journal of Technoethics (IJT)*, *10*(2), 57-70. https://doi.org/10.4018/IJT.2019070104

Lunnay, B., Borlagdan, J., McNaughton, D., & Ward, P. (2014). Ethical use of social media to facilitate qualitative research. *Qualitative Health Research*, *25*(1), 99-109. https://doi.org/10.1177/1049732314549031

Mackay, H. (2017). Social media analytics implications for journalism and democracy. *Journal of Information Ethics*, *26*(1), 34-48.

Mannheimer, S., & Hull, E. (2017). Sharing selves: Developing an ethical framework for curating social media data. *International Journal of Digital Curation*, *12*(2), 196-209. http://www.ijdc.net/article/view/518/494. https://doi.org/10.2218/ijdc.v12i2.518

Moreno, M., Goniu, N., Moreno, P., & Diekema, D. (2013). Ethics of social media research: Common concerns and practical considerations. *Cyberpsychology, Behavior and Social Networking*, *16*(9), 708-713. https://doi.org/10.1089/cyber.2012.0334

Nunan, D., & Yenicioglu, V. (2013). Informed, uninformed and participative consent in social media research. *International Journal of Market Research*, *55*(6), 791-808. https://doi.org/10.2501/IJMR-2013-067

Nunan, D., & Domenico, M. (2013). Market research & the ethics of big data. *International Journal of Market Research*, *55*(4), 505-520. https://doi.org/10.2501/IJMR-2013-015

Powell, T. (2019). Can quantitative research solve social problems? Pragmatism and the Ethics of Social Research. *Journal of Business Ethics*, *167*, 1-8. https://doi.org/10.1007/s10551-019-04196-7

Samuel, G., & Buchanan, E. (2020). Guest editorial: Ethical issues in social media research. *Journal of Empirical Research on Human Research Ethics*, *15*(1-2), 3-11. https://doi.org/10.1177/1556264619901215

Samuel, G., Derrick, G., & van Leeuwen, T. (2019). The ethics ecosystem: Personal ethics, network governance and regulating actors governing the use

of social media research data. *Minerva*, *57*, 317-343. https://doi.org/10.1007/s11024-019-09368-3

Shaw, D. (2016). Facebook's flawed emotion experiment: Antisocial research on social network users. *Research Ethics*, *12*(1), 29-34. https://doi.org/10.1177/1747016115579535

Sinclair, M. (2017). Evidence based midwifery. *Evidence Based Midwifery*, *15*(2), 39. https://www.rcm.org.uk/media/2787/evidence-based-midwifery-june-2017.pdf

Sormanen, N., & Lauk, E. (2016). Issues of ethics and methods in studying social media. *Media and Communication*, *4*(4), 63-65. https://doi.org/10.17645/mac.v4i4.793

Stevens, G., O'Donnell, V., & Williams, L. (2015). Public domain or private data? Developing an ethical approach to social media research in an inter-disciplinary project. *Educational Research and Evaluation*, *21*(2), 154-167. https://doi.org/10.1080/13803611.2015.1024010

Taylor, J., & Pagliari, C. (2018). Comprehensive scoping review of health research using social media data. *British Medical Journal*, *8*, 1-8. Available at: https://bmjopen.bmj.com/content/8/12/e022931. 10.1136/bmjopen-2018-022931

Taylor, J., & Pagliari, C. (2018). Mining social media data: How are research sponsors and researchers addressing the ethical challenges? *Research Ethics*, *14*(2), 1-39. https://doi.org/10.1177/1747016117738559

Tinati, R., Halford, S., Carr, L., & Pope, C. (2014). Big data: Methodological challenges and approaches for sociological analysis. *Sociology*, *48*(4), 663-681. https://doi.org/10.1177/0038038513511561

Varley-Winter, O., & Shah, H. (2016). The opportunities and ethics of big data: practical priorities for a national Council of Data Ethics. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, *374*(2083), 1-7. https://doi.org/10.1098/rsta.2016.0116

Warrell. J., & Jacobson, M. (2014). Internet research ethics and the policy gap for ethical practice in online research settings. Canadian *Journal of Higher Education*, *44*(1), 22-37. https://doi.org/ 10.47678/cjhe.v44i1.2594

Zimmer, M. (2010). "But the data is already public": On the ethics of research in Facebook. *Ethics and Information Technology*, *12*, 313-325. https://doi.org/10.1007/s10676-010-9227-5

Zwitter, A. (2014). Big Data ethics. *Big Data & Society*, *1*(2), 1-6. https://doi.org/10.1177/2053951714559253

Conference Papers

Bowser, A., & Tasi, J. (2015, May 2015*). Supporting ethical web research: A new research ethics review* (Paper presentation). WWW '15: Proceedings of the 24th International Conference on World Wide Web May 2015, pages 151-161, Florence, Italy. https://doi.org/10.1145/2736277.2741654

Chancellor, S., Birnbaum, M., Caine, E., Silenzio, V., & Choudhury, M. (2019, January 29). *A taxonomy of ethical tensions in inferring mental health states from social media* (Paper presentation). Conference on Fairness, Accountability, and Transparency (FAT* '19) (79-88), January 2019, ACM, New York, United States. https://doi.org/10.1145/3287560.3287587

Hutton, L., & Henderson, T. (2015, August 21). *Some challenges for ethics in social network research* (Paper presentation). NS Ethics' 15: NS Ethics '15: Proceedings of the 2015 ACM SIGCOMM Workshop on Ethics in Networked Systems Research, (13-15), London, England. https://doi.org/10.1145/2793013.2793023

Shilton, K., & Sayles, S. (2016, January 5-8). "We aren't all going to be on the same page about ethics:" Ethical practices and challenges in research on digital and social media (Paper presentation). The 49th Hawaii International Conference on System Sciences (HICSS), (1909-1918), Kōloa, Hawaii. https://doi.org/10.1109/HICSS.2016.242

JOURNAL OF DIGITAL SOCIAL RESEARCH

WWW.JDSR.IO ISSN: 2003-1998

VOL. 4, NO. 1, 2022, 23–48

MEMETIC MOMENTS: THE SPEED OF TWITTER MEMES

Naomi Smith[a] and Simon Copland[b]

ABSTRACT

This paper examines how speed shapes internet culture. To do so, it analyses 'memetic moments' on Twitter, short-lived and rapidly circulated memes that quickly reach saturation. The paper examines two 'memetic moments' on Twitter in 2018 and 2019 to assess how they develop over time. Each case study comprises a week's worth of relevant tweets that were analysed for temporal patterns. We analyse these 'memetic moments' through Lefebvre's (2004) work on rhythmanalysis, arguing that the temporal patterns of memes on Twitter can be understood through his concepts of repetition, presence and dialogue. While seemingly trivial, memetic moments underscore the didactic relationship between social media and news media while also providing a way to approach complex social issues.

Keywords: memes; Twitter; speed; rhythmanalysis.

[a] Federation University Australia, Australia
[b] Australian National University, Australia

1 INTRODUCTION

While much research has focused on the various affordances of online platforms, researchers are also increasingly focusing on the logic of speed at the core of online life and its effect on social media platforms' emotional tenor and tone. Speed is increasingly revealing itself as central to online life, with digital platforms requiring constant attention from users to survive economically. Speed drives what some have described as the 'attention' economy (Terranova, 2012; Venturini, 2019), in which platforms seek ways to keep users' attention in order to drive profits. Speed is integral to this process, with its logic dominating both technical and social aspects of digital platforms.

Technically, for digital platforms to provide an enticing user experience, they need to load quickly and reliably. Facebook, for example, famously restricted its initial rollout to ensure it had the server capacity to meet user demand (Losse, 2012). Without server capacity to meet each new influx of users, the site could crash, load slowly or unreliably and, as a result, increase user attrition. Site design, focusing on speed and the sensation of smoothness it brings, is key to one of the most successful digital companies to date, specifically Facebook. This logic is also modelled across a range of other online platforms. The reliance of online platforms on speed is facilitated by the implementation of high-speed internet infrastructure, such as broadband, that encouraged more data-heavy use, including photo and video sharing.

Speed is not just about user experience or the temporal rhythm at which things happen, but also the scale and spread of content. As a consequence, issues like moderation on social platforms become increasingly complex. The volume of content is vast and continues to grow daily. While the scale, growth and production of content is the focus of much scholarly discussion of the problem of online moderation (Gillespie, 2019; Roberts, 2019), we argue so too is the factor of speed. Speed is an important part of how information flows and circulates online. The logic of speed even permeates the rapid cycle nature of policy change in companies like Facebook, which often change policy on a fortnightly basis. Working towards a sociology of digital temporality is an important part of analysing the contemporary digital landscape.

Current analysis of the dynamics of speed frequently assumes it to be one of the more destructive elements of digital platforms. Large social media platforms have been associated with or blamed for the rapid spread of 'fake' or 'junk news' (Venturini, 2019), disinformation (O'Neil & Jensen, 2020), conspiracy theories (Tuters, Jokubauskaité & Bach, 2018) and hateful content (e.g. Massanari, 2017; Salter, 2018). Speed plays an integral role in the spread of this material, with some, in turn, arguing for digital platforms to attempt to 'slow down' discourse online. Suggestions have included the removal or limiting of 'sharing' functions such as the sharing button on Facebook or Retweets on Twitter (Mirani, 2020), as well as the removal of voting functions, such as the like button on Facebook or the voting

function on Reddit (Cook, 2019). These functions play an integral role in boosting and in turn spreading popular content. Recently, in the lead up to the 2020 US Presidential election, Twitter changed the retweet button to default to the "quote tweet" function, turning a one click process into a two click one, and hopefully adding a moment of consideration (or friction) into the retweeting process (Hatmaker, 2020).

It is the intersection of speed, memes and social media that we examine in this paper. Specifically, we examine two memetic moments on Twitter to interrogate the role of speed in the creation and circulation of memes on Twitter. In this paper, we contend that speed can also create a sense of joy, entertainment and even important discussion online. We examine the spread of two memes across 2018 and 2019 on Twitter. These memes focused on political issues in the United States and began after initial tweets, which were then parodied by thousands of users across quick timeframes. These copies and remixes created what we describe as 'emergent memes' or 'memetic moments', which not only function as entertainment, but also facilitate discussion of important political issues.

In this paper, we do two things. First, we analyse the dynamics of the spread of the two identified memes. We do this by conducting a temporal analysis of the development of these memes, conducting an in-depth reading of how these memes developed over time. In doing so, we explore the role of speed in the spread of this material and the dialectical relationship between Twitter memes and the mass media.

2 MEMES AND/AS DIGITAL CULTURE

Memes are a distinctive hallmark of digital culture. Memes are generally jokes that emerge as a by-product of digital culture. Internet memes, Nissenbaum and Shifman (2015) argue, are digital objects that follow a shared pattern or characteristics that allow them to be identified as a unified group. Scholars have argued for the importance of memes as a "prism for shedding light on aspects of contemporary digital culture" (Shifman, 2012, p. 190). Most scholarship examining memes focuses on template style image-macro memes such as the 'first world problems' or 'lol cat' memes that rely on a text/image combination, also called image macro memes (Wiggins & Bowers, 2015). Specifically, memes are a part of participatory digital culture, which Wiggins and Bowers (2012) argue possess a virtual physicality that renders them observable social artifacts in internet culture.

The memes share unifying aesthetic and linguistic cues; this includes the font choice and the central image. Some scholars have argued that other forms of viral content creation as also memes, such as the 'It Gets Better' videos produced in response to homophobic bullying (Gal, Shifman & Kampf, 2015). However, the majority of research that considers memes identifies their humour as a quality that makes them distinct from other forms of viral content. Template or image-based memes are often complex and intertextual artefacts of internet subcultures,

requiring insider knowledge to be correctly read and interpreted. At their core, memes are media-rich jokes that sprawl across form and content (Miltner, 2014). They are commonly conceptualised as a visual, not textual genre, although image-macros memes rely on both visual and text elements for interpretation. Research investigating the social implications of memes is a broad field, covering everything from rhetoric to meme warfare (Bapna & Lokhande 2021). Bapna and Lokhande identify memes as part of what they argue is the 'aestheticisation' of politics within India and beyond (2021, p. 201). They also argue that memes are a way of 'microdosing' political content amongst more general, humorous content on, for example, Instagram meme accounts. While not directly addressing speed, they highlight how messaging services like WhatsApp can quickly spread fake news using memes to millions of users.

This means that less attention is placed on emergent memes, or memetic moments, that are directly responsive to shared cultural or social events and do not utilise image macros. However, these types of memes are increasingly dominant, including on new applications like TikTok, in which sounds and videos fuel its meme-heavy visual culture. Abidin and Kaye (2021, p. 58) argue that audio memes are the "next frontier of meme cultures on the internet, presenting an 'aural turn' in meme ecologies". Audio use may use songs and original audio produced on the app to tell stories and as a way to organise content within the app. In this context, memes function as a way of cataloguing and finding similar information.

While we focus on Twitter as a case study, our analysis can be applied to memes on other platforms, which do not fit the typical image macro-based structure. In expanding the category of a meme in this way, we follow Phillips and Milner's (2017) concept of the ambivalent internet. They identify a number of 'memetic moments' both on Twitter and on other internet platforms, which they categorise as part of the 'ambivalent' internet (Phillips & Milner, 2017). They argue that these moments are often miscategorised as 'trolling' or part of the 'weird' internet. We agree that these types of behaviour are misunderstood and perhaps overlooked. These types of moments should also be understood as memes - in which an initial tweet provides a meme 'template' that is both funny and shareable and allows for a high degree of intertextuality. What distinguishes these memes from the template-based memes, which are the objects of previous research, is their reliance on text. The emergent memes on Twitter are not usually formed around an image, but rely on a readers' sophisticated knowledge of online communication norms and jokes to be read successfully. Emergent memes are the product of a particular subset of online cultures, one that is 'always on' and constantly consuming news and events. It is a culture that prioritises speed. Part of hacking the attention economy of the 'always on' internet culture is to be first, first to spot the joke, first to make the joke. In the following sections, we refer to these memes as 'memetic moments' due to their ephemeral nature. They are less persistent and recursive than image macro memes but point to a broader understanding of memes as social artifacts in participatory digital culture.

3 SPEED AND THE ATTENTION ECONOMY

Speed is increasingly being addressed as a core component of the function of the digital economy, described by theorists as the "attention economy" (Terranova, 2012; Venturini, 2019). However, speed has long been of attention to social theorists. Speed, discussed through the lens of time-space compression, occurs across the work of geographers (e.g. Harvey 1989; Thrift 1996) and sociologists (Beck 2000; Giddens 1990). These accounts grapple with the effects of a range of technologies, from the industrial to the personal, on the social world. In the work of Giddens (1990) and Beck (2000), this is also accompanied by an attendant focus on what they argue is the increasingly rapid pace of social and cultural change; technology has altered the rhythms of life, speeding them up, alongside the rate of change itself. Specific to social relations, Lash (2002) argues that speed means that previously fixed (in place) social bonds have given way to more transient communication bonds that are immediate yet distanciated. Virilio (2000) also identifies similar processes, observing that the 'speeding up' characteristic of late modernity means that social relations requiring time are no longer wholly dependent on a body located in space. While Virilio and Lash are writing before the advent of Twitter, the accelerated sociality he described, which is geographically distant yet incredibly fast, is characteristic of the type of engagement that Twitter encourages.

Lash (2002) also identified an emphasis on information over society as part of the transition to late modernity. This is consistent with early political economy analyses identified 'information' as the core commodity being sold and traded on the internet. However, more recent analyses have argued that attention has replaced information as the new commodity (Terranova, 2012). Digital platforms base their economic model on collecting and selling user data, primarily to advertisers (Srnicek, 2016). Platforms collect the specified data of their users -- including demographic information, interests, and the ways in which they participate -- and sell that data to advertisers to allow them to target ads to specific audiences on the site. Platforms require the attention of their users to maximise this model (Terranova, 2012; Venturini, 2019). Attention provides platforms with more data to sell and ensures greater usage of platforms, which allows for ads to be sold at a higher level. However, this attention is scarce. Users only have so much time in the day and often engage with different platforms at different times. In this context, the speed and spread of material have become integral to the operations of digital platforms. Speed provides users access to new material, keeping people's attention on a platform over time. However, speed is also seen as impoverishing social life. DeLuca, when discussing speed in the circulation of images, describes that in "the ceaseless circulation of images, speed annihilates contemplation, surface flattens depth, flow drowns moments, distraction disrupts attention, affect eclipses meaning, the glance replaces the gaze, reiteration erases originals, and the public screen displaces the public sphere" (2006, p. 87).

In the same way that DeLuca identifies speed as having a deleterious effect on the public sphere, similar arguments are made regarding the influence of speed on social media. Speed has been part of the analysis of social media dynamics, particularly as it concerns fake news. For example, Vosoughi, Roy and Aral (2018) found that fake news, being more novel than true news, spreads more quickly, and this is seen as key to social problems born of misinformation. Venturi (2019) argues that fake news is not new, mimicking longstanding propaganda techniques. The present difference is the speed at which this news can spread, facilitated by digital platforms, which provide the architectural scaffolding for its spread (Venturini, 2019). In sum, Venturini argues that social media prioritise attention and speed in both social and technical design, as well as encouraging users to develop affective relationships to news through liking and sharing. This allows news to travel at great speed through online social networks, becoming 'viral'.

Venturini analyses the operation of speed on digital platforms - with the attention economy being driven by economic, communicative, technological, cultural and political reasons. We see similar driving factors in the speed of memes within our own analysis. The varied literature on memes often references virality and popularity when discussing the spread of memes. These terms, we argue, can also be understood as proxies for speed. However, there is a dearth of research that addresses speed directly as part of the circulation of memes. In Jenkin's (2014) rhetorical analysis of image-macro memes, he argues that accounting for the speed at which memes circulate online is central to understanding the affects they produce. He argues that it is the circulation of memes across digital media platforms, which produce the structures that enable the ongoing remaking of memes. The cross-platform circulation of memes produces their plasticity. The pattern of circulation described by Jenkins is similar to the modes of circulation described by Lefebvre (2004) in his text *Rhythmanalysis*. Lefebvre (2004, p. 7) argues that "not only does repetition not exclude differences, it also gives birth to them; it produces them." Repetition creating difference is also what distinguishes memes from viral content. As Miltner explains, "This is the key difference between 'memetic' content and 'viral' content; if a piece of content is passed along intact and unaltered, it is considered to be viral. If a piece is altered or changed as it is passed along, it is considered to be a meme" (2017, p. 414). Mapping the rate at which memes circulate is difficult because, as Jenkins (2014, p. 446) highlights, "they often emerge in numerous places simultaneously and frequently spread like wildfire."

In the case studies that follow, we begin to assess how speed factors into emergent memes, or memetic moments on Twitter. In addition to the empirical analysis presented, we also draw on Lefebvre's theory of rhythmanalysis to understand the temporal structure, or the speed of memes. In addition to drawing on Lefebvre's work, we also extend Venturini's focus on 'junk'. While the mechanism Venturini identifies in the viral spread of information can cause problems, it also brings joy. The internet has always been a home for the weird, absurd and funny, alongside the loftier utopian ideals. This, we argue, is the kind

of attention economy we can see in our data set, driven by speed. In the next section, we examine our two case studies, examining two text-based memes.

4 CASE STUDIES: MEMETIC MOMENTS

This paper is specifically focused on the role that speed plays in the creation and spread of emergent text-based memes on Twitter. To investigate this, we selected two 'memetic moments' that took place on Twitter in 2018 and 2019 to judge the ways in which they develop over time. We selected one moment centred on US Congresswoman Alexandria Ocasio-Cortez (AOC) as well as the 30-50 feral hogs meme that was a response to back to back mass shootings in the United States in 2019 in El Paso, Texas, and Dayton, Ohio, which in total killed 31 people. In the case studies that follow, we have identified the authors of each initial tweet by name. These tweets have been widely reported on (Benwell & Paul, 2019; Rosenberg, 2018), and in one instance, the author of the original tweet has openly engaged with reporting on their viral moment (Reply All, 2019). We have chosen these case studies as they were widely reported outside of Twitter both during the initial memetic moment and after it had subsided. We will further detail our methods later in the article, but first, we want to outline the substantive events from which these memes emerged to better situate them in the current social and political context.

Alexandria Ocasio-Cortez, also known as AOC, is the Representative for New York's 14th congressional district. Ocasio-Cortez is a high-profile politician elected initially in the United States midterm elections of 2018. Ocasio-Cortez first drew interest due to her substantial social media platform, her relative youth (29 years of age at the time of election) and her progressive policies. She was considered to be a 'surprise' winner over a more established Democratic candidate in the Primary elections. Ocasio-Cortez is also relatively unique among politicians, having held a number of low-wage service jobs, such as bartending prior to her election. These factors, combined with Ocasio-Cortez's viral appearances in bill reading hearings, have meant that she has become a rising political star. It is in the context of this additional attention, and her working-class background, that Ocasio-Cortez became the centre of our first 'memetic moment'. In the lead up to taking up her position in Washington (a two-month delay between being elected and taking up one's seat), Ocasio-Cortez was public about her financial struggles - having no income to bridge the gap between election and commencement. At her 2018 congressional orientation in Washington, Ocasio-Cortez was covertly photographed by an unnamed Hill staffer. This photograph was sent to conservative political writer Eddie Scarry, who then posted it to Twitter with the caption "Hill Staffer sent me this pic of Ocasio-Cortez they took just now. I'll tell you something: that jacket and coat don't look like a girl who struggles." (see Figure 1)

Figure 1: @eScarry tweet

The response from Twitter users was quick, funny and revelled in the absurdity of the situation. Ocasio-Cortez responded sharply and quickly, and mainstream media outlets were quick to point to the tweet as an example of sexism in politics (Bell, 2018). As will be discussed below, *the hill staffer sent me this pic* is a variation of template memes (Wiggins & Bowers, 2015), and in turn, a valuable example to examine how speed affects memes on Twitter.

The second memetic moment selected for analysis is the 30-50 feral hog meme. The 30-50 feral hogs meme started as a response to a weekend of mass shootings in the United States. On August 3, 2019, in a Walmart in El Paso, Texas, a gunman shot and killed 22 shoppers while injuring 24 others. This was followed by another mass shooting on August 4, 2019, in Dayton, Ohio, where the shooter, using an automatic assault rifle, shot and killed nine victims and wounded 17 others. Two mass shootings so close together renewed debates about gun control in the United States, particularly around automatic, assault-style weapons used in both shootings. Subsequently, musician Jason Isbell tweeted, "If you're on here arguing the definition of "assault weapon" today you are part of the problem. You know what an assault weapon is, and you know you don't need one."

The tweet was popular, being liked and retweeted thousands of times. Amongst the replies, one stood out. William McNabb, who was not a high-profile

user at the time, responded, "Legit question for rural Americans – How do I kill the 30-50 feral hogs that run into my yard within 3-5 mins while my small kids play?" (see Figure 2)

Figure 2: @JasonIsbell and @WillieMcNabb tweets

The next day the jokes began, and the absurdity of the response, along with the format of the original tweet, made it ideal meme material. The meme quickly became intertextual - remixed as song lyrics, as alternative movie titles, and with already existing memes. This meme was also followed by serious political and policy discussion, with stories following about the feral hogs problem that exists in much of rural America (see, for example, Reply All, 2019).

While these memetic moments are focused on humour, they are also political in nature, as well as touching on broader questions of class, gender and violence in contemporary American life. In the latter parts of this paper, we will examine how these memes develop on Twitter, with a particular focus on how quickly they emerge and disappear.

5 METHODOLOGY

We collected data associated with two internet memes across 2018 and 2019. Data was collected from Twitter using the package Twint (2021). For each meme, we collected one week's worth of data, searching for terms associated with the meme, with data collection beginning on the date of the initial tweet that sparked the memetic moment. Data were initially collected for a period longer than one week, but our analysis has been limited to this time period as very little engagement occurred outside it. Data collected included raw tweets and associated metadata, such as retweet count, favourite count, and mentions of other users. The meme associated with the American Congresswoman Alexandra Occasio Cortez included

two terms, while the 30-50 feral hogs meme only included one term. The terms and dates of the data analysis were as follows:

1.) “Hill staffer sent me this pic” and "girl who struggles": 15th November 2018 – 22nd November 2018

2.) “30 - 50 feral hogs”: 4th August 2019 – 11th August 2019

These searches did not include hashtags, although hashtagged material does appear within each data set. We chose these search terms as they represented the core ‘memeable’ aspects of the original tweet, or the parts of the original tweet required for subsequent memes to be recognised and referencing the original. After we collected the data, we then cleaned each data set, removing duplicate tweets, tweets in foreign languages and any tweets not associated with the relevant meme. In total, this created a dataset of 8,585 and 54,086 tweets, respectively.

There are some limitations in these data sets. First, our data set only contains original tweets associated with the terms and does not include retweets. While the data set contains the number of retweets (at time of collection) for each tweet, we cannot see who has retweeted these tweets, and this makes further analysis of these retweets impossible. This limits our understanding of who was engaged in the relevant memes, particularly those who participated by sharing other people’s material. Second, our data set is limited to tweets containing the exact terms searched for each meme. This means that tweets that contain spelling mistakes, or, more importantly, tweets that may include the creative twists on the original tweet, are not included in the data set. This likely represents tweets that are memes of the original meme. The search terms also limit our examination of conversations associated with the memes. If a user replies to a tweet about the meme, for example, but does not include the relevant words, these tweets are not included in the data set. However, as our focus in this research is on the spread of memes connected to these specific terms, this does not necessarily limit our examination of conversations and engagement around the content too heavily.

We have analysed the data using the analysis tool Tableau. Tableau provides quick and clean analysis options for digital data. Using this tool, we have been able to analyse the development of each meme across time, as well as study the key users associated with each meme. We have been able to use this tool to then conduct an in-depth analysis of the development of each meme over time, with a specific focus on the starting points of the meme - i.e. the moments in which memes were sparked. As we detail below, we argue that the moments that spark these memes are essential to understanding both the context and their temporal structure.

6 DATA ANALYSIS

The first clear observation associated with these datasets relates to the size of each. As already noted, the datasets consisted of 8,585 (Hill staffer sent me this pic/girl who struggles)and 54,086 ("30-50 feral hogs) tweets, respectively. We can see a large discrepancy between the meme associated with the American politician Alexandria Ocasio-Cortez (AOC) and the 30-50 feral hogs tweet. As we will discuss later, this likely represents that the meme associated with AOC only reached a relatively small and niche audience of more politically minded tweeters, while the 30-50 feral hogs meme was able to go beyond this smaller community. This, we theorise, is in part due to gun violence affecting far more people in the United States, resulting in the meme having greater spread across the country.

Our first key question in our data analysis was to ask how quickly did the relevant memes spread online and what were the key moments associated with this spread? To answer this question, we created temporal graphs of the two datasets, examining the number of tweets per hour following the initial tweet, which started the memetic moments.

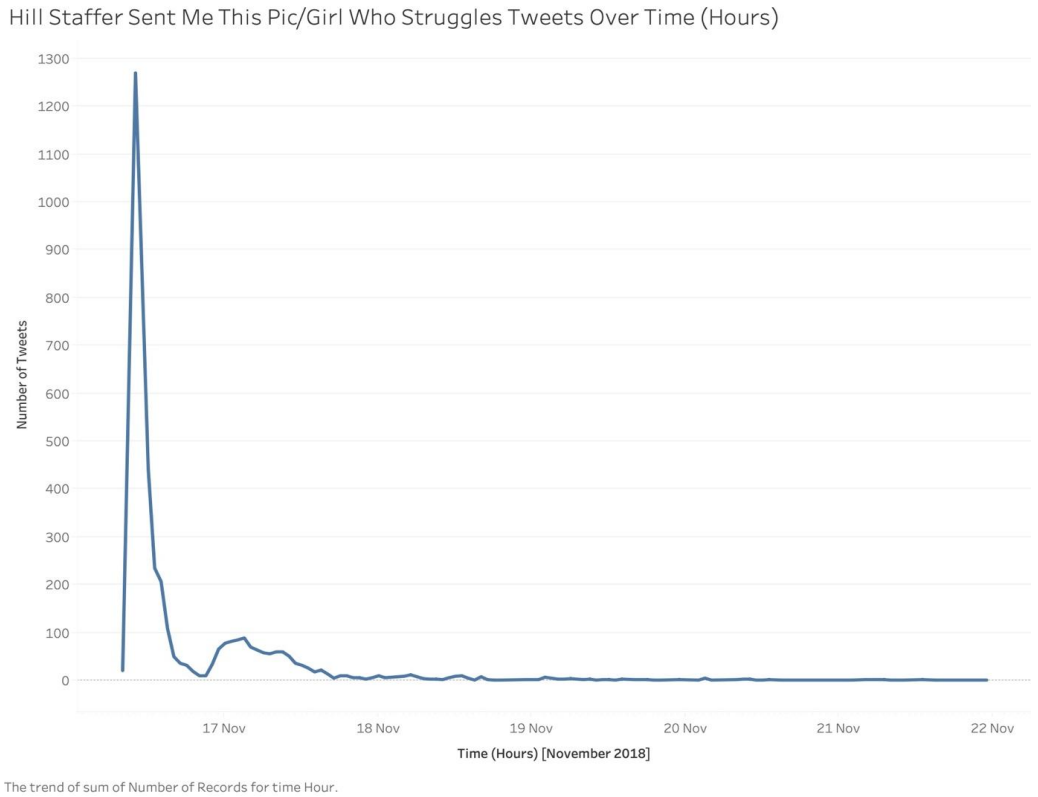


Figure 3: "hill staffer sent me this pic" and "girl who struggles" tweets over time

As is observable in Figures 3 (above) and 4 (below), the two memes developed in a similar fashion. On the left-hand side of each graph is the initial tweet, which is then followed by a very large spike in activity. These spikes consist of 1,270 an hour for the AOC meme, and 3,715 tweets in the case of the 30-50 feral hogs meme.

This spike can last as little as two to three hours, as in the meme associated with AOC, or approximately 12 hours, as in the case of the 30-50 feral hogs memes. Spikes in activity all occur during the daytime hours in the United States, where each meme originated. We then see a sharp drop off of activity as Americans go to sleep, although some other parts of the world pick up on the discussion during their working hours. We then see a small bump of activity on the following day in the United States, although it never reaches peaks anywhere close to the original period of activity. Activity associated with each meme then trails off sharply.

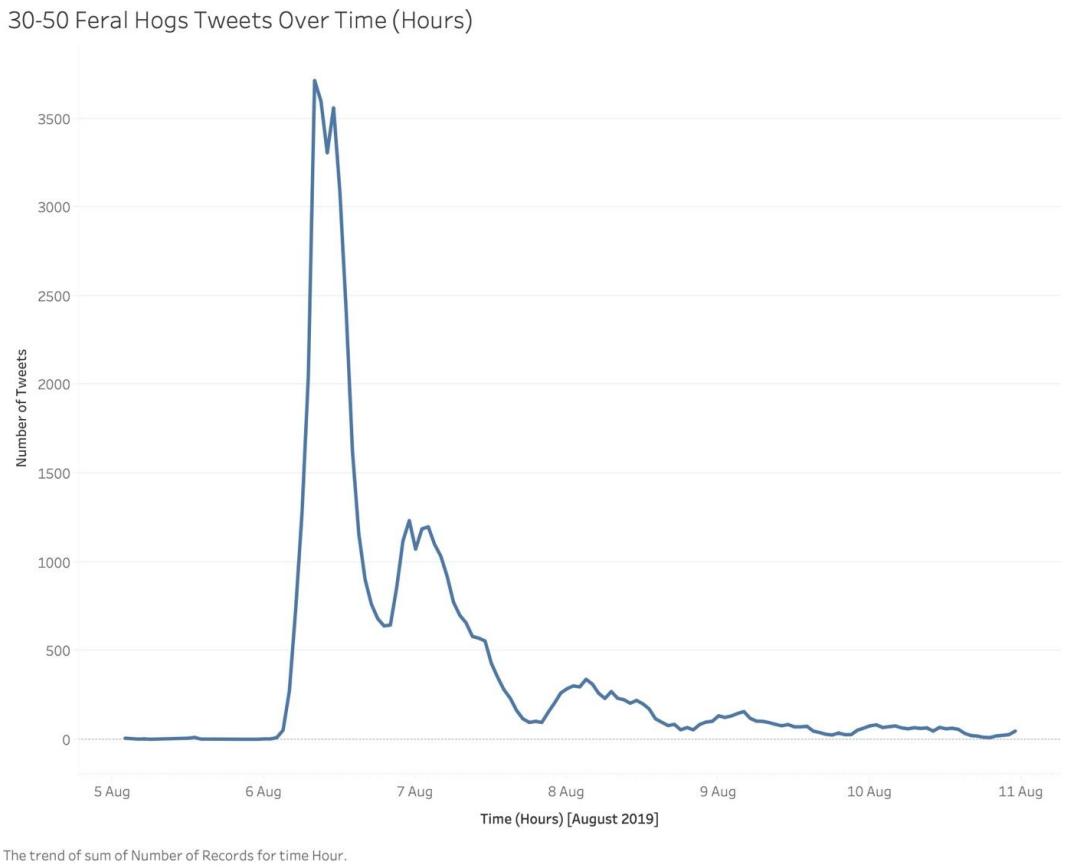


Figure 4: 30-50 Feral Hogs tweets over time

What we can see is a very quick spread of these memes, followed by an almost as quick decline. Similar to Venturini's (2019) notion of 'junk', these memetic moments represent bubbles - like sugar highs - that burst quickly and then almost vanish entirely. Extending Lefebvre's work on cyclical and linear analyses of rhythm, we can understand these memes as part of the cyclical nature of Twitter as they exist in opposition to the linear or to the "monotonous, tiring and even intolerable" (2004, p. 76)

As part of our analysis of the role of speed in these memetic moments, we want to analyse how the memes develop over time. We are particularly interested in the early stages – the moments that result in the immediate spikes of popularity for each of these three memetic moments. As Jenkins (2014) notes, mapping the

rate of circulation of memes is often difficult due to their emergence in numerous spaces simultaneously. We note these challenges through our initial temporal analysis with Tableau, which, while giving us a sense of spread over time, does not analyse links between each tweet, simply positioning each along a timeline as individual nodes. We cannot see, for example, if and how one tweet is inspired by the one before it. Tableau also allows us to examine each tweet in closer depth along the timeline, providing opportunities for a qualitative analysis of tweets at particular moments. Using this capability, we have studied the most popular tweets (in terms of both likes and retweets) for each meme that occur at the base of each of these spikes – i.e. the tweets that are the first spark of the creation and spread of the meme. We have done this to identify themes and structures associated with each meme. This provides us with insight into the temporal patterns of these memes.

Through this analysis, we identify two ways these memes spread at the speed they do. The first is viewable in the "hill staffer sent me this pic" and "girl who struggles" meme. This meme has the fastest incline of the two memes, with an immediate escalation with virtually no lead-in time. This is driven primarily by a small number of high-profile 'verified' Twitter users, who provided a template for the meme.

Figure 5: Bob Schooley tweet

The most popular tweet[1] in the early stages of the meme development is from Bob Schooley, an American screenwriter, who, at the time of writing, is a verified Twitter user with over 100,000 followers. Schooley tweeted a direct copy of the initial tweet, with a replaced image (see Figure 5), collecting 3,193 likes and 320 retweets. Here we see the first iteration of the meme, in which Schooley replaces the image of Ocasio-Cortez with a woman on a fashion show runway wearing an oversized hat and large jacket. The extravagance of the runway look implicitly pokes fun at the original tweet.

Figure 6: Andrew Lawrence tweet

Other popular tweets at this stage follow a similar pattern. The two other most popular tweets in this early stage of the meme are tweeted mere minutes after Schooley's tweet. Andrew Lawrence (the Deputy Director of Rapid Response at Media Matters who has over 50,000 followers) and Asawin Suebsaeng (a senior political reporter at The Daily Beast with over 80,000 followers), tweeted content with 240 and 160 likes, respectively. Lawrence takes a slightly different approach[2], featuring an image of a woman wearing a barrel (see Figure 6), and replacing the

[1] https://twitter.com/Rschooley/status/1063188511589752833
[2] https://twitter.com/ndrew_lawrence/status/1063189114323910656

word 'jacket and coat' with 'barrel and suspenders'. Meanwhile, Suebsaeng's tweet[3] (see Figure 7), features a scene from the popular TV show The Simpsons, in which character Milhouse photographs Lisa Simpson in a future scene in which she is popular, wealthy and cheered on by an adoring crowd.

Figure 7: Asawin Suebsaeng tweet

This meme acts as a variation of a 'template meme' (Wiggins & Bowers, 2015). However, in this instance, the text becomes the original template, with the image varied across each version of the meme. This meme acts in a similar way to other image-based memes, such as *distracted boyfriend* meme or the image of Bernie Sanders at the Joe Biden inauguration, in which users shared the same image repeatedly for comedic sake (and at times to make social and political points). What is different in this meme is that it is the text that stays the same, while it is the image that varies with each share. In doing so, it could be read as a version of junk (Venturini, 2019), with the meme presenting a repetitive, simplified, critical response to the initial tweet from @eScarry. However, as Lefebvre (2004) notes, repetition in the meme does not exclude difference, but instead gives birth to it. Each user chooses a unique image to accompany the text of the meme, creating a meme that is both repetitive and similar at the same time, making it easily recognisable and copied from other users.

[3] https://twitter.com/swin24/status/1063189546677035008

This repetition gives birth to a range of political critiques around the treatment of women in politics. Initially, these users used the meme to make a political point. For example, Schooley follows up his initial tweet with a 'thread' criticising conservative media and their obsession with criticising Ocasio-Cortez appearance. The involvement of Lawrence and Suebsaena also indicates the political nature of the meme, as both are closely embedded in the US political scene, from which the first tweet originated. The meme, in turn, becomes a political point in and of itself, with the ridiculousness of the images chosen highlighting the ridiculous standards to which women such as Ocasio-Cortez are expected to live up to.

Figure 8: @iwriteallday tweet

This meme was sparked by a small number of highly politically engaged Twitter users, who saw the original tweet, aimed to make a political point about it, and in doing so, provided a template for other users to follow. The meme can be picked

up quickly because of the ease of this template, with other users able to quickly grab other images from the web to participate. It becomes, in turn, a quick way to both have some fun online, while also making a political point, mixing the absurd with genuine political critique.

Notably, the trend of the early stages of this meme follows throughout its development. The dataset contains three other highly prominent tweets from @iwriteallday, @davidmackau and @matthewacherry. Each three of these tweeters posted a series of tweets parodying the original tweet. @iwriteallday, for example, posted a tweet[4] copying the original text with an image of a man standing in his bedroom wearing ugg boots, a hoodie and holding a handbag (see Figure 8). @davidmackau, a prominent Australian Twitter user, parodied the tweet[5] by copying the original text followed by a picture of Anne Hathaway in the movie Devil Wears Prada (see Figure 9).

Figure 9: @davidmackau tweet

[4] https://twitter.com/IWriteAllDay_/status/1063199471520047105
[5] https://twitter.com/davidmackau/status/1063200091664732161

Finally, @matthewacherry posted several parody tweets, the most popular of which included a photo of one of the main characters from the Netflix show Big Mouth, Nick, a teenage boy going through puberty; in this image wearing a fancy jacket at school[6] (see Figure 10).

Figure 10: @MatthewACherry tweet

These users are relatively prominent journalists and political operatives, with high follower numbers and verification on Twitter. The meme is developed and spread primarily by a small number of politically engaged users. These users are "in the know" and each take and spread the template across a short period of time.

The second way in which we see the memes develop is an inverse of the approach from the "hill staffer sent me this pic" and "girl who struggles" meme. Instead of being driven by a small number of high-profile verified users, the 30-50 feral hogs meme is instead driven by a high number of users who are not 'high profile', but who have still posted extremely popular tweets. We theorise that the ability of non-verified users to lead this meme development is in part due to the high-profile nature of the original conversation. Specifically, @WillieMcNabb's initial tweet was a reply to a tweet from the very popular Jason Isbell, providing visibility. Isbell's initial tweet gained over 50,000 likes and 7,000 retweets, giving it @WillieMcNabb's reply an immediately large audience.

[6] https://twitter.com/MatthewACherry/status/1063207962410434560

What we see following the initial conversation is a range of users riffing off @WillieMcNabb's tweet, with some being popular and others falling flat. This meme took some time to build, with users tweeting the term for several hours before we see the main spike of the meme. Similar to the "hill staffer sent me this pic" and "girl who struggles" meme, the 30-50 feral hogs meme then takes off sharply. This is driven by several different riffs on the content, with users coming up with a range of different ways to poke fun at the original tweet. The most popular of these at the early stage of the meme is from @BarbiturateCate[7] (see Figure 11), which replicates a template from a dating site, in which users can pick between being 'male', 'female' or '30-50 feral hogs', and seeking 'male', 'female' or 'a yard with unsupervised small children to run in to within 3-5 minutes'. @BarbiturateCat notes the immense popularity of her tweet herself, replying to it some hours later "Ok I think this is my best tweet I must now retire from Twitter dot com forever."

Figure 11: @BarbiturateCat

Other users take this in different directions. For example, @kamilumin tweets "30-50 FERAL HOGS IN YOUR AREA CHAT NOW", satirising common pop-up advertisements, which advertise sex work services by stating that 'women are in your area' and encouraging men to 'chat now!'. Others just make fun of the scenario that @WillieMcNabb sets up. For example, @tinybaby uses a popular tweet[8] format to ridicule McNabb (see Figure 12). This format of tweet mocks people for saying things that no one else is thinking – starting with the word "Nobody" followed by silence and then repeating that with "No one" and "Not a single person" until the

[7] https://twitter.com/BarbiturateCat/status/1158450759576502273
[8] https://twitter.com/tinybaby/status/1158418843393331200

initial content is thrown in. This highlights the seemingly ridiculous nature of the initial tweet, poking fun at McNabb for saying something that no one else would even contemplate.

Figure 12: @tinybaby tweet

Notably, unlike the "hill staffer sent me this pic" and "girl who struggles" meme, these tweets are largely non-political in nature. While many replies to @WillieMcNabb's tweet criticise his position around gun control, the meme itself heads off in a non-political route. Instead, users take the seeming ridiculousness of the initial tweet into further ridiculous directions. One of the most popular tweets is from the user @nomiddlesliders and reads, "take me down to the paradise city where the hogs are feral and there's 30-50". This tweet replaces the lyrics of the popular song *Paradise City* by Guns N' Roses, but in doing so, injects a moment of humour around a difficult debate. At the time of writing, @nomiddlesliders has only 439 followers, yet this tweet amassed over 59,000 likes and 10,000 retweets, once again highlighting the capacity of 'regular users' to spread this meme.

In comparison, the *hill staffer sent me this pic/girl who struggles* meme followed a template approach, in which the same text was replaced by different images. This meme takes off in a range of different directions. While the term "30-50 feral hogs" remains central, each tweet has a different theme - from jokes about the dating industry, attempts to mock the initial tweeter, and riffs off song titles. The flexible format of this memetic moment allows for a wide variety of creative responses.

While initial tweets do not follow a specific template, the 30-50 feral hogs meme then gets adapted to popular memes on social media platforms such as Twitter. The author @jasonarnopp tweeted a call out for people to tweet in '30-50 feral hogs movies'[9] (see Figure 13). This is part of a recurring meme that occurs on Twitter where users replace movie or other popular culture titles with a suggested

[9] https://twitter.com/JasonArnopp/status/1158494270170501120

word. @jasonarnopp's tweet received over 1,000 replies with people sending in movie titles ranging from "When Harry Met 30-50 Feral Hogs" to "How to lose 30-50 feral hogs in ten days". Other prominent users in the data were also associated with similar tweets. @wildheartreads, for example, started a '30-50 feral hogs book titles' thread - i.e. "Interview with 30-50 Feral Hogs", which received over 100 replies, while @four4thefire also tweeted a separate '30-50 feral hogs movie titles' thread, which also received over 100 tweets.

Figure 13: @JasonArnopp tweet

The 30-50 feral hogs meme operated in an inverse nature to the "hill staffer sent me this pic" and "girl who struggles" meme, with 'regular' users driving the content in a non-template form. Unlike the previous example, 30-50 feral hogs develops in a range of different directions, with users using their creativity to riff off the meme. However, we theorise that the very potential of the meme is driven by the initial (unwitting) involvement of Jason Isbell, whose initial tweet provided significant exposure to Willie McNabb's reply. The speed of this meme is driven by this initial exposure, with users riffing off this and taking it in a range of different directions. After initial slowness, the meme then takes a life of its own, spreading quickly across Twitter, becoming incorporated in other social media memes.

7 DISCUSSION

As part of our data collection process, we manually cleaned the tweets collected, removing irrelevant content, and content in languages other than English. To do this, we read each tweet to determine its relevance and suitability for analysis. However, this task also sensitised us to a surprising aspect of the data. The data contained almost no angry, hateful, sexists, racist or otherwise abusive tweets. This was unexpected. Twitter receives a lot of attention for the abusive behaviour on its platform, which appears to be widespread and part of a systemic failure of design and governance. The topics of the memetic moments we collected data for, Ocasio-Cortez and gun control, are political hot button issues. Ocasio-Cortez is abused for her gender, race and political beliefs. Gun control and gun violence is a particularly

divisive issue. Because of this, we were not expecting to see the data comprising primarily good-natured involvement.

The speed at which the memes grew and 'burst' seemed to create a space for low-stakes, non-serious and wholesome engagement. The memetic moments also reinforced the standing of social media spaces like Twitter as networked publics (Ito, 2008), ones that have their own rules and ways of surfacing and responding to content outside of Twitter's attempts at algorithmic curation.

It can be easier to dismiss memes as artefacts of internet vernacular, but memetic moments are also additionally generative outside of the confines of Twitter itself. Lefebvre's (2004) work is again useful here, specifically his distinction between the present and presence. He defines the difference between these two states as follows: "The present is a fact and effect of commerce; while presence situates itself in the poetic: value, creation, situation in the world and not only in the relations of exchange" (Lefebvre 2004, p. 47). More simply, in a digital sense, the present belongs to the commercialised attention economy, to algorithms that attempt to hack user engagement. Presence is the ambivalent or weird internet (Philips & Milner 2017), the emergence of memetic moments that are creative, sometimes poetic and fundamentally linked to the social world, not just relations of exchange. The generation of dialogue also categorises Lefebvre's concept of presence; this is evident in the two memes chosen as case studies for this paper. For example, the 30-50 wild hogs created further discussions about gun control and environmental management. Reply All's (2019) deep dive into the meme highlighted how wild hog populations are exploding and all but unmanageable for many farmers in the south of America. Wild hogs not only destroy farmed crops but can also have a devastating impact on local ecosystems. Effective management of wild hogs is mired in bureaucracy and politics, leaving landowners to manage the problem on their own, usually with the help of high-powered automatic weapons. While the meme hinges on the ridiculousness of citing rampaging wild hogs as a reason for not advancing restrictions on automatic weapons, the memetic moment also has a 'long tail' of discussion and analysis long after the initial meme bubble has burst. The creativity evident in the production of memes on Twitter affords a social presence. Similar is true for the memetic moments centring on Ocasio-Cortez, which generated a number of opinion pieces about gender (Bell, 2018), dress and class (De Valle, 2018). While memetic moments are spontaneously generated sites of engagement, they also serve as a springboard for deeper discussion across other forms of media, underscoring the political and cultural power of memes.

The 'long tail' of memetic moments also illustrates how intertwined 'traditional' media has become with social media. Traditional media outlets rely heavily on Twitter to both generate and editorialise content. Each memetic moment/bubble detailed in this paper generated a slew of media think pieces (e.g. Bell, 2018; De Valle, 2018). Likewise, comments on Twitter about these events are also used to illustrate journalistic content, like a digital vox pops.

8 CONCLUSION

In this paper, we identified memetic moments as primarily text-based and as an emergent property of activity on social media. Previous research has focused on social network analysis of Twitter and memes. These studies (Dang et al. 2019; Jafari Asbagh 2014; Segev et al. 2015; Xu et al. 2016) have largely focused on political discussion and the spread of hateful material or disinformation on the platform.

Despite their similarity, the memetic moments in this paper demonstrate how humour functions as part of the speed of online platforms. Speed is not just a destructive function of digital platforms, but can also create moments of creative engagement and dialogue in online spaces. The creative, repetitious and quickly-paced circulation of memes on Twitter creates a sense of 'presence' (Lefebvre, 2004). Repetition creates enough difference for engagement, while still producing identifiable content to create a clear 'moment'. This facilitates a quick uptake, with users building off the success of previous tweets in order to turn these moments into memes. As we noted, this occurs in two primary ways - one driven by high profile users, and another by 'regular' users, who build off a high-profile conversation to turn it into a meme. While our analysis is limited to only two memes, we argue that this provides insight into the temporal rhythms of memes on Twitter, extending our understanding of how aspects of time, such as speed, function in digital cultures.

As Segev et al. (2015) highlight, memes are often considered trivial and/or mundane parts of the online environment. Memes are playful, silly and sometimes fleeting. Aside from their political resonance, these memetic moments remind us of the positive promise of the internet that often seems like a lost dream, collaborative, fun, intimate. The condition of speed that produces virality can be largely blamed for many of the more toxic elements of social media, including the persistent spread of misinformation. Still, speed can also work to knit publics together. Alongside understanding the harms of social media and the internet more broadly, it also seems important to focus on how social media can produce afford presence, and as an extension, dialogue. Despite the ongoing critique of speed in scholarly literature, speed does not straightforwardly impoverish communication, but can be part of the joyful, creative and poetic rhythms of the social world (Lefebvre, 2004).

FUNDING STATEMENT AND ACKNOWLEDGMENTS

The author(s) received no financial support for the research, authorship, and/or publication of this article. The authors would like to acknowledge the two anonymous reviewers for their helpful feedback and Dr Clare Southerton for advice and feedback.

REFERENCES

Abidin C. and Kaye, B. (2021). Audio Memes, Earworms, and Templatability: The 'Aural Turn' Of Memes" in eds. Arkenbout, C, Jack Wilson and Daniel de Zeeuw, *Critical Meme Reader: Global Mutations of the Viral Image* (pp. 58-68). Amsterdam: Institute of Network Cultures.

Ackland, R. (2013). *Web social science: concepts, data and tools for social scientists in the digital age*. London: SAGE.

Bapna A., and Lokhande, A. (2021)."Post-Truth Design Tactics in Memes in India: 198 Decoding Malicious Design Practices in Right-Wing Meme Factories" in eds. Arkenbout, C, Jack Wilson and Daniel de Zeeuw, *Critical Meme Reader: Global Mutations of the Viral Image* (pp. 198-216). Amsterdam: Institute of Network Cultures.

Beck, U. (2000). *The Brave New World of Work*. Cambridge: Polity Press.

Bell, C. (2018, November 16). Journalist criticised for 'sexists' Ocasio-Cortez tweet. *BBC News*. Retrieved from: https://www.bbc.com/news/blogs-trending-46233260.

Benwell, M. and Paul, K. (2019, August 6). What about the 30-50 feral hogs? Man's defence of assault weapons goes viral. *The Guardian*. Retrieved from: https://www.theguardian.com/us-news/2019/aug/05/feral-hogs-memes-twitter-30-50-running-into-my-yard-small-kids.

Cook, H. (2019, October 9). Facebook removing likes is good for social media. *Mumbrella*. Retrieved from: https://mumbrella.com.au/facebook-removing-likes-is-good-for-social-media-601557.

Dang, L., Chen, Z., Lee, J., Tsou, M.-H. & Ye, X. (2019). Simulating the spatial diffusion of memes on social media networks. *International Journal of Geographical Information Science, 33*(8), 1545–1568, https://doi.org/10.1080/13658816.2019.1591414

DeLuca, K. M. (2006). "The Speed of Immanent Images," in ed. Hope, D. *Visual Communication: Perception, Rhetoric, and Technology* (pp. 79-90). New York, NY: Hampton Press.

De Valle, G. (2018, November 16). The real reason conservative critics love talking about Alexandria Ocasio-Cortez's clothes, *Vox*. Retrieved from: https://www.vox.com/the-goods/2018/11/16/18099074/alexandria-ocasio-cortez-clothes-eddie-scarry.

Gal, N. Shifman, L. & Kampf, Z. (2016). "It Gets Better": Internet memes and the construction of collective identity. *New Media & Society, 18*(8), 1698–1714, https://doi.org/10.1177/1461444814568784.

Giddens, A. (1990). *The Consequences of Modernity*. Cambridge: Polity Press.

Harvey, D. (1989). *The Condition of Postmodernity*. Oxford: Blackwell.

Hatmaker, T. (2020, October 10). Changing how retweets work, Twitter seeks to slow down election misinformation. *TechCrunch*. Retrieved from:

https://social.techcrunch.com/2020/10/09/twitter-retweet-changes-quote-tweet-election-misinformation/.

JafariAsbagh, M., Ferrara, E., Varol, O., Menczer, F., & Flammini, A. (2014). Clustering memes in social media streams. *Social Network Analysis and Mining 4*(1), 237, https://doi.org/10.1007/s13278-014-0237-x.

Jenkins, E. S. (2014). 'The Modes of Visual Rhetoric: Circulating Memes as Expressions', *Quarterly Journal of Speech*, *100*(4), 442–466, doi:10.1080/00335630.2014.989258.

Lash, S. (2002). *Critique of Information*. London: Sage.

Lears, R. (2019). *Knock Down the House.* Netflix.

Lefebvre, H. (2004). *Rythmanalysis: Space, Time and Everyday Life*. New York: Continuum.

Losse, K. (2012). *The Boy Kings: A Journey into the Heart of the Social Network.* New York: Simon and Schuster.

Massanari, A. (2017). #Gamergate and The Fappening: How Reddit's algorithm, governance, and culture support toxic technocultures. *New Media and Society*, *19*(3), 329–346, https://doi.org/10.1177/1461444815608807.

Miltner, K. (2014). "There's no place for lulz on LOLCats": The role of genre, gender, and group identity in the interpretation and enjoyment of an Internet meme. *First Monday 19*(8), https://doi.org/10.5210/fm.v19i8.5391.

Miltner, K. (2018). Internet memes. In J. Burgess A. Marwick, & T. Poell *The sage handbook of social media* (pp. 412-428). London: SAGE. https://www.doi.org/10.4135/9781473984066.n23

Mirani, L. (2020, June 29). No more going viral: why not apply social distancing to social media? *The Guardian.* Retrieved from: https://www.theguardian.com/commentisfree/2020/jun/29/social-distancing-social-media-facebook-misinformation.

Nissenbaum, A. & Shifman, L. (2017). Internet memes as contested cultural capital: The case of 4chan's /b/ board. *New Media & Society, 19*(4), 483–501, https://doi.org/10.1177/1461444815609313

O'Neil, M. & Jensen M. (2020). Australian Perspectives on Misinformation. *News and Media Research Centre: University of Canberra.* Retrieved from: https://www.canberra.edu.au/research/faculty-research-centres/nmrc/research/australian-perspectives-on-misinformation.

Phillips, W. & Milner, R. M. (2017). *The Ambivalent Internet: Mischief, Oddity, and Antagonism Online.* Cambridge, UK: Polity Press.

Reply All (2019, October 10). "#149 30-50 Feral Hogs". *Gimlet Media.* Retrieved from: https://gimletmedia.com/shows/reply-all/n8hw3d.

Rosenberg, E. (2018, November 16). Alexandria Ocasio-Cortez wore clothing, a journalist tweeted a photo, and the internet pounced. *The Washington Post.* Retrieved from: https://www.washingtonpost.com/nation/2018/11/16/eddie-scarry-tweet-about-alexandria-ocasio-cortezs-clothing-is-latest-critique-about-class/.

Salter, M. (2018). From geek masculinity to Gamergate: the technological rationality of online abuse *Crime, Media, Culture, 14*(2), 247–264, https://doi.org/10.1177/1741659017690893.

Segev, E., Nissenbaum, A., Stolero, N., & Shifman, L. (2015). Families and Networks of Internet Memes: The Relationship Between Cohesiveness, Uniqueness, and Quiddity Concreteness. *Journal of Computer-Mediated Communication, 20*(4), 417–433, https://doi.org/10.1111/jcc4.12120.

Souroush V., Roy D. & Aral, S. (2018). The Spread of True and False News Online. *Science, 359*(6380), 1146-1151, doi: 10.1126/science.aap9559.

Srnicek, N. (2017). *Platform Capitalism*. Cambridge, UK: Polity Press.

Terranova, T. (2012). Attention, economy and the brain. *Culture Machine,* 13, 1–19. Retrieved from: https://culturemachine.net/wp-content/uploads/2019/01/465-973-1-PB.pdf.

Tuters M., Jokubauskaité E. & Bach D. (2018). Post-truth protest: how 4chan cooked up the Pizzagate bullshit. *M/C Journal, 21*(3), https://doi.org/10.5204/mcj.1422.

Thrift, N. (1996). New Urban Eras and Old Technological Fears: Reconfiguring the Goodwill of Electronic Things', *Urban Studies, 33*(8), 1463– 93, https://doi.org/10.1080/0042098966754.

Twint (2021). Twint Project. Retrieved from: https://github.com/twintproject/twint.

Venturini, T., Jacomy, M. & Pereira D. (2016). *Visual Network Analysis*. Retrieved from: http://www.tommasoventurini.it/wp/wp-content/uploads/2014/08/Venturini-Jacomy_Visual-Network-Analysis_WorkingPaper.pdf.

Venturini, T. (2019). From fake to junk news, the data politics of online virality. In D. Bigo, E. Isin and Ruppert E. (eds), *Data politics: worlds, subjects, rights* (pp. 123-144). London: Routledge.

Virilio, P. (2000). *Open Sky*. London: Verso.

Wiggins, B.E., & Bowers, G.B. (2015) Memes as genre: A structurational analysis of the memescape. *New Media & Society*, *17*(11), 1886–1906, https://doi.org/10.1177/1461444814535194.

JOURNAL OF DIGITAL SOCIAL RESEARCH

WWW.JDSR.IO ISSN: 2003-1998

VOL. 4, NO. 1, 2022, 49–72

IT'S CROWDED AT THE BOTTOM: TRUST, VISIBILITY, AND SEARCH ALGORITHMS ON CARE.COM

Elizabeth Fetterolf[a]

ABSTRACT

Trust, visibility, and the deepening of existing inequalities are major themes within the platform care work literature. However, no study to date has applied these themes to an analysis of worker profiles. I investigate both how workers communicate trustworthiness through their profiles on Care.com, the world's largest care work platform, and which of these profiles are rendered more and less visible to clients. Through a qualitative content analysis of profiles (n=60) sampled from the top and bottom search results in three different US zip codes, I find that visibility is often related to connectivity, response time, and positive reviews, and who is rendered visible mirrors preexisting inequalities. The language of "passion" for the job is common across top and bottom profiles, indicating a contradiction between the deemphasis on professionalization and the high level of connectivity and responsiveness present in top profiles.

Keywords: platform care work; visibility; trustworthiness.

[a] University of Oxford, United Kingdom.

1 INTRODUCTION

In the past several years, the plight of gig economy workers has been increasingly covered in the news (e.g., Bussewitz & Olson, 2020; Conger et al., 2020), academic literature (see Vallas & Schor, 2020 for a review), and within popular culture, as seen by the new documentary *The Gig Is Up* (Walsh, 2021). A clear theme has emerged: the rise of platform-facilitated gig work has, under the guise of increased flexibility, created a precarious, underpaid workforce whose livelihood depends on opaque algorithms and the large technology companies who rely on their labour (Gray & Suri, 2019). But within this large and growing literature, an important kind of platform labour has been overlooked: care work. Before the introduction of platforms, this feminized, racialized workforce performed ununionized and largely unregulated work as nannies, maids, and home healthcare workers; as Ai-Jen Poo, founder of the National Domestic Workers Alliance states, they are "the original gig workers" (Poo, 2017, n.p.).

COVID-19 has highlighted the importance of many kinds of care work, and childcare specifically was put front and centre when school closures across the world left many working parents in need of assistance. The burden has disproportionately fallen on mothers, who experienced an increase in unpaid domestic work and a decrease in well-being (Zhou et al., 2020), a phenomenon that has received significant media attention (e.g., Dickson, 2020; Grose, 2021; Hsu, 2020). However, the pandemic has also taken a massive tole on the "essential and untrusted" childcare workers who are subject to increased levels of surveillance on care work platforms like Care.com and SitterCity (Ticona, 2020). In the low-trust context of the pandemic, opaque algorithms determine which workers are seen and which ones are not.

The present study has two interrelated but distinct aims: I investigate both how workers communicate trustworthiness on Care.com through their profiles, and which of these profiles are rendered more and less visible to clients. Using Noble's (2018) analysis of search algorithms and Ticona and Mateescu's (2018a) work on the key role of worker visibility and trustworthiness on Care.com, I sample profiles from the top and bottom of a general search in three different United States locations and conduct a qualitative content analysis. I find that, despite non-white workers' feelings of hypervisibility on the platform (Ticona & Mateescu, 2018a), they appear largely less visible than their white counterparts. I also find that, in line with existing literature, education, quick response time, and high ratings are used to convey trustworthiness, and yet profile bios often downplay the work of childcare, a phenomenon that provides possibilities for new avenues of research in this emerging literature.

2 LITERATURE REVIEW

2.1 Care Work as Gig Work

Despite the characterization of care and domestic workers as the "original" gig workers (Poo, 2017), care work has been understudied in gig economy research (Kaine et al., 2020; Ticona & Mateescu, 2018a; Ticona et al., 2018). In-home care and domestic work have a long history of being left out of the picture; domestic and childcare workers have been excluded from the labour movement (Burnham & Theodore, 2012; Poo, 2017), nannies provide "hidden" support for white-collar women entering the workforce (Macdonald, 2011), and racialized domestic workers are expected to act "invisible" within white homes (Glenn, 1992). In this section, I will highlight the ways in which platforms mirror past iterations of paid care work. Then, I will draw on the emerging platform care work literature to detail two major changes that platforms have prompted: an emphasis on visibility and a deepening of preexisting inequalities.

First, I will define my use of "gig work" for the purposes of this paper. The definitions of the gig economy and gig work are flexible and have at times been contested (Montgomery & Baglioni, 2021). In their review of the platform labour literature, Schor and Vallas (2020) make a distinction between gig workers who are contracted via a platform but perform the services offline and platform workers who perform short tasks entirely online. In contrast, Bajwa et al. (2018) embrace a broader definition of gig work, using the term to encompass workers who are not employees, are paid by task, and whose work is mediated in some way by a platform. In this review, I use the latter, broader conception of "gig work," while recognizing that the off-platform delivery of the service is a distinctive component of care work.

While some earlier scholars of the gig economy argued that platforms would transform work (Parker et al., 2016) or, in some cases, abolish it (Sundararajan, 2016), platform care work shares much in common with its earlier iterations. A 2018 report on gig work reports low wages, non-employee status, lack of unions, and general precarity among workers (Bajwa et al., 2018). However, these challenges are not new to in-home care workers. In Cooke's (1950/2015) investigative reporting on the "Bronx Slave Market," an area in New York where Black domestic workers would gather to sell their labour by the hour to white housewives, her depictions of the women's uncertainty regarding their next "gig," clients' attempts to renege on pre-agreed payments, lack of unionization opportunities, and strategies for navigating the "informal" yet highly codified marketplace mirror modern studies of gig workers, such as those by Graham and Anwar (2019) and Gray and Suri (2019).

Like modern-day gig work, low wages, and precarious status in the "gray economy" are historic characteristics of in-home care work. A widely cited study by England et al. (2002) demonstrated that relative to other professions, care workers' wages were significantly lower, even when controlling for education and experience.

In-home childcare and domestic workers are rarely full-time employees of the families that hire them, and these arrangements are often precarious, unpredictable, and inconsistent (Burnham & Theodore, 2012). The shape and character of this work has undergone changes that contextualize the introduction of platforms; for example, the former half of the 20th century saw a shift from governesses and servants to nannies and domestic workers (Glenn, 1992) and the rise of neoliberalism came with increased outsourcing of intimate life (Hochschild, 2013).

If invisibility and precarity have historically been characteristics of in-home care work, platforms like Care.com and Sitter City attempt to make it more visible and less "under the table" through increased surveillance of workers (Flanagan, 2019; McDonald et al., 2021; Tandon & Rathi, 2021; Ticona & Mateescu, 2018a). They do so through trust-focused branding, background checks, optional platform-mediated payment mechanisms, and client-facing literature that discourages paying workers in cash (Tandon & Rathi, 2021; Ticona & Mateescu, 2018a; Ticona et al., 2018). Trust was also key in pre-platform care work arrangements, and Flanagan (2019) points out that agencies facilitated trust between families and workers. However, care work platforms differ in their strategy; by making these measures "optional," platforms ensure that risk is transferred from the company to the individual workers and clients (van Doorn, 2017). This abdication of risk is one way that care work platforms protect themselves, while exerting control over their workers (McDonald et al., 2021).

One result of this dual focus on visibility and trust is that care work platforms exacerbate existing inequalities between workers, based on race, class, gender, and immigration status (Flanagan, 2019; Ticona, 2020; Ticona & Mateescu, 2018a; Ticona et al., 2018; van Doorn, 2017). Background checks and mechanisms for issuing pay slips protect the company, while creating barriers for undocumented workers (Ticona & Mateescu, 2018a). In a study of domestic work platforms, van Doorn (2017) found that the language of the "sharing" economy gave the illusion of increased worker freedom and meritocracy, while masking the racialized and gendered legacy that the platforms were built upon. In the United States, this legacy involves women of colour, often immigrants, being employed as domestic servants and doing the "dirty work" of the home for white, middle-class housewives (Cooke, 1950/2015; Glenn, 1992).

This history of entwined sexism, racism, and classism continues to play out on care work platforms today. Inequalities between workers share similarities to what Schor (2017) has called the "crowding out" effect on platforms like TaskRabbit and Airbnb, in which highly educated middle-class workers occupy jobs that used to go to low-income workers without college degrees. The childcare sector has been historically dominated by working-class women of colour and defined by a flow of migrants from the Global South to the Global North (Ehrenreich & Hochschild, 2003). White, college-educated nannies have been more prominent in the sector in recent years (Wu, 2016) and class signifiers on care work platforms, such as education, hobbies, and languages spoken, could potentially

deepen this divide (Ticona et al., 2018). Ollier-Malaterre et al. (2019) observe that workers' ability to successfully balance connectivity and response time has become its own class signifier, a concept that they call "digital cultural capital." Ticona et al. (2018) identified lack of consistent internet access as a barrier for some low-income users – one of their participants cited this as a reason she stopped using the platform. These barriers may be further exacerbated by care work platforms' focus on "personality matching," a process that can be defined by race and class norms (Ticona & Mateescu, 2018a).

Flanagan's (2019) historical analysis sheds light on the "personality matching" process of 19th and 20th century employment agencies, with an emphasis put on moral "character" and values, rather than references or qualifications. This connects to Ticona and Mateescu's (2018a) interviews, in which workers described curating their profiles for potential clients. Just as domestic servants were advertised by their agencies based on a racialized, classed notion of "character" (Flanagan, 2019), on care work platforms, related kinds of class norms may be playing out in presentation of "personality."

2.2 Algorithms and Visibility

While agencies used to provide a medium for raced and classed "personality matching," on care work platforms algorithms perform this work. Algorithms are often opaque to both the researchers who study them and the workers whose time and income they dictate, and there is a large literature devoted to analysing algorithms' mutually shaping relationship with gig workers (e.g., Chen, 2018; Gray & Suri, 2019; Wood et al., 2019). There is an even larger literature that details how algorithms shape everything from the healthcare individuals receive (e.g., Obermeyer et al., 2019), to the news they consume (e.g., Thurman et al., 2018), to their selection of potential partners (e.g., Sharabi, 2021). Furthermore, scholars like Noble (2018) and Eubanks (2018) demonstrate that while algorithms are often seen as neutral tools, they reflect the racism, sexism, and classism of both their creators and society more broadly. Searches for "Black girls" on Google bring up pornographic images (Noble, 2018) while predictive policing algorithms rely on existing crime data and target neighbourhoods of colour, ignoring the fact the data itself comes from previous over-policing of these communities (Brayne, 2017). Bucher's (2016) conception of "the algorithmic imaginary" illustrates the ways in which algorithms inspire feelings in the individuals whose lives they shape, and in turn those same feelings mould the algorithms; this iterative process is key to understanding how systemic oppression creates unjust algorithms and algorithms create new forms of oppression.

However, as Benjamin (2019) points out, these processes can be hidden a lack of transparency on the part of their creators, creating what she calls an "anti-Black box." This opacity is especially relevant to gig workers, given that they typically have little information about the algorithms that shape their work (Vallas & Schor,

2020). Workers using "on-demand" platforms, like ride-hailing and delivery apps (Ticona et al., 2018), are surveilled and managed by frequently changing algorithms, leading to both stress and exhaustion (Newlands, 2021; Wood et al., 2019) as well as possibilities for resistance and "fissures" in algorithmic power (Chen, 2018; Ferrari & Graham, 2021).

On marketplace platforms, algorithms provide visibility based on where a worker appears in a client's search results (Graham & Anwar, 2019; Ticona et al. 2018; Wood & Lehdonvirta, 2021). Analyses of social media algorithms have identified visibility as a kind of double-edged sword for users (Arriagada & Ibáñez, 2020; Bucher, 2012), both the goal and a risk (Rega & Medrado, 2021). In an analysis of Facebook's EdgeRank algorithm, Bucher (2012) turned Foucault's panopticon on its head; for Facebook users, invisibility presents the true threat. For workers on marketplace platforms, this claim holds. Processes of "algorithmic shortlisting" (Williams et al., 2021) or "algorithmic amplification" (Wood & Lehdonvirta, 2021) are key determinants of success. Interviews with workers indicate that they view getting positive reviews and ratings as key to maintaining visibility within the search (D'Cruz & Noronha, 2016; Graham & Anwar, 2019; Ticona & Mateescu, 2018a).

The current literature I have surveyed focuses primarily on either the experiences of workers (e.g., Graham & Anwar, 2019; Tandon & Rathi, 2021; Ticona & Mateescu, 2018a, 2018b) or an analysis of how the platforms themselves function (e.g., Flanagan, 2019; McDonald et al., 2021; van Doorn, 2017). Though there is discussion of how workers perceive their own visibility to clients, there is little empirical investigation of how visible they appear in practice. And while profiles are the crux of how workers get hired and communicate trustworthiness, there are currently no attempts to analyse the profiles themselves as qualitative data. Finally, though there is a rich literature on algorithmic inequality on all kinds of platforms, the ways in which search algorithms could reinforce existing inequalities among care workers have been underdiscussed. This study attempts to fill these gaps by analysing worker profiles on Care.com.

2.3 Case Study: Care.com

I focus on Care.com over other care work platforms, like UrbanSitter or SitterCity, for two main reasons. First and foremost, Care.com is currently the world's largest and most-used care work platform, hosting 31.7 million members across 20 different countries (Care.com, Inc., 2019). Second, like many care work platforms, much of its branding and public-facing materials centre on trust and trustworthiness (Ticona & Mateescu, 2018a); this is encapsulated by the banner on its homepage, which states "find trusted caregivers for your every need" (Care.com, 2021). Though the site offers eldercare, pet care, tutoring, special needs care, and cleaning services, I look at its largest offering, childcare, and limit my analysis to the United States, where most of its paying members reside (Care.com, Inc., 2019).

While this study does not claim to be internationally generalizable, other scholars have noted that themes of reputation, reviews, trust, and exacerbated inequality are relevant on care work platforms in many countries (Flanagan, 2019; McDonald et al., 2021; Tandon & Rathi, 2021).

Care.com, founded in 2006, has been called "Amazon for caregivers" (Farrell, 2014), a comparison that invokes the platform's marketplace nature, in which workers can be viewed as products. Indeed, many of Care.com's affordances are similar to Amazon's (along with other gig work platforms), such as the importance of ratings and reviews, and the "Book Now" feature which promises a quick, frictionless experience for the client. However, Care.com is structured less like Amazon and more like freelancer platforms such as Upwork and Fiverr, where employers can post jobs and search for workers based on a series of criteria such as distance, pay range, and availability. Workers create profiles that detail demographic information (like gender and education, but notably not race), availability, a personal bio, past reviews, and a series of platform-determined icons and badges (for example profile, see Figure 1).

A marketplace platform, Care.com manages the hiring process by "sorting, ranking, and rendering visible large pools of workers" (Ticona et al., 2018, p. 2). However, Care.com's "Book Now" feature, which allows clients to book workers instantaneously based on their listed availability (Care.com, 2021), shares similarities with on-demand platforms on which workers have less control over the client matching process (Shapiro, 2017; van Doorn, 2017).

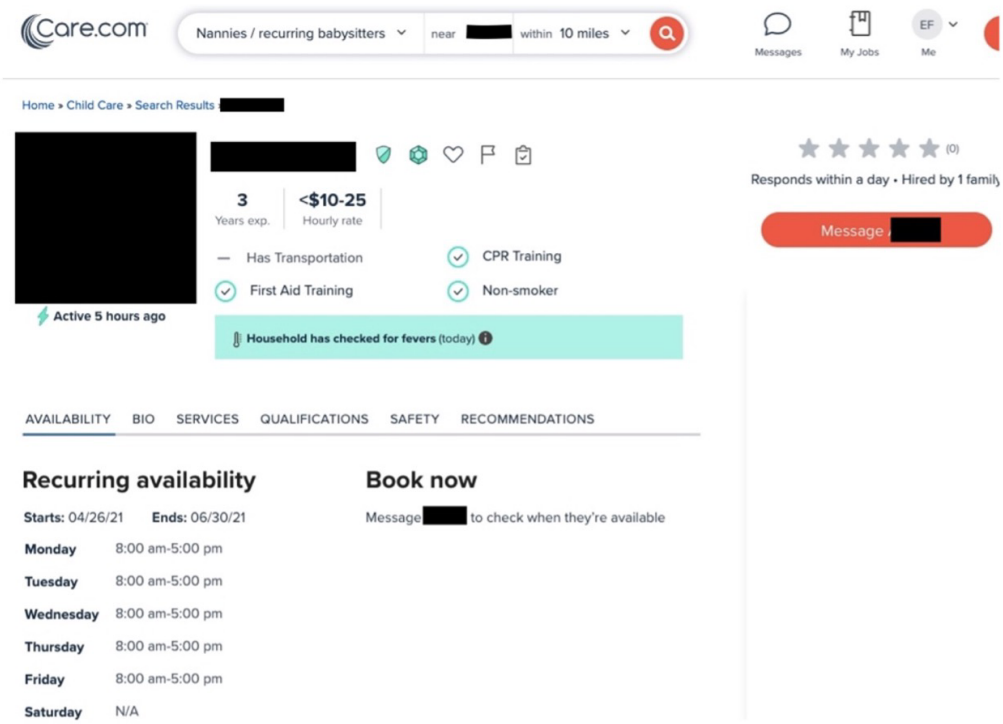


Figure 1. Care.com example profile (Care.com, 2021)

Reputation, reviews, and trust are key to Care.com's functionality. Facilitating "trust between strangers" (Ticona & Mateescu, 2018a, p. 4388) is crucial due to the personal, home-based nature of childcare. Care workers are required to prove their trustworthiness via signifiers, both controlled by the platform and communicated through their bios (Ticona & Mateescu, 2018a). They undergo platform-mediated background checks called "CareChecks," but families are also encouraged to pay for additional background and department of motor vehicles record checks (Gerson, 2019). Workers can also convey trustworthiness by verifying personal information like their cell phone number, email address, and social media accounts, curating their profile (which includes their bio, education, and qualifications), staying active on the platform, and maintaining a quick response time (Ticona & Mateescu, 2018a). As on other marketplace platforms, positive reviews, a five-star rating, and appearing high up in the search are crucial to getting clients (D'Cruz & Noronha, 2016; Graham & Anwar, 2019; Wood & Lehdonvirta, 2021; Wood et al., 2019).

Care.com recently created a "Household Fever Check" status, an optional badge indicating when workers had last checked for fevers at home; no such verification currently exists for clients (Ticona, 2020). Ticona and Mateescu (2018a) observe this lack of reciprocity in nearly all platform-facilitated trustworthiness signifiers. Graham and Anwar (2019) noted a similar asymmetry of information in their interviews with Upwork workers, which limited bargaining power. In the case of care work platforms, it could also compromise worker safety (Ticona & Mateescu, 2018b), given that employer abuse is common (Burnham & Theodore, 2012). Given the features outlined, and the existing literature, trustworthiness and visibility emerge as two crucial, interrelated themes to explore.

3 THEORETICAL FRAMEWORK

To examine this relationship, I apply Noble's (2018) work on how Google search algorithms have shaped trust and knowledge as a theoretical framework. In her work on Google searches, Noble (2018) argues that individuals perceive items that appear at the top of a search as inherently more trustworthy than those that do not. However, she notes that search algorithms are not neutral assessors of the "best" option, but deeply influenced by power, financial resources, and oppression. Research on Google searches have shown that around 70% of search traffic comes from the first page of results, with around 60% of clicks focusing on the top five items (Petrescu, 2014).

Algorithms on marketplace gig work platforms are distinct from search engines in that the client is searching for workers, not simply information. However, I argue that Noble's general theory, that the most visible search results are rendered trustworthy by the algorithm's "objectivity" which is necessarily opaque to the user, is applicable to Care.com. And on this platform, the immediate stakes of this trust are often higher than a Google search for information.

On Care.com, trustworthiness and visibility are intertwined but not mutually exclusive; visibility represents one way in which workers are rendered trustworthy by the platform, but there are other steps they must take to communicate this quality. Here, I use Noble's argument that visibility within a search algorithm engenders implicit trust that the top results will represent what is true. Noble (2018) argues that Google's "enclosure of the public domain" (p. 50) has changed how individuals view information. Most people see Google as an objective receptacle for knowledge and therefore, what appears at the top of the search is trustworthy. I argue that on Care.com, workers appearing at the top of the search renders them more trustworthy; therefore, visibility and trustworthiness are intertwined on the platform, a key finding from Ticona and Mateescu's (2018a) study of workers. Applying Noble's (2018) theory, as well as past work on marketplace platforms and Google searches (Graham & Anwar, 2019; Petrescu, 2014; Wood & Lehdonvirta, 2021), I assume that potential clients trust what they're being shown, and likely will not scroll past the first several pages.[1] I seek to analyse who is rendered implicitly trustworthy by the algorithm, and who is not. Since visibility on the platform represents only one measure of trustworthiness, I also want to examine the signifiers that workers themselves control, such as their qualifications, vaccination status, and the text of their bios.

In sum, this study attempts to understand how workers render themselves trustworthy, while simultaneously being rendered as trustworthy or less-so by an opaque, commercial algorithm. I undertook a qualitative content analysis of Care.com worker profiles, guided by three research questions: *What are the characteristics of profiles at the top of the search? What are the characteristics of the profiles at the bottom? How do profiles at both the top and the bottom communicate trustworthiness to potential clients?*

4 METHODS[2]

I conducted a qualitative content analysis of 60 Care.com worker profiles from three different locations in the US. I sampled the top 10 profiles along with a random sample of 10 profiles from the bottom 20% in each area, in order to investigate which profiles are made most visible and which are unlikely to be seen. The random sample from the bottom was obtained by selecting the last profile on each page in the bottom 20% of total pages. Given that the profiles at the very end of each search might be blank, long inactive, or very new, I chose instead to sample from a section of the search results that few prospective clients were likely to see, given that each zip code had at least several hundred profiles. Search algorithms change frequently based on time, date, and location, so my samples represent exploratory "snapshots"

[1] This assumption, on which my theoretical framework heavily relies, should be tested by future research—to my knowledge, an empirical study of clients has not been conducted yet.

[2] This project was reviewed and approved by the Oxford Internet Institute's Departmental Research Ethics Committee.

(Noble, 2018); longitudinal and more extensive sampling are required to draw generalizable conclusions.

4.1 Data

I chose search locations based on zip code, with the awareness that most search results would be from the surrounding areas. Due to algorithmic opacity (Wood & Lehdonvirta, 2021) it is unclear exactly how search results are sorted on Care.com, but I wanted to simulate the experience of a client from that particular neighbourhood searching for a caregiver. I chose to focus on neighbourhoods in Brooklyn, New York, and Atlanta, Georgia, partly because these were two of the locations used in Ticona and Mateescu's (2018a) interview study. Fremont, California, was selected primarily because it is an affluent suburb with a majority non-white population. Locations were chosen with the goal of variation in terms of race, class, geographic location, and urbanity (see Table 1 for a breakdown).

Table 1. Demographic and Geographic Breakdown of Search Locations[3]

	Brooklyn, NY (Park Slope)	**Atlanta, GA (Northwest)**	**Fremont, CA (Fremont)**
Racial demographics	72.1% White, 7.2% Black, 6% Asian, and 4.9% Hispanic (nonwhite)	88.6% Black, 8.5% White, 2.5% Hispanic (nonwhite), 1.4% Multiracial, 0.5% Asian, and 0.4% Other	58.2% Asian, 24.1% White, 7.8% Hispanic (nonwhite), and 3.4% Multiracial
Median household income	$122,002	$28,017	$127,374
Home ownership	38.2% (majority renters)	47% (majority renters)	61.6% (majority homeowners)
Urban or suburban?	Urban	Urban	Suburban
United States region	Northeast	South	West Coast
City included in Ticona and Mateescu (2018a)?	Yes	Yes	No

4.2 Sampling

I created a free client account, used an incognito window to prevent cookies from affecting the results, and conducted general caregiver searches. To create an account, I provided only a name and email address. I used three filters: I set the pay

[3] Data are from Data USA (2018a, 2018b) and City-Data (2019).

range as wide as possible ($10–50/hour) so that rate would not affect my results, I chose "recurring" versus "one time," and I set the search radius to five miles in New York and Atlanta and 10 miles in Fremont, given the smaller number of profiles.[4] The searches in each city were conducted on different days during the last week of March 2021.

4.3 Analysis

I analysed the text and image descriptions of the 60 profiles in NVivo, using a qualitative content analysis as outlined by Schreier (2014). I coded for platform-mediated trustworthiness signifiers identified by Ticona and Mateescu (2018a), such as five-star ratings, positive reviews, CareChecks, response time, platform activity, qualifications, and education (for an example profile, see Figure 1). Care.com states that Premium accounts increase visibility (Care.com, 2017), so I also coded for whether profiles had a Premium badge, as well as demographic data listed in the profile (age, gender, and languages spoken).[5] I first developed a codebook with two main hierarchical levels: worker demographics and trustworthiness signifiers. Then, I refined the codebook in a trial round, developing thematic codes for the bios and focusing on formal ones, such as ratings and response time, for profile features. Finally, I conducted two rounds of coding, with an interval of one week between each (Schreier, 2014).

5 RESULTS

5.1 Tracking visibility: Characteristics of top and bottom profiles

5.1.1 Response Time and Connectivity

Response time, availability, and activity on the platform proved to be key differences between top and bottom profiles across locations. Workers' response times are listed only if they respond within a few days. Every top profile either had either a quick response time listed, had been active on the platform within hours or days, listed recurring availability, or had all three characteristics. Having both a listed response time and recent activity on the platform was one of the most consistent features of top profiles (see Table 2). In contrast, bottom profiles typically listed no response time. There was one notable exception, which will be discussed below. Client interaction is required to generate response time, so this could be because profiles at the bottom did not have chances for interactions due to lack of visibility. However, some bottom profiles did indicate that they'd been hired by at least one

[4] Care.com requires the client to choose either "one time" or "recurring" before searching.

[5] Given that workers do not typically disclose race but do post a picture, clients' assumptions could affect perception. Therefore, I deemed it necessary to include "perceived race" in my analysis, dependent on the fact that (like many clients) I am a white person making assumptions.

family. Regardless, the relationship between response time and visibility highlights a potentially vicious cycle: more visible profiles list response times, but profiles can only improve their response time through client interaction, which requires increased visibility.

Book Now, a more "on-demand" feature, did not prove common, with profiles that listed it making up only 11% of the total sample. However, all but one of the profiles that did include the feature were in the top 10, which could indicate some algorithmic amplification of profiles using this feature. Further research on this point is required.

A common feature among bottom profiles was the status "Active Over 1 Month Ago," which appeared in 8/10 bottom profiles in Atlanta, 7/10 in Fremont, and 2/10 in Park Slope, with the latter number likely due to Brooklyn's much larger volume of profiles. This could indicate that users who don't engage regularly with the platform are made less visible. Nearly every top 10 profile had been active in at least the past few days, enforcing the idea that regular use of the platform and quick response time are key to visibility.

Table 2. Breakdown of Response Time among Top and Bottom Profiles

	Responds within minutes	Responds within hours	Responds within a day	Responds within a few days	No response time listed
Atlanta (Top 10)	0	3	4	3	0
Brooklyn (Top 10)	0	3	2	4	1
Fremont (Top 10)	0	3	3	4	0
Atlanta (Bottom 10)	0	0	0	0	10
Brooklyn (Bottom 10)	1	0	0	0	9
Fremont (Bottom 10)	0	0	0	0	10

5.1.2 Reputation

Interview studies with marketplace gig workers have found ratings crucial to visibility (Graham & Anwar, 2019; Ticona et al. 2018; Wood & Lehdonvirta, 2021), and the majority of profiles at the top of each location had reviews and five-star ratings. Similarly, most of the profiles at the bottom had no rating or reviews, or, in rare cases, a negative review (see Table 3). The combination of no ratings, no

reviews, and no response time was a common characteristic among bottom profiles across locations. Only one profile had a negative review (one star), and it appeared at the bottom. These results supported Ticona and Mateescu's (2018a) findings regarding worker perceptions that reputation was essential to visibility.

However, a departure from their findings emerged, especially among the top profiles in various locations. There were exceptions at the top and the bottom, with several profiles without five-star ratings listed at the top in all three cities, and at least one five-star profile at the bottom in Brooklyn and Fremont. For example, despite worker sentiment that lack of a five-star rating would preclude them from being boosted by the search algorithm (Ticona & Mateescu, 2018a), only half of the top profiles in Atlanta had one. Patterns among these exceptions provide an interesting window into how visibility might be related to inequalities among workers, as outlined below.

Table 3. Breakdown of Profile Ratings in Top and Bottom 10

	Profiles with five-star rating	Profiles with less than five-star rating	Profiles with no rating
Atlanta (Top 10)	5	1	4
Brooklyn (Top 10)	8	0	2
Fremont (Top 10)	8	0	2
Atlanta (Bottom 10)	0	1	9
Brooklyn (Bottom 10)	2	0	8
Fremont (Bottom 10)	2	0	8

5.1.3 Notable Exceptions Defined by Race

The exceptions to typical characteristics of top and bottom profiles (like response time, activity, and ratings) were often defined by perceived race. There were several instances of white-presenting workers missing key trustworthiness signifiers (such as response time or five-star ratings) in the top 10, as well as Black- and non-white-presenting workers with said trustworthiness signifiers appearing at the bottom. For example, Inez[6] a Black-presenting woman in Brooklyn, had the fastest response time of the entire sample ("within minutes") but appeared in the bottom. Kayleen,

[6] All names are pseudonyms and identifying details have been changed.

a white-presenting woman in Atlanta, had no reviews or rating, but still appeared number two in Atlanta, a majority Black city. She was listed above other more experienced caregivers with five-star ratings, faster response times, and many reviews. In Fremont, Lela and Katie, both young Asian-presenting women, were recently active on the platform, had five-star ratings and at least one positive review, but appeared in the bottom 10. However, Anne, a White-presenting woman of a similar age, had no reviews, rating, and no safety trainings, yet appeared near the top.

5.1.4 Black-presenting Workers Less Visible

These exceptions tie into another finding: Black-presenting workers consistently showed up in the bottom 10. In Brooklyn there were three Black women in the sample, all listed at the bottom. In Fremont there was one Black woman and one Black man, both at the bottom. In Atlanta, where the population of both the neighborhood and the city is majority Black (see Table 1), there were only three Black women in the top 10, out of the eight Black women in the sample. Even in a city with a majority Black population, Black-presenting women specifically were still less visible than their white-presenting peers. Furthermore, some of the Black-presenting women who ended up at the bottom had profiles that possessed many other trustworthiness signifiers, such as five-star reviews and quick response time. For example, Marta, who was listed in the bottom 10 in Brooklyn, had three five-star reviews, a 100% "Would Rehire" rate, and ample recurring availability listed. These findings indicate that despite workers' attempts to communicate trustworthiness (discussed further below), Black-presenting workers could be rendered less visible, and therefore less trustworthy by the algorithm.

This provides a counterpart to Ticona and Mateescu's (2018a) finding that their Black informants engaged in high levels of "visibility management" on the platform, curating their profiles and qualifications for a presumably white clientele. Ironically, while they may bear the personal burden of managing their perceived visibility, within my sample, Black women were rendered less visible by the algorithm.

5.1.5 Premium Accounts

A little over a quarter of the total profiles displayed "Premium" membership, which is available for purchase but not required. However, there seemed to be a relatively equal distribution of Premium profiles across the top and the bottom, with the exception of Atlanta, where 4/5 premium profiles appeared in the top 10. This finding casts some doubt on Care.com's statement that premium members will be "ranked higher in the search results" (Care.com, 2017, n.p.).

5.2 Communicating Trustworthiness

5.2.1 Ubiquity of Higher Education

The vast majority of profiles (53/60) had at least some college experience. On Care.com, education level is another "optional" piece of information that workers must choose whether or not to disclose. However, if they chose not to list their education, that portion of the profile will read as "not listed." Of the seven profiles that had no higher education listed, only two were in the top 10, and both profiles had multiple positive reviews, five-star ratings, and quick response times.

Furthermore, explicit mention of degrees, student status, or educational experience was one of the most common themes in the bios. Many bios made mention of current graduate school studies, college majors, and plans for further education. They often restated the name of their school, despite the fact that this information was already listed on their profile. Mention of education (or having a college degree) was consistent across top and bottom profiles.

5.2.2 COVID Safety

COVID safety was not as prevalent as expected, but profiles that utilized household fever checks or stated vaccine status were highly visible. I analysed three sub-categories within the code "COVID Safety:" household fever check, mention of COVID safety practices (including vaccination status), and statements of COVID boundaries (for example, only wanting to work with families that work remotely). Only five profiles of the 60 actually utilized the Household Fever Check, a finding that was surprising. However, all five of these profiles appeared in the top 10, providing support for the idea that workers who opt into this extra layer of surveillance could be rendered more visible by the platform. Five profiles total listed that they were fully vaccinated, and all appeared in the top 10. The small number of profiles that displayed vaccine status could be due to the sampling dates—at the end of March 2021 many individuals in these three states were not yet eligible. Several profiles mentioned COVID-safe practices, like social distancing or regular testing, but these were scattered among the top and the bottom. Only one profile in Brooklyn stated any kind of COVID boundaries—an experienced, older white-presenting woman in the top 10, who had multiple five-star reviews.

5.2.3 Passion for the Job

The majority of profiles communicated trust by framing childcare as more than just a job. The most common code that came out of analysis of the worker bios was an asserted "passion" for the job, which often was justified by a naturalized love of childcare. The majority of profiles in all three locations used phrases like, "I have a passion for childcare," "I really enjoy helping others," or "it is my passion to take care of people." Profiles used adjectives like "fun-loving," "compassionate," and

"nurturing" and the word "love" appears 43 times within the total sample. Many profiles made reference to childcare being a "fun" job and workers frequently referred to their own families in framing their work experience. One worker cited their history of customer service jobs and remarked, "I'm excited to switch to something I actually like." Here the worker contrasts childcare, framed as a "passion," with a different kind of people-facing job, framed as work.

Within the category of "passion for the job," I identified a sub-code, in which profiles alluded to a naturalized love of childcare, mentioning "innate skills," or asserting that they have "always been pulled towards caregiving." Phrases such as "natural rapport," "natural inclination," or "calling" framed childcare as a profession chosen for the worker, not by them. A worker from Atlanta referred to the fact that she had "always wanted" a big family and "always" nurtured a love of children; both were framed as unquestionable, naturalized facts and followed by a pitch for why this caregiver was right for the job. Clients also employed this language in their reviews. A five-star one stated, "she works with children because she has a genuine, deep-rooted love for what she does." The quote both emphasizes the importance of this worker's "authenticity" as success (and therefore, trustworthiness) and subtly implies that money isn't her primary motivation, love is.

This theme represents one way that profiles communicate trustworthiness to potential clients: asserting that childcare is not "just" a job, but rather tied to "inherent" qualities of one's personality, such as nurturance, compassion, curiosity, and fun. This theme was common in both profiles at the top and the bottom—stating a passion for the job did not appear related to visibility.

6 DISCUSSION

Overall, class norms, disparities in visibility based on race, connectivity, and the ubiquity of declarations of "passion for the job" were some of the most salient findings. In this section, I will discuss these further and tie them to the theoretical work of Noble (2018), along with Schor (2017), Ollier-Malaterre et al. (2019), and others. Schor's (2017) theory of "crowding out" on gig economy platforms could have some traction on Care.com. The ubiquity of some form of higher education on the platform was one of the most consistent features across profiles, with little variation by location. However, instead of workers with higher education being rendered more visible by the platform, workers with higher education made up the vast majority of profiles. This result could indicate that on Care.com a college education might not just be an advantage, but an unofficial requirement. It also supports the idea that "personality matching" that occurs during selection is often rooted in "class norms" (Ticona et al., 2018). A divide between primarily white, college-educated nannies and working-class women of color has already been identified (Wu, 2016), but these findings indicate that Care.com could exacerbate it. More representative data and interviews specifically with workers without college degrees on Care.com could help support this finding.

Listed response time and recent platform activity as key features distinguishing top from bottom profiles indicates that connectivity could be a potential mechanism through which inequality is exacerbated. This lines up with interviews with other types of platform workers, who speak of the need to be constantly present on the platform in order to get jobs (D'Cruz & Noronha, 2016; Wood et al., 2019). Given that lower-income workers might not have quick and easy internet access, (Ticona et al., 2018) visibility's potential ties to constant connectivity could be another way in which the platform strengthens existing inequalities. Personal caregiving duties, familial responsibilities, age, technology literacy, and commitments to other jobs could be additional barriers to connectivity. Beyond literacy, digital confidence, and ability to successfully "crack the code" of the platform itself could be an overarching factor. This quality is similar to Ollier-Malaterre and colleagues' (2019) concept of "digital cultural capital;" managing one's connectivity and presence on multiple platforms is a skill that not only has immense potential monetary value but is frequently defined by class status. Therefore, low-income workers may be at a disadvantage on the platform when it comes to seamlessly managing their online presence in the face of other obstacles.

Black-presenting women's overall lack of visibility parallels Noble's (2018) findings regarding Google searches—Black women are often hurt by search algorithms due to the ways in which racism and gendered oppression are built into their design. Ticona et al. (2018) noted that Black workers feel "hyper-visible" on the platform, but ironically this may not be translating to visibility via the search. Search algorithms are not neutral but shaped by racism and sexism (Noble, 2018). Care.com's algorithm is not publicly available, but here I provide a potential theory of how these divisions might be playing out. Since race is not measured as a demographic variable among workers, it is unlikely (but not impossible) the algorithm itself is directly suppressing the profiles of Black workers. However, as Noble (2018) notes with her analysis of PageRank, Google's algorithm was racist in part because it relied on what was most popular in a deeply racist society, therefore enforcing stereotypes about Black women. A similar process could be occurring on Care.com: profiles that have more engagement (views, messages, favorites) become more visible to clients. But in engaging with workers, clients (especially white ones) might favor white-presenting workers over non-white ones.[7] Hence, without explicitly filtering based on race, the algorithm could render non-white workers less visible, and therefore less trustworthy, to potential clients.

"Passion for the job" as a major theme among top and bottom profiles framed childcare as a passion first, and a job second. This parallels care work platforms' emphasis on "personality matching" (Flanagan, 2019; McDonald et al., 2021; Ticona & Mateescu, 2018a), and indicates that workers may feel the need to frame childcare as a "passion," rather than a "gig" to appear both trustworthy and

[7] Racism among white clients seeking childcare and domestic work is documented by scholars like Ehrenreich and Hochschild (2003) and Glenn (1992).

appealing to clients. This provides support for the argument that "uberization" models don't fit neatly onto Care.com (Ticona et al., 2018); workers on many other gig platforms are not expected to convey their genuine love for the work to clients.

The "naturalized" way in which this passion for the job is presented also highlights the gendered nature of care work, another feature that makes it distinct from other, more commonly studied gig economy platforms. Much has been written on how care work has been historically gendered and naturalized, partly to maintain a low-paid mostly female workforce in the public sphere and an unpaid one in the private sphere (for an overview, see Bhattacharya, 2017). The overwhelming focus on a naturalized love of childcare highlights the gendered nature of the platform—my total sample was 96% female, which is in line with Care.com's 94% female workforce (Care.com, Inc., 2019).[8] Framing care work as a passion rather than a gig may speak to its gendered nature, and it also may speak to the rhetoric of the so-called "sharing economy," (Sundararajan, 2016) an image that has been used to erase the work of gig work, framing a precarious sector as one of free, mutual exchange. Framing care work as a "labor of love," and gig work as "sharing" both serve an extractive purpose, and the combination of these two narratives converge on Care.com

However, this narrative of childcare as passion presents a contradiction, given that the most visible profiles displayed high levels of connectivity and established reputation through ratings and reviews. Workers must be passionate about the work, but they must also be timely, curated, and ready to trade privacy for platform-mediated trustworthiness signifiers. Therefore, this "personal branding" on the platform could also take the form of downplaying the job's "gig-like" nature while still being held to the same exacting standards as workers on other marketplace platforms (Graham & Anwar, 2019; Wood et al., 2019; Wood & Lehdonvirta, 2021). In this sense, the worker must support the gig system by obscuring the labor that goes into it. To be successful (and perceived as trustworthy) they must frame care work as genuine love or passion, rather than a means of making a living. Ironically, work that is framed as passion may be especially vulnerable to exploitation, as Jaffe (2021) argues in her critique of neoliberalism's "do what you love" ethos. Further qualitative research could explore the potential contrast between how workers present their "passion" for the job on the platform and their day-to-day experience of the work.

Finally, the relative lack of emphasis on COVID safety was a surprising finding. However, given that profiles disclosing vaccine status were listed exclusively at the top during a time in which many individuals were not yet eligible, I would expect this trend to continue. Given that there is not currently a way for workers on Care.com to find out information about clients' COVID-safe practices

[8] Care.com displays gender on each worker's profile. There were only two men in my total sample. Interestingly, one of them (a young, white-presenting man) was ranked number one in Brooklyn. The other (a Black-presenting young man) was in the bottom 10 in Fremont.

or vaccine status, the practice of listing health information could continue to become a one-sided worker-surveillance tool.

As Noble (2018) notes, search results are characterized by constant change. Hence one limitation of this study is that my samples represent snapshots, rather than the full picture. Similarly, my samples are small and not meant to be representative. For the top search results, the selectivity is crucial to understanding who appears most visible. However larger samples from the bottom 20% of the search could be useful in future studies. Further research could collect longitudinal data: running the same search at various periods in time, collecting data, and attempting to replicate these findings.

Another limitation is a reliance on my own, biased assumptions regarding perceived race. However, given that workers do not actually list their race on the platform, perceived race plays a large role in how they are evaluated by clients (Ticona & Mateescu, 2018a). Lack of information about Care.com clients is an additional limitation. This study assumes that clients will be inclined to view top search results as more trustworthy and appealing. To my knowledge, no existing study asks clients directly how they evaluate and choose workers on care work platforms. Further research could explore how they establish trustworthiness on the platform.

Overall, my findings support arguments that care work platforms may exacerbate inequalities (Ticona & Mateescu 2018a; van Doorn, 2017), specifically here with visibility. I found that to appear visible and hence trustworthy, Care.com workers, like those on other marketplace platforms, must maintain stellar reputations and quick response times. However, unlike freelancer platforms, delivery services, or ride-hailing apps, workers on Care.com cite passion as a primary job qualification, a phenomenon that speaks to the gendered nature of care work, the illusion of the "sharing" economy, and the specific challenges facing workers on care work platforms. These challenges could and should be explored in future research within this emerging literature.

FUNDING STATEMENT AND ACKNOWLEDGMENTS

Many thanks to Dr. Mark Graham at the Oxford Internet Institute, who oversaw this research, Collin Knopp-Schwyn, who provided crucial feedback and copyediting support, and the two anonymous reviewers, whose thoughtful comments greatly improved this paper's arguments.

REFERENCES

Arriagada, A., & Ibáñez, F. (2020). "You need at least one picture daily, if not, you're dead": Content creators and platform evolution in the social media ecology. *Social Media + Society*, *6*(3). https://doi.org/10.1177/2056305120944624

Bajwa, U., Gastaldo, D., Di Ruggiero, E., & Knorr, L. (2018). The health of workers in the global gig economy. *Globalization and Health*, *14*, Article 124. https://doi.org/10.1186/s12992-018-0444-8

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code*. Polity.

Bhattacharya, T. (Ed.). (2017). *Social reproduction theory: Remapping class, recentering oppression*. Pluto Press. https://doi.org/10.2307/j.ctt1vz494j

Brayne, S. (2017). Big data surveillance: The case of policing. *American Sociological Review*, *82*(5), 977–1008. https://doi.org/10.1177/0003122417725865

Bucher, T. (2012). Want to be on the top? Algorithmic power and the threat of invisibility on Facebook. *New Media & Society*, *14*(7), 1164–1180. https://doi.org/10.1177/1461444812440159

Bucher, T. (2016). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, *20*(1), 30–44. https://doi.org/10.1080/1369118X.2016.1154086

Burnham, L., & Theodore, N. (2012). *Home economics: The invisible and unregulated world of domestic work.* National Domestic Workers Alliance. https://www.domesticworkers.org/wp-content/uploads/2021/06/HomeEconomicsReport.pdf

Bussewitz, C., & Olson, A. (2020, July 5). Gig workers face shifting roles, competition in pandemic. *Washington Post*. https://www.washingtonpost.com/business/gig-workers-face-shifting-roles-competition-in-pandemic/2020/07/05/3589759e-bec8-11ea-8908-68a2b9eae9e0_story.html

Care.com. (2017, June 14). *What is the premium membership on Care.com?*. Retrieved April 21, 2021, from https://www.care.com/c/questions/29145/what-is-the-premium-membership-on-carecom/

Care.com. (2021). [Homepage]. Retrieved April 24, 2021, from https://www.care.com/ppr.do

Care.com, Inc. (2019). *2018 annual report and form 10K*. United States Securities and Exchange Commission. Retrieved May 9, 2021, from https://www.annualreports.com/HostedData/AnnualReports/PDF/NYSE_CRCM_2018.pdf

Chen, J. Y. (2018). Thrown under the bus and outrunning it! The logic of Didi and taxi drivers' labour and activism in the on-demand economy. *New Media & Society*, *20*(8), 2691–2711. https://doi.org/10.1177/1461444817729149

City-Data. (2019). *30032 zip code (Candler-McAfee, Georgia) profile - Homes, apartments, schools, population, income, averages, housing, demographics, location, statistics, sex offenders, residents and real estate info*. Retrieved April 18, 2021, from https://www.city-data.com/zips/30032.html

Conger, K., Satariano, A., & Isaac, M. (2020, March 18). Pandemic erodes gig economy work. *The New York Times*. https://www.nytimes.com/2020/03/18/technology/gig-economy-pandemic.html

Cooke, M. (2015, October 31). The Bronx slave market. *Viewpoint Magazine*. https://viewpointmag.com/2015/10/31/the-bronx-slave-market-1950/ (Original work published 1950)

Data USA. (2018a). *Fremont, CA*. Retrieved April 18, 2021, from https://datausa.io/profile/geo/fremont-ca

Data USA. (2018b). *Park Slope, Carroll Gardens & Red Hook PUMA, NY*. Retrieved April 18, 2021, from https://datausa.io/profile/geo/park-slope-carroll-gardens-red-hook-puma-ny

D'Cruz, P., & Noronha E. (2016). Positives outweighing negatives: The experiences of Indian crowdsourced workers. *Work Organisation, Labour & Globalisation*, *10*(1), 44–63. https://doi.org/10.13169/workorgalaboglob.10.1.0044

Dickson, E. (2020, July 3). Coronavirus is killing the working mother. *Rolling Stone*. https://www.rollingstone.com/culture/culture-features/working-motherhood-covid-19-coronavirus-1023609/

Ehrenreich, B., & Hochschild, A. R. (2003). Introduction. In B. Ehrenreich & A. R. Hochschild (Eds.), *Global woman: Nannies, maids, and sex workers in the new economy* (pp. 1–14). Metropolitan Books.

England, P., Budig, M., & Folbre, N. (2002). Wages of virtue: The relative pay of care work. *Social Problems*, *49*(4), 455–473. https://doi.org/10.1525/sp.2002.49.4.455

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Farrell, M. B. (2014, August 14). Care.com, the big business of babysitting. *Globe Magazine*. https://www.bostonglobe.com/magazine/2014/08/14/care-com-big-business-babysitting/4Fjpf5q3YUSw3rMn9GraOM/story.html

Ferrari, F., & Graham, M. (2021). Fissures in algorithmic power: Platforms, code, and contestation. *Cultural Studies*, *35*(4–5), 814–832. https://doi.org/10.1080/09502386.2021.1895250

Flanagan, F. (2019). Theorising the gig economy and home-based service work. *Journal of Industrial Relations*, *61*(1), 57–78. https://doi.org/10.1177/0022185618800518

Gerson, E. S. (2019, July 8). *Background checks: What every family needs to know before hiring a caregiver*. Care.com. Retrieved April 22, 2021, from https://www.care.com/c/stories/3147/nanny-background-checks/

Glenn, E. N. (1992). From servitude to service work: Historical continuities in the racial division of paid reproductive labor. *Signs*, *18*(1), 1–43. https://doi.org/10.1086/494777

Graham, M., & Anwar, M. A. (2019). The global gig economy: Towards a planetary labour market? *First Monday*, *24*(4). https://doi.org/10.5210/fm.v24i4.9913

Gray, M. L., & Suri, S. (2019). *Ghost work: How to stop Silicon Valley from building a new global underclass*. Houghton Mifflin Harcourt.

Grose, J. (2021, February 4). America's mothers are in crisis. *The New York Times*. https://www.nytimes.com/2021/02/04/parenting/working-moms-mental-health-coronavirus.html

Hochschild, A. R. (2013). *The outsourced self: What happens when we pay others to live our lives for us*. Picador.

Hsu, A. (2020, September 29). *'This is too much': Working moms are reaching the breaking point during the pandemic*. NPR. https://www.npr.org/2020/09/29/918127776/this-is-too-much-working-moms-are-reaching-the-breaking-point-during-the-pandemi

Jaffe, S. (2021). *Work won't love you back: How devotion to our jobs keeps us exploited, exhausted, and alone*. Bold Type Books.

Kaine, S., Flanagan, F., & Ravenswood, K. (2020). Future of work (FoW) and gender. In A. Wilkinson & M. Barry (Eds.), *The future of work and employment* (pp. 119–138). Edward Elgar Publishing. https://doi.org/10.4337/9781786438256.00016

Macdonald, C. L. (2011). *Shadow mothers: Nannies, au pairs, and the micropolitics of mothering*. University of California Press. https://doi.org/10.1525/9780520947818

McDonald, P., Williams, P., & Mayes, R. (2021). Means of control in the organization of digitally intermediated care work. *Work, Employment and Society*, *35*(5), 872–890. https://doi.org/10.1177/0950017020969107

Montgomery, T., & Baglioni, S. (2021). Defining the gig economy: Platform capitalism and the reinvention of precarious work. *International Journal of Sociology and Social Policy*, *41*(9–10), 1012–1025. https://doi.org/10.1108/IJSSP-08-2020-0400

Newlands, G. (2021). Algorithmic surveillance in the gig economy: The organization of work through Lefebvrian conceived space. *Organization Studies*, *42*(5), 719–737. https://doi.org/10.1177/0170840620937900

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press. https://doi.org/10.2307/j.ctt1pwt9w5

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Ollier-Malaterre, A., Jacobs, J. A., & Rothbard, N. P. (2019). Technology, work, and family: Digital cultural capital and boundary management. *Annual Review of Sociology*, *45*(1), 425–447. https://doi.org/10.1146/annurev-soc-073018-022433

Parker, G. G., Alstyne, M. W. V., & Choudary, S. P. (2016). *Platform revolution: How networked markets are transforming the economy—and how to make them work for you*. W. W. Norton & Company.

Petrescu, P. (2014, October 1). *Google organic click-through rates in 2014*. Moz. https://moz.com/blog/google-organic-click-through-rates-in-2014

Poo, A.-J. (2017, June 16). *Out from the shadows: Domestic workers speak in the United States*. OpenDemocracy. https://www.opendemocracy.net/en/beyond-trafficking-and-slavery/out-from-shadows-domestic-workers-speak-in-united-states/

Rega, I., & Medrado, A. (2021). The stepping into visibility model: Reflecting on consequences of social media visibility – a Global South perspective. *Information, Communication & Society*. Advance online publication. https://doi.org/10.1080/1369118X.2021.1954228

Schor, J. B. (2017). Does the sharing economy increase inequality within the eighty percent?: Findings from a qualitative study of platform providers. *Cambridge Journal of Regions, Economy and Society*, *10*(2), 263–279. https://doi.org/10.1093/cjres/rsw047

Schreier, M. (2014). Qualitative content analysis. In U. Flick (Ed.), *The SAGE handbook of qualitative data analysis* (pp. 170–183). SAGE. https://doi.org/10.4135/9781446282243.n12

Shapiro, A. (2018). Between autonomy and control: Strategies of arbitrage in the "on-demand" economy. *New Media & Society*, *20*(8), 2954–2971. https://doi.org/10.1177/1461444817738236

Sharabi, L. L. (2021). Exploring how beliefs about algorithms shape (offline) success in online dating: A two-wave longitudinal investigation. *Communication Research*, *48*(7), 931–952. https://doi.org/10.1177/0093650219896936

Sundararajan, A. (2016). *The sharing economy: The end of employment and the rise of crowd-based capitalism*. MIT Press.

Tandon, A., & Rathi A. (2021). Care in the platform economy: Interrogating the digital organization of domestic work in India. In B. Dolber, M. Rodino-Colocino, C. Kumanyika, & T. Wolfson (Eds.), *The gig economy: Workers and media in the age of convergence* (pp. 47–57). Routledge. https://doi.org/10.4324/9781003140054-6

Thurman, N., Moeller, J., Helberger, N., & Trilling, D. (2019). My friends, editors, algorithms, and I: Examining audience attitudes to news selection. *Digital Journalism*, *7*(4), 447–469. https://doi.org/10.1080/21670811.2018.1493936

Ticona, J. (2020) Essential and untrusted. *Dissent*, *67*(4), 12–18. https://doi.org/10.1353/dss.2020.0080

Ticona, J., & Mateescu, A. (2018a). Trusted strangers: Carework platforms' cultural entrepreneurship in the on-demand economy. *New Media & Society*, *20*(11), 4384–4404. https://doi.org/10.1177/1461444818773727

Ticona, J., & Mateescu, A. (2018b, March 29). How domestic workers wager safety in the platform economy. *Fast Company*. https://www.fastcompany.com/40541050/how-domestic-workers-wager-safety-in-the-platform-economy

Ticona, J., Mateescu, A., & Rosenblat, A. (2018) *Beyond disruption: How tech shapes labor across domestic work & ridehailing*. Data & Society.

https://datasociety.net/wp-content/uploads/2018/06/Data_Society_Beyond_Disruption_FINAL.pdf

Vallas, S., & Schor, J. B. (2020). What do platforms do? Understanding the gig economy. *Annual Review of Sociology*, *46*(1), 273–294. https://doi.org/10.1146/annurev-soc-121919-054857

van Doorn, N. (2017). Platform labor: On the gendered and racialized exploitation of low-income service work in the 'on-demand' economy. *Information, Communication & Society*, *20*(6), 898–914. https://doi.org/10.1080/1369118X.2017.1294194

Walsh, S. (Director). (2021). *The gig is up* [Film]. Dogwoof Pictures.

Williams, P., McDonald, P., & Mayes, R. (2021). Recruitment in the gig economy: Attraction and selection on digital platforms. *The International Journal of Human Resource Management*, *32*(19), 4136–4162. https://doi.org/10.1080/09585192.2020.1867613

Wood, A. J., Graham, M., Lehdonvirta, V., & Hjorth, I. (2019). Good gig, bad big: Autonomy and algorithmic control in the global gig economy. *Work, Employment and Society*, *33*(1), 56–75. https://doi.org/10.1177/0950017018785616

Wood, A. J., & Lehdonvirta, V. (2021). *Platform precarity: Surviving algorithmic insecurity in the gig economy*. SSRN. https://doi.org/10.2139/ssrn.3795375

Wu, T. (2016). More than a paycheck: Nannies, work, and identity. *Citizenship Studies*, *20*(3–4), 295–310. https://doi.org/10.1080/13621025.2016.1158358

Zhou, M., Hertog, E., Kolpashnikova, K., & Kan, M.-Y. (2020). *Gender inequalities: Changes in income, time use and well-being before and during the UK COVID-19 lockdown*. SocArXiv. https://doi.org/10.31235/osf.io/u8ytc

JOURNAL OF DIGITAL SOCIAL RESEARCH

WWW.JDSR.IO ISSN: 2003-1998

VOL. 4, NO. 1, 2022, 73–97

#VACHINA: HOW POLITICIANS HELP TO SPREAD DISINFORMATION ABOUT COVID-19 VACCINES

Raquel Recuero[ab] and Felipe Bonow Soares[b]

ABSTRACT

This paper focuses on how Brazilian politicians helped to spread disinformation about Covid-19 vaccines, discussing legitimation strategies and actors that played a significant role on Twitter and Facebook. Based on data gathered through CrowdTangle and Twitter API, we selected the 250 most shared/retweeted posts for each dataset (n=500) and examined if they contained disinformation, who posted it, and what strategy was used to legitimize this discourse. Our findings indicate that politicians and hyperpartisan accounts have a key influence in validating the Brazilian president's populist discourse through rationalization (pseudo-science) and denunciation (against the vaccine). The political frame also plays an important role in disinformation messages.

Keywords: anti-vaccine; Covid-19; disinformation; social media; discourse.

[a] Federal University of Pelotas, Federal University of Rio Grande do Sul, Brazil.
[b] MIDIARS, Brazil.

1 INTRODUCTION

The Covid-19 pandemic was a catastrophic event around the world. However, some countries felt it harder than others. From the 5.4 million deaths the world experienced in two years, Brazil had, alone, over 600 thousand. The chaotic response to the disease was credited to the spread of disinformation, the populist discourse of the president, and uncoordinated actions from the country's Health Ministry[1]. For example, because the pandemic spread so widely in Brazil, the country was a laboratory for the development of many Covid-19 vaccines[2]. So, it was also given the chance to buy vaccines early on. However, the Brazilian's far-right president, Jair Bolsonaro, refused it[3], arguing Brazilians already had "the cure"[4] for Covid-19 (the usage of hydroxychloroquine, which was already a debunked claim by scientific experiments) (Hallal and Victora, 2021). This particular context also helped the spread of disinformation campaigns about the pandemic (Soares et al., 2021), similar to what was observed in other far-right countries (Palau, 2021). This disinformation phenomenon targeted many topics, including vaccines (Galhardi et al., 2020), which increased the hesitancy for some to get vaccinated (Bivar et al., 2021).

Because of these problems, Brazil did not acquire vaccines until the end of 2020 (Hallal and Victora, 2021). By then, there were two efforts to develop a vaccine available to Brazilians. One was made by João Doria, the governor of the state of Sao Paulo, that mobilized the Butantan institute (a biologic research facility located in Sao Paulo) to work with the Sinovac laboratory in China. The other was made by Fiocruz, another research institute, situated in the state of Rio de Janeiro, that was working with AstraZeneca for the same purpose. Doria is a political adversary of Bolsonaro, as he constantly challenged the conspiracy theories created by the president (Santos and Fossá, 2020). At the same time, Bolsonaro attacked the vaccines, claiming they were "Chinese"[5], "water with sugar"[6] and "didn't work"[7]. In fact, until now, Bolsonaro refuses to get vaccinated[8].

[1] https://www.newyorker.com/news/daily-comment/brazils-covid-19-crisis-and-jair-bolsonaros-presidential-chaos (Access on 09/02/2021)
[2] https://www.nytimes.com/2020/08/15/world/americas/brazil-coronavirus-vaccine.html (Accessed on 09/02/2021)
[3]https://www.nytimes.com/2021/01/18/world/americas/brazil-covid-variants-vaccinations.html (Accessed on 09/02/2021)
[4] https://www.nytimes.com/2020/06/13/world/americas/virus-brazil-bolsonaro-chloroquine.html (Accessed on 09/02/2021)
[5] https://www.bloomberg.com/news/articles/2020-10-21/bolsonaro-slams-chinese-vaccine-his-government-said-it-would-buy (Accessed on 09/02/2021)
[6] https://www.washingtonpost.com/world/2021/04/15/brazil-china-vaccine-coronavirus-coronavac/ (Accessed on 09/02/2021)
[7] https://www.theguardian.com/world/video/2021/jan/15/brazils-president-casts-doubts-on-covid-vaccine-as-second-wave-hits-video (Accessed on 09/02/2021)
[8] https://www.washingtonpost.com/world/2021/09/19/jair-bolsonaro-unga-vaccine-honor-system/ (Accessed on 09/02/2021)

Bolsonaro, like many other politicians, relies on social media channels to mobilize their supporters and to talk to his audience (Silva, 2020), as he distrusts traditional media outlets as "fake news". Research has pointed this is a strategy used by populist leaders, as social media platforms offer affordances that are particularly important for this field, such as the possibility to influence more people (Cesarino, 2020), and has been used in western democracies for a long time (Boulianne, Koc-Michalska and Bimber, 2020). Some of these leaders, like Bolsonaro, have insufflated disinformation, as they use social media to antagonize real content (Soares et al., 2021). Thus, it is important to understand how, in political contexts such as this, disinformation discourse is legitimated and by whom. The Brazilian situation provides a unique opportunity to analyze how discursive strategies are mobilized to spread disinformation about Covid-19 vaccines. Although we look at one country, the issue we explore in this paper is relevant in a global sense. In the last few years, far-right populist politicians emerged in many countries (Wodak, 2015; Mudde and Kaltwasser, 2017; Rooduijn, 2019), as well as health-related and scientific disinformation turned out to be a key issue during the Covid-19 pandemic worldwide (Araujo and Oliveira, 2020; Brennen et al., 2020; Casarões and Magalhães, 2021). The intersection between discourse, populism, and disinformation that we explore in this research represents a real challenge for countries around the world – in particular, those that face a context like Brazil with populist politicians in positions of leadership.

The context we analyze exemplifies why studying the disinformation discourse, its legitimation process, and connections to populism are so important, as political leaders support disinformation campaigns that may harm the population. Our study explores a cross-platform context (Twitter and Facebook) based on the following research questions: (1) Who are the key users for disinformation spread? (2) How did they frame the disinformation about vaccines to gain legitimation? And (3) What are the connections of disinformation discourse to populist discourse?

2 DISINFORMATION DISCOURSE, COVID-19 AND SOCIAL MEDIA PLATFORMS

Disinformation has been a major problem during the Covid-19 pandemic, particularly because of how social media affordances end up fueling its spread. We define disinformation as the phenomenon based on the spread of false and misleading content to influence public discourse (Benkler, Faris and Roberts, 2018). This disinformation discourse is often mobilized through fabricated content, false claims mixed with truthful content, or false connections between truthful statements (Derakhshan and Wardle, 2017). Some authors have also pointed certain characteristics of disinformation discourse, such as “call to action” or a sense of “urgency to spread”, and the usage of discursive strategies to gain legitimation (Recuero, Soares and Vinhas, 2021). Disinformation discourse has been used for

political purposes, as public authorities may engage in sharing them for political gain (which was seen in Brazil by Ricard and Medeiros, 2020; Alcantara and Ferreira, 2020) or for the manipulation of public opinion in their favor (Rogers and Niederer, 2020).

The fast spread of disinformation on social media may be connected to how people legitimate these discourses online. Disinformation often circulates due to a very engaged audience (Vosoughi, Roy and Aral, 2018). Although most of this content may circulate among ordinary accounts, influencers such as politicians, celebrities and users with a great audience may increase its spread (Brennen et al., 2020). In fact, some authors found that politicians, celebrities, organizations, and other influencers played an important role in the circulation of disinformation about Covid-19 on Twitter (Recuero, Soares and Zago, 2021; Shahi, Dirkson and Majchrzak, 2021).

In this context, different studies have shown an alignment between populist discourse and disinformation about Covid-19, particularly from the far-right (Wondreys and Mudde, 2020; Stecula and Pickup, 2021). This means that characteristics of the populist discourse have been found in disinformation content. The contemporary right-wing populism has been studied by several authors, and particularly, by Wodak (2015). This form of populism, represented by leaders like Trump, Marie LePen and others, have a strong basis on the affordances and self-mediation allowed by social media (Wodak and Kryzanowski, 2017). The authors also explain that this type of populism relies on traditional media support through reports of scandals, as they often lack traditional parties' structures to promote them. When exploring these scandals, contemporary right-wing populists reinforce their anti-establishment rhetorics and survive on social media as an alternative space (Wodak, 2015). This populist discourse is based on elements such as the opposition to elites, the anti-science discourse, and the idea that people who align with populists are "virtuous people" that fight against the corruption of societal values (Roudjin, 2019; Mede and Schäfer, 2020; Mudde and Kaltwasser, 2012).

Jair Bolsonaro, current Brazilian president, is considered a far-right populist (Mendonça and Caetano, 2020; Watmough, 2021) with anti-science views (Oliveira, 2020). Besides, Bolsonaro is often linked to disinformation spread, particularly in the context of the Covid-19 pandemic (Araujo and Oliveira, 2020; Casarões and Magalhães, 2021; Ricard and Medeiros, 2020). In this context, disinformation in Brazil reproduced both political populism (Roudjin, 2019; Mendonça and Caetano, 2020) and science-related populism (Mede and Schäfer, 2020), which are important for the Covid-19 context. The first one was used to criticize political elites and the mainstream media who stand in Bolsonaro's way and to reinforce his "macho" image (Watmough, 2021), as he claimed the virus would not concern him given his "athletic history". The second one was used to criticize WHO and other health authorities, to reproduce anti-vaccine discourses and to promote unproven drugs (Araujo and Oliveira, 2020). Finally, like many other populist leaders, Bolsonaro also relies on social media for support and

disinformation on these platforms to influence public opinion in his favor (Soares et al., 2021; Kallil, 2019).

3 DISCOURSE AND LEGITIMATION

One key point to understand how disinformation circulates in contexts such as the Covid-19 pandemic is to find out how this type of content gains legitimation. The disinformation, like other types of discourse, often relies on strategies that articulate credibility to produce effects on society. Legitimation is a key concept for Critical Discourse Analysis (CDA), a perspective that focuses on how discourse builds and legitimizes power relations in society (Fairclough, 2001; Van Dijk, 2006, 2009). Legitimation strategies, in this sense, are tools used through discourse to build credibility. While often seen in a positive light (Vaara and Tienar, 2008), legitimation is a critical process that has effects in society as it helps naturalize the power (and dominant) relations built through discourse.

The study of legitimation strategies focuses on the linguistic characteristics of the discourse. These strategies have been studied by several authors. Van Leeuwen (2007), for example, explains four major categories of legitimation, which are: (1) Authorization, which is the legitimation through authority; (2) Moral Evaluation, the legitimation through the reference to moral values; (3) Rationalization, which refers to the cognitive validity, the social knowledge, and other rational arguments; (4) Mythopoesis, which is the legitimation through narratives and stories, which build upon legitimate and non-legitimate actions. These categories can be used to legitimate or de-legitimate a topic. Reyes (2011), on the other hand, analyzes the strategies specifically used by politicians during their public discourse, which he classifies as: (1) the usage of emotions (and particularly fear). Legitimation through emotions is built upon skewing the involved actors to the audience. Politicians often represent their opponents with negative attributes, creating a narrative of "us" versus "them", the others. (2) The construct of a hypothetical future. In this case, the discourse predicts an imminent threat that requires action to preserve the future. (3) Rationality, when the discourse is based on "rational" arguments, logic. (4) Voices of the experts, a strategy directly connected to "authorization" (Van Leeuwen, 2007), where the legitimation comes from the expertise, the authority that legitimates de discourse. (5) Altruism, when the legitimation is proposed through an altruistic, not driven by personal interest's discourse. While Van Leeuwen's (2007) typology focuses on general strategies, Reyes' (2011) work, however, focuses on the discourse produced by politicians to convince their audience. These two classifications provide some light on the discussion of how discourse gains legitimation and thus, how it can be used to manipulate and influence social practices that are not beneficial to people.

In a broader political context, other specific discursive strategies are used to build legitimation. One relevant strategy is the use of humor. Humor is key to the rhetoric of political memes, as humor contributes to the consolidation of shared

meanings on social media (Chagas et al., 2019). In this context, Chagas et al. (2019) argue that the trivialization of socially relevant topics is often a way to solve problems of adequacy in political discourse – consequently, engaging more actors and fueling political participation. In the context of disinformation spread, Crilley and Chatterje-Doody (2020) identified that humor was a fundamental legitimation strategy used by RT (Russia Today), especially in content related to Russian foreign policy. In addition, during the Covid-19 pandemic, Basch et al. (2021) identified that memes and parodies were used to spread anti-vaccine content on TikTok.

Social media platforms, in this context, offer an interesting setting to study how discourse is articulated, particularly because they play an important role in disinformation spread (Allcott and Gentzkow, 2017; Vosoughi, Roy and Aral, 2018). Also, the interaction tools provided by these platforms, such as "likes" or "retweets", for example, may provide insights on how much of an impact each post has on a particular community, and are also often seen as a form to legitimate discourses (Metaxas et al. 2018; Glozer, Caruana and Hibbert, 2019).

Several studies have discussed how disinformation is legitimized on social media platforms based on these ideas. Igwebuike and Chimuanya (2021), for example, studied which were the recurrent strategies used by Nigerian "fake news" producers on social media. They discussed how Van Leeuwen's (2007) strategies were used, suggesting that authorization and appeal to emotion, using moralization and rationalization were key for the phenomenon. Similarly, Recuero, Soares and Vinhas (2021), discussing legitimation strategies for disinformation on WhatsApp and Twitter, found that they were mostly based on emotional framing and incentives to share, but also that the most used strategies (again, based on Van Leeuwen, 2007) differ on each platform. While Twitter had more authorization and moral evaluation, WhatsApp messages used mostly mythopoesis. These works show that the disinformation discourse on social media platforms relies on different forms of legitimation. However, while they point to the disinformation strategies used, they do not explore how these strategies are mobilized and by whom. Finally, it is important to understand how discourse is articulated to gain legitimation to disinformation about Covid-19 in different contexts and, particularly, among the ones where governments favor disinformation, like Brazil.

4 METHODS

In this paper, we focus on discussing the strategies used to legitimate disinformation about the Covid-19 vaccines and the usage of political influence to circulate them in Brazil. We focus on Twitter and Facebook to explore three research questions: (1) Who are the key users for disinformation spread? (2) How did they frame the disinformation about vaccines to gain legitimation? And (3) What are the connections of disinformation discourse to populist discourse?

To explore these questions, we collected tweets through Twitter API (n=890,501) and Facebook public posts through CrowdTangle (n=111,807)

containing the expression "Vacina" + "chinesa" or + "China" (Chinese or China plus Vaccine, in Portuguese) between July 2020 and April 2021. These keywords were selected to focus on a particular public debate that was often fueled by disinformation. The expression "Chinese vaccine" was heavily used by actors spreading disinformation, but it was also by other users and the Brazilian mainstream media to refer to the Sinovac vaccine – similarly to how they used "Oxford Vaccine" to refer to the AstraZeneca vaccine. Sinovac vaccine (which was produced in a collaboration with the Brazilian Butantan Institute, as we explained) was the first and the most widely available vaccine in Brazil (and most of South America) for a long time, differently from the countries of the north global. Disinformation narratives often focused on the role of China in the pandemic, emphasizing that China is a Communist country, and reproducing the "Chinese virus" narrative and conspiracy theories that the virus was intentionally created by China. Because of this connection between politics and health (the vaccine), we believe this case provides important insights to understand how political discourse influenced disinformation about the pandemic.

From these original datasets, we selected the 250 most retweeted tweets and the 250 most shared Facebook posts to analyze. Although limited in number, this sample of 250 tweets represented 45.5% of the total retweets and the 250 Facebook messages accounted for 50.4% of the total shares on Facebook. The creation of a sample is a limitation, as we do not look at less prevalent tweets and Facebook posts, but our sample provides fair representativeness of our dataset, as it accounts for over 400,000 RT and almost 1.5 million FB shares. Besides, we chose the most shared/retweeted messages because, as we explained in the previous section, retweets and shares can indicate discourses that received more attention and were perceived as more legit by the audience (Metaxas et al. 2018; Glozer, Caruana and Hibbert, 2019).

We used Content Analysis (Krippendorff, 2012) to identify disinformation and to explore the data. At first, we visited the messages and the profiles qualitatively to think about the categories to code the messages. We then created the coding framework. After that, two authors independently classified all the messages from the dataset. That is, each message of the dataset was double-coded (independently). We classified the messages based on (1) the presence of disinformation; (2) the type of account that posted the message, and (3) the discursive strategy.

In the first step, we visited each tweet and each Facebook post, examined the content, and searched for fact-checking about it. Then we classified each message regarding the presence of disinformation. Our second step was to classify the type of account that posted the message. To discuss the type of account, we created the following categories: (1) politicians, (2) activists, (3) media; (4) hyperpartisan media, (5) celebrities, and (6) others. These categories were created based on how the account identifies itself (see Table 1).

Table 1. Categories for accounts

Category	Characteristics
Politicians	Politically affiliated users. Example: political candidates, congressmen, president, ministers, etc.
Celebrities	Famous users not politically affiliated. Example: actors, humorists, artists etc.
Media	Traditional media: News outlets, journalists etc.
Hyperpartisan media	Accounts that claimed to be media outlets or to share "news" but were clearly politically affiliated.
Activists	Users and organizations with a purpose of political activism. Ex.: political parties.
Others	Ordinary accounts and users that did not fit in any of the previous categories.

Finally, we classified the discourse strategy that framed the message. This classification was based on the following categories: (1) rationalization, (2) humor, (3) opinion, (4) denunciation, and (5) other (see Table 2). These categories emerged from the original analysis of the data collected through Van Leeuwen's (2007) and Reyes' (2011) strategies. We first examined the dataset and discussed how the data was legitimized, creating the categories that were further tested through independent coding.

Table 2. Categories for Legitimation Strategies

Type of strategy	Characteristics
Rationalization	Rational argument, logical, based on theories or facts to legitimate the discourse.
Humor	Legitimation through a joke, irony, or other forms of humor.
Opinion	Opinion as the strategy to legitimate the discourse.
News	Neutral tone, presented as "news".
Denunciation	Denounces something.
Other	The message did not present any of the categories. For example: surveys.

We calculated Cohen's Kappa to test the reliability of our independent classification of the three categories in the 500 messages (Freelon, 2010). Table 3 provides a

breakdown of the coding reliability. The Kappa is considered high for most of the classifications, and substantial for the strategies, which was expected as this was a more subjective evaluation. We discussed the classifications that we did not agree on in the independent coding to reach a final classification for those messages.

Table 3. Categories for Legitimation Strategies

Category	Cohen's Kappa	
	Twitter	Facebook
Type of account	0.89	0.87
Disinformation	0.76	0.75
Legitimation strategies	0.65	0.72

Finally, we also examined the messages through Connected Concepts Analysis (Lindgren, 2016), focusing on disinformation and non-disinformation discourses to better understand their content. CCA is a method that uses content analysis as the basis. The messages are initially examined in terms of the frequency of words. To make this analysis more meaningful, similar words are classified into concepts (for example: president and Bolsonaro were put in the same category). Therefore, in the first step of CCA, we can identify the most prevalent concepts for each group of messages (disinformation and non-disinformation). This step is relevant because it provides clues to interpret the discourse of a group based on the main topics and concepts. Further on, we use CCA to identify the co-occurrences of these concepts. In this step, we understand how the topics are framed by each group based on the connections between concepts (for example, a strong connection between "China" and "Communism" indicates a particular way to frame the discussion). We created network graphs to better visualize the co-occurrences of concepts and explore to which discourse they are associated with. These graphs were created using Social Network Analysis metrics (Wasserman and Faust, 1994), to represent the most central concepts (biggest nodes) and their co-occurrences (through connections), as well as their tendency to appear together (color). In summary, CCA is a useful method to make sense of a large corpus by identifying key concepts and discursive frames based on quantitative elements (frequency and co-occurrences) and visual resources (network graph). Therefore, based on CCA, we can explore how each group of messages generally framed the discussion about vaccines on Twitter and Facebook.

5 ANALYSIS

In this section, we present our results and further discuss them. For this discussion, we compare the strategies used by disinformation and "information" content, to understand how disinformation is different for each platform.

5.1 Twitter

Most of the tweets in our sample contained disinformation. While tweets that shared "information" were more retweeted on average, the median of RT is higher for disinformation tweets. This indicates that some tweets that challenged disinformation were heavily shared but were outliers in our dataset. On the other hand, the similarities between the mean and median RT for disinformation tweets indicate that this category had a larger number of accounts with high visibility, possibly motivated by a very engaged audience, as pointed by other authors (Vosoughi, Roy and Aral, 2018). Table 4 summarizes the data.

Table 4. Breakdown of tweets in our sample per type of account

Type of account	**Disinformation**			**Information**		
	Tweets	Mean RT	Median RT	Tweets	Mean RT	Median RT
Politicians	31 (17.7%)	1468	917	11 (13.3%)	1134	790
Celebrities	32 (18.2%)	1172	802.5	12 (16%)	1956	570
Traditional Media	0	-	-	17 (22.6%)	2019	754
Hyper-partisan media	76 (43.3%)	1152	855	6 (8%)	1198	662
Activists	2 (1.1%)	701	701.5	1 (1.3%)	722	722
Others	33 (18.9%)	1098	796	29 (38.6%)	4208	1173
Total	175 (70%)	1197	825.5	75 (30%)	2623	794

Note: The percentage for the categories was separated based on the presence of disinformation (i.e., 17.7% of the tweets containing disinformation were posted by politicians). The percentage for the total considered the entirety of our sample (70% of the tweets in our sample contained disinformation).

Hyperpartisan media (highest number of tweets) and politicians (highest mean and median of retweets) played a key role in spreading disinformation. Celebrities also account for many tweets and received a high average of RT. The high number of RT of tweets posted by politicians and celebrities is likely caused by their well-stablished audience, as Shahi, Dirkson and Majchrzak, (2021) argued.

Tweets without disinformation were mostly shared by "others". While this may seem odd, it happened because these were usually viral tweets from common users, which explains why their mean and median are also the highest. Traditional media, celebrities and politicians played important roles in challenging disinformation.

We also identified differences in the legitimation strategies. Table 5 provides a breakdown of how these strategies were used in each group of tweets. While

disinformation relied mostly on denunciation and rationalization, tweets without disinformation mostly used humor as a legitimation strategy.

Table 5. Legitimation strategies on Twitter

Type of account	Disinformation			Information		
	Tweets	Mean RT	Median RT	Tweets	Mean RT	Median RT
Rationalization	54 (30.8%)	1667	853	14 (18.6%)	1576	788.5
Humor	16 (9.1%)	939	661.5	32 (42.6%)	3854	1614.5
Opinion	24 (13.7%)	1042	845.5	6 (8%)	713	677
News	4 (2.2%)	723	658.5	11 (14.6%)	1274	622
Denunciation	76 (44%)	990	830	8 (10.6%)	777	570
Other	-	-	-	4 (5.3%)	1262	1029.5

Note: The percentage for the categories was separated based on the presence of disinformation (i.e., 30.8% of the tweets containing disinformation used rationalization).

Disinformation messages that relied on denunciation as a strategy often denounced the vaccine as something untested and unreliable since it "came from China, who also created the virus", putting the origin of the vaccine as a relevant and negative matter. These messages also urged the audience to share this content to inform "everyone else". Tweets that used rationalization often relied on a pseudo-scientific argument, connected to conspiracy theories (for example, citing doctors who defended the use of ivermectin and hydroxychloroquine as the cure for covid, so that vaccines were not necessary). Many of these tweets reproduced Bolsonaro's discourse about vaccines. While denunciation was the most common strategy used in disinformation tweets, rationalization was the most effective in terms of retweets. This data seems to suggest that while denunciation may raise awareness, it was the pseudo-scientific rationalization that got more endorsements on Twitter. The relevance of rationalization might be related to the affordances of Twitter as a platform that provides proper space for social or macro-level discussion, which favors the use of "logical" arguments (Recuero, Soares and Vinhas, 2021).

The use of opinion was also a relevant strategy for disinformation, as it was the third most frequent category and received the second-highest mean and median of retweets. This strategy was often linked to celebrities and politicians, who generally have many followers and thus, receive more retweets (Brennen et al., 2020). Hyperpartisan media accounts also contributed to this category by reproducing others' opinions. Opinion tweets would reproduce some falsehood claiming that it was "the opinion" of the user (this was often connected to the decision to vaccinate with the "Chinese vaccine" or "Doria's vaccine" or not).

Of the tweets with "information", the most common category was humor, which was also the category with the highest mean and median of retweets. Humor was the most frequent legitimation used by ordinary users' tweets that went viral

("others"), although traditional media and celebrities also reproduced some of this strategy. These tweets relied on humor to challenge disinformation and seem to succeed. Many tweets made fun of people discussing the vaccine origin as a reason to not get vaccinated, for example "'I don't trust the Chinese vaccine'- My dear, you trust MEN", or "'Do you trust the Chinese vaccine?'- I've eaten too much highway gas station food to have this type of frills". These strategies might be effective because they use trivialization and humor to engage other users that shared similar beliefs (Chagas et al., 2019).

The second most common category was rationalization, a strategy that also received the second-highest mean of retweets. Rationalization was used to promote scientific and technical arguments to defend the vaccines. This strategy was employed mostly by politicians, but there were also viral tweets from ordinary users and some tweets from traditional media. News was mostly mobilized by the mainstream media and turned out to be a relevant strategy to challenge disinformation.

This data shows some interesting information. While humor is used as a legitimation strategy both by information and disinformation tweets, it is among real content that this approach thrives. On the other hand, opinion and denunciation had a stronger presence on the disinformation group, suggesting they are more relevant strategies for disinformation discourse. Finally, rationalization is a category that appears on both groups and seems to be equally legitimated by both, although it had a stronger impact on disinformation.

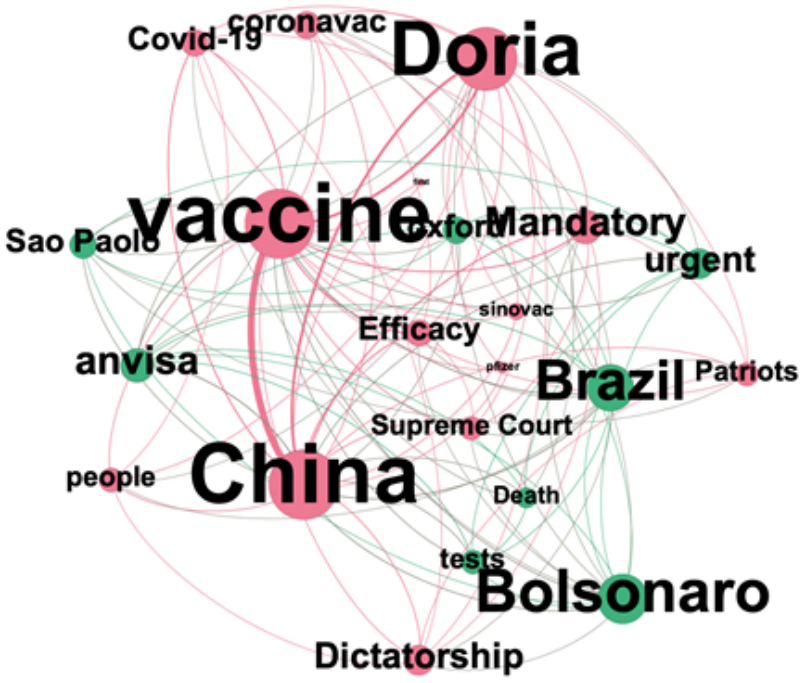


Figure 1. Connected component analysis for disinformation tweets.

To better understand the context of these tweets, we use Connected Concept Analysis to explore the main themes in the tweets to better understand the use of legitimation strategies. Figure 1 provides a systematization of these connections in the tweets containing disinformation.

When looking into these themes, we first see that the vaccine is strongly associated with Doria, Sao Paulo governor, and China, which was expected. They are also connected to concepts such as "dictatorship" and "mandatory", indicating a negative frame in disinformation tweets. Many of these tweets claimed that Doria would make vaccination mandatory in Brazil (which was false) and only "Bolsonaro" could stop this. These connections are mainly associated with tweets that relied on opinion and denunciation.

We also found a strong connection between "China" and the "efficacy" of the "vaccine", which was associated with rationalization and denunciation. In this case, we found conspiracies associating China as the "creator" of the virus to sell the "ineffective" "vaccine". Associations between "China" and the low quality of the vaccine "coronavac" and "sinovac", as well as comparisons to "Oxford" (the AstraZeneca vaccine) and "Pfizer" were also strong in this dataset. These tweets would present the origin of the vaccine as an indication of its quality.

Another important concept was "patriots", a word often used by the far right and Bolsonaro's supporters, to refer to his nationalist anti-elite discourse (Mendonça and Caetano, 2020; Watmough, 2021). In this case, we found that most tweets connected to an anti-vaccination discourse, connecting to concepts such as "individual rights". These tweets underline the far-right rhetoric present on the disinformation dataset, as many of these actors recognize themselves as "the good people" that fight for the president against the conspiracies and the "corrupted" institutions such as the Brazilian "Supreme Court". In this sense, both opinion and denunciation rely on nationalist arguments, supporting Bolsonaro and portraying China as untrustworthy. Many tweets reinforce Bolsonaro's image and rely on populist anti-establishment rhetoric (Wodak, 2015; Mudde and Kaltwasser, 2017; Rooduijn, 2019) to criticize democratic institutions such as the Brazilian agency who authorized the vaccines to be used (ANVISA) and the Supreme Court (both amongst the most central themes in disinformation tweets).

When we look at the content of the most shared tweets with real content, there is a different picture (see Figure 2). In this dataset concepts such as "American", "English", "Russian", "Oxford" and "China" are often mentioned. These concepts come from different tweets that made fun of people who wanted to choose the vaccine based on its origin, claiming that the origin did not matter. Although these tweets directly challenged the disinformation content, they also focused on the origin of the vaccine as something important. Other tweets ridicule one of the pseudo-scientific treatments defended by the anti-vaccine discourse, "ozone" therapy. There are also concepts connected to the "trust" in vaccines and referring to the Bolsonaro versus Doria political fight.

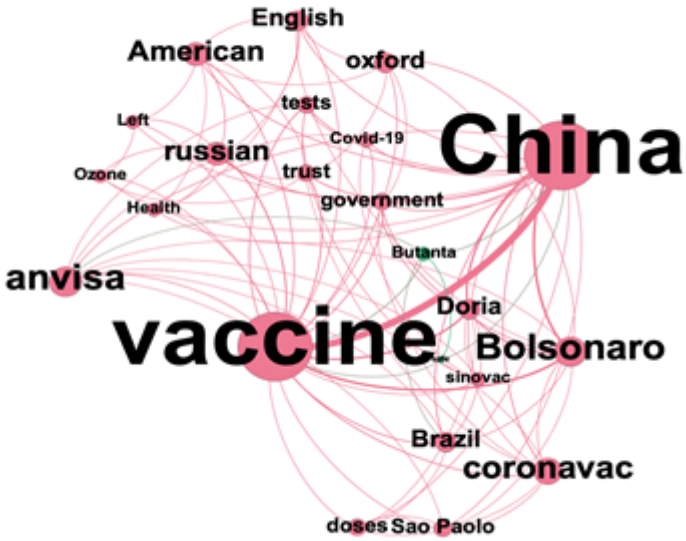


Figure 2. Connected concept analysis for tweets without disinformation.

5.2 Facebook

We now focus on the sample of Facebook posts. While the Facebook dataset had more equally distributed types of content, disinformation had a higher mean and median of shares. Table 6 summarizes the data.

Table 6. Breakdown of Facebook posts in our sample per type of account

Type of account	Disinformation			Information		
	Posts	Mean shares	Median shares	Posts	Mean shares	Median shares
Politicians	72 (56.2%)	7816.3	3605.5	62 (50.8%)	4561	3479
Celebrities	6 (4.6%)	8216	2670	9 (7.3%)	7281	7067
Traditional Media	0	-	-	20 (16.3%)	5071	2644.5
Hyperpartisan media	30 (23.4%)	5531.9	4439.5	17 (13.9%)	8351	4010
Activists	15 (11.7%)	4841.4	3149	7 (5.7%)	3041	2887
Others	5 (3.9%)	3746.6	2682	6 (4.9%)	9225	2548
Total	128 (51%)	6793	3660	122 (49%)	5554	3372

Note: The percentage for the categories was separated based on the presence of disinformation (i.e., 56.2% of the posts containing disinformation were posted by politicians). The percentage for the total considered the entirety of our sample (51% of the posts in our sample contained disinformation).

When we examined the types of accounts that shared disinformation, we found that most of them were politicians, all of them clearly stating their support for Bolsonaro on their pages. Similarly, the hyperpartisan media category comprised mostly far-right outlets and activists accounts were usually groups/pages focused on Bolsonaro or the far-right (for example: "[City's] Conservative Right"). In terms of shares, celebrities and politicians had the two highest averages, which indicates how much

disinformation is influenced by users with a large audience on Facebook. Hyperpartisan media had the highest median, indicating that this category had the more equally engaged audience.

On the "information group", politicians were also the most present category. We also found in this category three pages of Bolsonaro's supporters. In this case, the pages shared real content when it was framed according to their political agenda (for example, the news that ANVISA, the Brazilian regulatory agency for vaccines had halted the tests for Coronavac because there was a death among volunteers. However, it was omitted that the volunteer death had occurred by suicide). While the information was not false, it was used to hurt the vaccine's credibility. Among hyperpartisan accounts that shared real content, we also identified five that worked in the same light as the pro-Bolsonaro's politicians accounts.

This data suggests that politicians and hyperpartisan accounts are key for both disinformation and information spread, which indicates that the debate around vaccines was politically framed on Facebook (similar to the findings of other studies – see Recuero, Soares and Zago, 2021; Soares et al., 2021; Araujo and Oliveira, 2020). Although in small numbers, celebrities also played an important role in terms of shares (the highest average for disinformation and the highest median for information). These categories underly the importance of pages from individuals with high visibility to spread content about vaccines.

Table 7. Legitimation strategies on Facebook

Type of account	Disinformation			Information		
	Tweets	Mean RT	Median RT	Tweets	Mean RT	Median RT
Rationalization	40 (31.2%)	8030	5084.5	17 (15.7%)	6598	3768
Humor	7 (5.7%)	20126	2664	16 (12.5%)	8119	4584
Opinion	26 (20.3%)	6633	3452.5	13 (10.6%)	6633	3714
News	11 (8.5%)	4496	7305	26 (21.3%)	4496	2691.5
Denunciation	44 (34.3%)	4634	3174.5	50 (40.9%)	4650	2885
Other	0	-	-	0	-	-

Note: The percentage for the categories was separated based on the presence of disinformation (i.e., 31.2% of the posts containing disinformation used rationalization).

Rationalization and denunciation were the most used strategy in the disinformation group on Facebook. Usually, these posts articulate conspiracies against the president and focus on the Chinese origin of the vaccine and the pandemic. Rationalization often relied on pseud-scientific arguments. These posts articulate scientific-related populism to criticize pharmaceutic companies and health organizations (Oliveira, 2020), suggesting they are involved in a conspiracy.

Denunciation posts claimed that far-right politicians were right and the conspiracy was happening to make the “Chinese” vaccine mandatory. Populist rhetoric was present for this sample as posts claimed users to "protest" against the scientific and political elites that oppose Bolsonaro, promoted distrust in the vaccines and xenophobic discourse about China.

On Facebook, information posts heavily relied on denunciation. Most of these were posted by politicians to criticize Bolsonaro and other far-right politicians for vilifying Coronavac. Posts containing news, mostly from traditional media, were also prevalent in this group. Humor posts making fun of anti-vaccine discourse received many shares, which indicates that this strategy was to consolidate shared meanings and challenge anti-vaccine discourse on social media (Chagas et al., 2019).

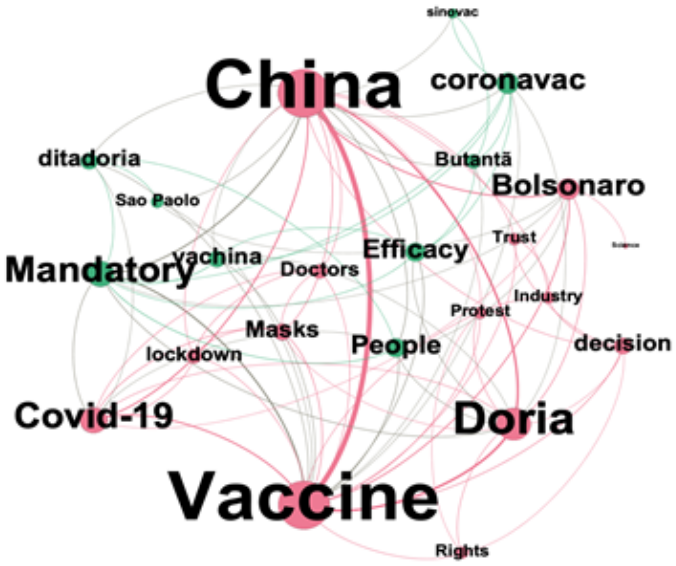


Figure 3. Connected Concepts analysis for Disinformation on Facebook.

Using CCA, we were able to understand more about the context of the most shared posts. Most disinformation posts associated the vaccine with China and conspiracies based on its origin (see Figure 3). They also claimed that the vaccine would be made mandatory by Doria (Sao Paulo governor), using words such as "ditadoria" (Doria dictator) and "vachina" (vaccine + China), frequently relying on denunciation. We observed that much of the effort against vaccination used denunciation to fuel political polarization (Bolsonaro anti-vaccine views versus Doria efforts to produce a vaccine) and rationalization to focus on the origin of the vaccine to frame it as ineffective. Facebook posts also questioned the usage of masks and the lockdowns measures imposed by several Brazilian governors to avoid the growth of the pandemic. This political framing of the combat measurements and the vaccines appear to take a turn to favor Bolsonaro’s views against democratic institutions (such as the courts, the state governors, the congress and many others),

reinforcing populist rhetoric (Wodak, 2015). Besides, many posts used populist rhetoric to claim individual "rights" not to get vaccinated, often based on denounces of conspiracies (Globalism and China creating the virus) and rational pseud-scientific discourse (often citing "scientists" who claimed the vaccines were not necessary). In these cases, the populist discourse was used to support conspiracy theories against vaccines and question scientific discourse (Roudjin, 2019; Mede and Schäfer, 2020; Mudde and Kaltwasser, 2012).

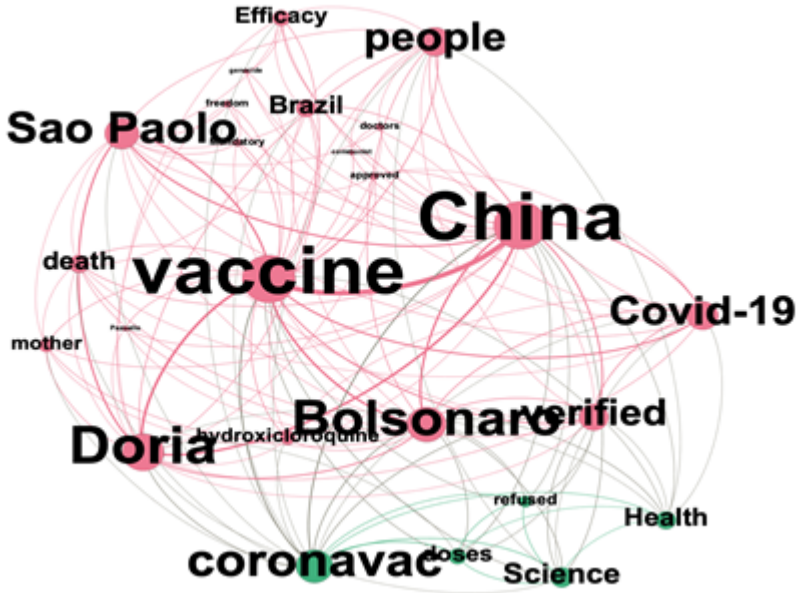


Figure 4. Connected Concepts analysis for Information on Facebook.

The discussion of posts without disinformation was mostly centered on the political struggle between Sao Paulo Governor Doria and Bolsonaro (Figure 4). Part of the content relied on denunciation strategies to accuse Bolsonaro of refusing to buy doses of the Coronavac vaccine. Other posts accused Bolsonaro of conspiracy and genocide, and some pointed that his mother was vaccinated with the same vaccine he refused to buy for the people, topics also strongly related to denunciation strategy. There was also an association between vaccine, China, hydroxychloroquine, and Communism that was often present in humor posts that questioned these relations created by disinformation. Consequently, humor was a strategy to criticize the populist discourse and the conspiracy theories associated with the Covid-19 vaccination in Brazil. The focus on the "deaths" and words like "refused" connected to Bolsonaro discuss also how he refused to buy the vaccines beforehand, an argument that was related to denounces against Bolsonaro and rationalization to reinforce scientific discourse.

6 DISCUSSION

In this paper, we focus on three research questions. We will organize the discussion section based on exploring these questions based on our findings. Regarding the first one, "Who are the key users for disinformation spread?", we found out that politicians, celebrities and hyperpartisan accounts were very important to the spread of disinformation on Facebook and Twitter.

Politicians were key for disinformation spread on Facebook, while hyperpartisan media was the most central category on Twitter. Their presence underlines the politically framed discursion about the vaccine and the support of the far-right for disinformation discourse in Brazil. Although less central in terms of number of tweets and posts, celebrities were also relevant to spread disinformation in terms of number of retweets and shares received. The relevance of politicians and celebrities in retweets and shares highlights the influence of prominent public figures in the top-down spread of disinformation about vaccines – a result that is in line with Brennen et al. (2020). As for actors challenging disinformation, traditional media was important on both platforms, while others (mostly common users) were central on Twitter and politicians were important on Facebook. This also points to the relevance of the affordances of each platform. The presence of viral tweets from ordinary users in our sample is due to Twitter's more public nature, where messages are more easily spread because they are not locked into groups or pages. Facebook, on the other hand, had more shares on tweets from celebrities' pages and groups, probably because these already have a large audience.

Our second research question was about "how did these accounts frame the disinformation about vaccines to gain legitimation?". We observed that **denunciation and rationalization were the strategies most frequently used for disinformation**. Denunciation messages usually referred to conspiracies about China and the vaccines, framing these through the political fight between Governor Doria and president Bolsonaro. Denunciation was also used to alert against the mandatory vaccination that would be determined by Doria (which was not true), calling people to protest “the dictator”. Denunciation uses emotion to engage other users by appealing to a "possible" future where people would be coerced by the government to take the "Chinese" vaccine (a key discursive strategy identified by Reyes, 2011). Our findings indicate that **denunciation may be a key form to obtain legitimation from far-right populists’ discourses in the pandemic context**, where “the virtuous” need to fight against the corruption of the political elites and “globalism” in particular (Roudjin, 2019; Mede and Schäfer, 2020; Mudde and Kaltwasser, 2012).

Rationalization, while also referring to the same conspiracy theories, used a logical strategy to reinforce pseud-scientific claims. **Rationalization suggests that anti-science populism was also very important in this context because of how these messages framed companies producing the vaccines and health authorities as**

untrustworthy (Mede and Schäfer, 2020; Oliveira, 2020). Although it accounted for fewer messages than denunciation, rationalization was key in terms of retweets and shares, which indicates that other users are more likely to endorse posts containing logical arguments to promote anti-vaccine discourse.

Another important point in this discussion is that some politicians and hyperpartisan accounts shared real content. However, **this content was framed to support disinformation**. For example, headlines (often from the mainstream media) that reproduced Bolsonaro's statement that "the government won't buy any vaccines that weren't tested and approved" are not disinformation. Nevertheless, in the context that Bolsonaro frequently discredited vaccines (especially Sinovac), these messages are framed to support disinformation discourses. This finding also indicates a concern with the journalistic practice of simply reproducing politicians' statements on headlines that might end up favoring disinformation discourses.

As for strategies to challenge disinformation, **the usage of humor was one of the most important findings**. Tweets and Facebook posts that used humor to legitimate real content had the highest average and median of retweets and shares in our sample. This is also a strategy based on the appeal to emotion (Reyes, 2011), as they make fun of disinformation content and use a strategy of trivialization to reject anti-vaccine discourse (Chagas et al., 2019). Humor was also the most prevalent strategy in the most retweeted tweets without disinformation, which is related to the discussion about Twitter's affordances and the virality of tweets from ordinary users. On Facebook, denunciation was also an important strategy that relied on criticizing Bolsonaro and his supporters for reproducing anti-vaccines discourse and boycotting Coronavac in Brazil. This strategy was associated with the high presence of politicians in our Facebook sample, which indicates the relevance of political polarization in fueling the discussion about vaccines in Brazil. This result is in line with other studies about the role of political polarization in the discussions about the pandemic on social media (Calvillo et al., 2020; Clarke et al., 2021; Gramacho and Turgeon, 2021).

Focusing on the third question, the connections between disinformation and populism, we found that **most of the legitimation strategies for disinformation relied on populist themes and concepts**, such as the "patriots" that were ready to fight against the elites (the Supreme Court, the corrupt communist governor Doria) and would fight for their leader (Bolsonaro). These discourses have a strong connection to populism, as they are also legitimated by the discursive formations that they refer to – such as the communist conspiracy against the "good people" (Wondreys and Mudde, 2020; Stecula and Pickup, 2021). These discourses often relied on denunciation strategy to legitimate their claims. We also found several pseudo-scientific claims connecting conspiracy theories about China, the virus, and scientists who developed a vaccine to mislead the public. These conspiracy theories were backed by populist discourse through opposition to traditional scientists who are framed as corrupted (Roudjin, 2019; Mede and Schäfer, 2020).

Most of the disinformation posts in our sample were framed to support Bolsonaro's claims about vaccines, which underlies the **appropriation of social media by the far-right** (Wodak, 2015) and the **usage of disinformation as a strategy to support populist leaders**. Many of these messages questioned the origin of the vaccine (China) and associate it with Communism, corruption, and conspiracies against the "virtuous people" (Wondreys and Mudde, 2020; Wodak, 2015; Oliveira, 2020). These arguments are in line with the far-right discourse in Brazil, mostly fueled by the struggle between Bolsonaro and Doria.

7 CONCLUSION

Our results point to far-right politicians, hyperpartisan outlets and celebrities as key actors in spreading and amplifying disinformation about vaccines on Facebook and Twitter. The political framing was decisive for the disinformation campaign against the vaccines. As for discursive strategies, denunciation was often used to fuel political polarization by criticizing politicians and democratic institutions that promoted vaccines. Besides, the usage of pseudo-scientific discourse (rationalization) was central to questioning vaccines' safety and efficacy and promoting distrust in vaccines.

Our findings also shed light on how populist leaders are using disinformation to support their discourse, as populism was often used to reinforce an "us" versus "them" context. This was used to oppose Bolsonaro and Communist China, Sao Paulo governor Joao Doria and the Brazilian Supreme Court (political populism); and to challenge Brazilian health authorities such as ANVISA and scientific institutions by promoting anti-vaccine discourse (scientific populism).

Another important finding is the usage of humor as an important strategy to fight disinformation. It seems that humor can legitimize the pro-vaccine discourse much more than other strategies and other actors, resulting in more spread both on Twitter and Facebook (although it was less prevalent in the latter). Denunciation was also a relevant strategy to challenge disinformation on Facebook. Both humor and denunciation deal with emotions, which might be a key element in challenging disinformation. Future studies may focus on this type of appeal to emotion to investigate how it can help fight disinformation.

This study has limitations. We used specific keywords for data collection that might have favored messages containing disinformation, as framing Coronavac as "Chinese vaccine" was part of disinformation discourses in Brazil. Besides, Twitter API and, especially, CrowdTangle have limitations, as we collect limited samples of tweets and Facebook posts. Therefore, the generalizations made in this study are based on a limited sample of tweets and Facebook posts in a particular period. Consequently, certain messages might have received more interactions after data collection. Besides, tweets and Facebook posted that were excluded or removed before data collection are missing in our dataset. Furthermore, our decision to analyze the most shared tweets and Facebook posts was due to the aim of this study,

but this also means that we did not analyze a large part of our datasets. In addition, some accounts were suspended, and posts excluded, therefore we relied only on the data provided by Twitter API and CrowdTangle to evaluate their content.

FUNDING STATEMENT AND ACKNOWLEDGMENTS

This research received funding from the Rio Grande do Sul State Foundation for Research Support (FAPERGS), Grant No. 19/2551-0000688-8 and the National Council for Scientific and Technological Development (CNPq) Grant No. 301433/2019-4.

REFERENCES

Alcantara, J., and Ferreira, R. (2020). 'A infodemia da "gripezinha": uma análise sobre desinformação e coronavírus no Brasil', *Chasqui. Revista Latinoamericana de Comunicación, 1*(145), pp. 137-162, https://doi.org/10.16921/chasqui.v1i145.4315

Allcott, H., and Gentzkow, M. (2017). 'Social Media and Fake News in the 2016 Election', *Journal of Economic Perspectives*, 31 (2), pp. 211-36, http://doi.org/10.1257/jep.31.2.211

Araujo, R. F., and Oliveira, T. M. (2020). 'Desinformação e mensagens sobre a hidroxicloroquina no Twitter: da pressão política à disputa científica', *Atoz: Novas Práticas em Informação e Conhecimento, 9*(2), pp. 196-205, http://dx.doi.org/10.5380/atoz.v9i2.75929

Basch, C. H., Meleo-Erwin, Z., Fera, J., Jaime, C., and Basch, C. E. (2021). 'A global pandemic in the time of viral memes: COVID-19 vaccine misinformation and disinformation on TikTok', *Human Vaccines and Immunotherapeutics*, 17(8), pp. 2373-2377, https://doi.org/10.1080/21645515.2021.1894896

Brennen, J., Simon, F., Howard, P., and Nielsen, R. (2020). '*Types, Sources, and Claims of COVID-19 Misinformation*'. Reuters Institute Report, https://reutersinstitute.politics.ox.ac.uk/types-sources-and-claims-covid-19-misinformation

Boulianne, S., Koc-Michalska, K., and Bimber, B. (2020). 'Right-wing populism, social media and echo chambers in Western democracies', *New Media and Society*, 22(4), pp. 683–699. https://doi.org/10.1177/1461444819893983

Calvillo, D. P., Ross, B. J., Garcia, R. J. B., Smelter, T. J., and Rutchick, A. M. (2020). 'Political Ideology Predicts Perceptions of the Threat of COVID-19 (and Susceptibility to Fake News About It)', *Social Psychological and Personality Science, 11*(8), pp. 1119–1128, https://doi.org/10.1177/1948550620940539

Bivar, G. C. C., Aguiar, M. E. S. C., Santos, R. V. C., and Cardoso, P. R. G. (2021). 'Covid-19, the anti-vaccine movement and immunization challenges

in Brazil: A review', *Scientia Medica*, 31(1), e39425, https://doi.org/10.15448/1980-6108.2021.1.39425

Casarões, G., and Magalhães, D. (2021). 'The hydroxychloroquine alliance: how far-right leaders and alt-science preachers came together to promote a miracle drug', *Rev. Adm. Pública*, 55(1), pp.197-214, http://dx.doi.org/10.1590/0034-761220200556

Cesarino, L. (2020). 'How Social Media Affords Populist Politics: Remarks on Liminality Based on the Brazilian Case' *Trabalhos em Linguística Aplicada*, 59(1), pp. 404-427, http://dx.doi.org/10.1590/010318136861916202004 10

Chagas, V., Freire, F., Rios, D., and Magalhães, D. (2019). 'Political memes and the politics of memes: A methodological proposal for content analysis of online political memes' *First Monday*, 24(2), https://doi.org/10.5210/fm.v24i2.7264

Clarke, E., Klas, A., and Dyos, E. (2021). 'The role of ideological attitudes in responses to COVID-19 threat and government restrictions in Australia', *Personality and individual differences*, 175, 110734, https://doi.org/10.1016/j.paid.2021.110734

Crilley, R., and Chatterje-Doody, P. N. (2020). 'From Russia with Lols: Humour, RT, and the Legitimation of Russian Foreign Policy' *Global Society*, 35(2), pp. 269-288, https://doi.org/10.1080/13600826.2020.1839387

CrowdTangle Team (2021). *CrowdTangle*, Facebook, Menlo Park, California, United States. URL: https://apps.crowdtangle.com/search/

Derakhshan, H., and Wardle, C. (2017). 'Information Disorder: Definitions', *Proceedings of Understanding and Addressing the Disinformation Ecosystem*. Annemberg: University of Pennsylvania, pp. 5-12.

Evangelista, R. and F. Bruno (2019). 'WhatsApp and political instability in Brazil: targeted messages and political radicalisation', *Internet Policy Review 8*(4), https://doi.org/10.14763/2019.4.1434

Fairclough, N. (2001). *Discurso e mudança social*. Brasília: Editora UnB.

Freelon, D. (2010). ReCal: Intercoder reliability calculation as a web service. International Journal of Internet Science, 5(1), 20-33.

Galhardi, C. P., Freire, N. P., Minayo, M. C. S., and Fagundes, M. C. M. (2020). 'Fact or Fake? An analysis of disinformation regarding the Covid-19 pandemic in Brazil', *Ciência and Saúde Coletiva*, *25*(2), pp. 4201-4210, https://doi.org/10.1590/1413-812320202510.2.28922020

Glozer, S., Caruana, R., and Hibbert, S. A. (2019). 'The Never-Ending Story: Discursive Legitimation in Social Media Dialogue', *Organization Studies*, 40(5), pp. 625–650, https://doi.org/10.1177/0170840617751006

Gramacho, W. G., and Turgeon, M. (2021). 'When politics collides with public health: COVID-19 vaccine country of origin and vaccination acceptance in Brazil', *Vaccine*, 39(19), pp. 2608-2612, https://doi.org/10.1016/j.vaccine.2021.03.080

Hallal, P.C., and Victora, C.G. (2021). 'Overcoming Brazil's monumental COVID-19 failure: an urgent call to action', *Nat Med*, 27(933), https://doi.org/10.1038/s41591-021-01353-2

Igwebuike, E. E. and Chimuanya, L. (2021). 'Legitimating Falsehood in Social Media: A Discourse Analysis of Political Fake News', *Discourse and Communication*, 15(1), pp. 42-58, https://doi.org/10.1177/1750481320961659

Krippendorff, K. (2012). *Content Analysis*: An Introduction to Its Methodology. Sage Publications.

Kallil, I. (2019). *Who are Jair Bolsonaro's voters and what they believe*. Report. https://doi.org/10.13140/RG.2.2.35662.41289

Lindgren, S. (2016). 'Introducing Connected Concept Analysis: A network approach to big text datasets', *Text and Talk*, 36(3), pp. 341-362, https://doi.org/10.1515/text-2016-0016

Mede, N. G., and Schäfer, M. S. (2020). 'Science-related populism: Conceptualizing populist demands toward science' *Public Understanding of Science*, *29*(5), pp. 473–491, https://doi.org/10.1177/0963662520924259

Menczer, F., and Hills, T. (2020). 'The Attention Economy', *Scientific American*, 323(6), pp. 54–61, https://www.scientificamerican.com/article/information-overload-helps-fake-news-spread-and-social-media-knows-it/

Mendonça, R. F., and Caetano, R. D. (2021). 'Populism as Parody: The Visual Self-Presentation of Jair Bolsonaro on Instagram', *The International Journal of Press/Politics*, *26*(1), pp. 210–235, https://doi.org/10.1177/1940161220970118

Metaxas, P. T.; Mustafaraj, E., Wong, K., Zeng, L., O'Keefe, M., and Finn, S. (2018). 'Do retweets indicate interest, trust, agreement?', *preprint arXiv*, https://arxiv.org/abs/1411.3555

Mudde, C., and Kaltwasser, C. R. (2017). *Populism*: A very short introduction. Oxford: Oxford University Press.

Oliveira, T. (2020). 'Desinformação científica em tempos de crise epistêmica: circulação de teorias da conspiração nas plataformas de mídias sociais', *Fronteiras-estudos midiáticos, 22*(1), pp. 21-35, https://doi.org/10.4013/fem.2020.221.03

Recuero, R., Soares, F., and Vinhas, O. (2021). 'Discursive strategies for disinformation on WhatsApp and Twitter during the 2018 Brazilian presidential election', *First Monday*, 26(1), https://doi.org/10.5210/fm.v26i1.10551

Recuero, R., Soares, F., Zago, G. (2021). 'Polarization, Hyperpartisanship, and Echo Chambers: How the disinformation about COVID-19 circulates on Twitter', *Contracampo*, 40(1), https://doi.org/10.22409/contracampo.v40i1.45611.

Reyes, A. (2011). 'Strategies of legitimization in political discourse: From words to actions', *Discourse and Society*, 22(6), pp. 781–807, http://doi.org/10.1177/0957926511419927

Ricard, J., and Medeiros, J. (2020). 'Using misinformation as a political weapon: COVID-19 and Bolsonaro in Brazil', *Harvard Kennedy School (HKS) Misinformation Review*, https://doi.org/10.37016/mr-2020-013

Rogers, R. and Niederer, S. (2020). 'The politics of social media manipulation'. Amsterdam University Press B. Available at: https://library.oapen.org/bitstream/handle/20.500.12657/42884/9789048551675.pdf?sequence=1andisAllowed=y

Rooduijn, M. (2019). 'State of the field: How to study populism and adjacent topics? A plea for both more and less focus', *European Journal of Political Research, 58*, pp. 362-372, https://doi.org/10.1111/1475-6765.12314

Santos, M.D., and Fossá, M.I. (2020). 'A Disputa pelo poder político em meio à pandemia de Covid-19, análise do confronto entre Joao Doria e Jair Bolsonaro', *Panorama*, 10(1), http://doi.org/10.18224/PAN.V10I1.8297

Shahi G. K., Dirkson A., and Majchrzak, T. A. (2021). 'An exploratory study of COVID-19 misinformation on Twitter', *Online Soc Netw Media*. 22, 100104, http://doi.rg/10.1016/j.osnem.2020.100104

Silva, Y. (2020). *Bolsonaro and Social Media: A Critical Discourse Analysis of the Brazilian President's Populist Communication on Twitter*. Master Thesis. Master's Programme in Social Sciences, Digital Media and Society specialization. Upsala Universitet. Available at: https://www.diva-portal.org/smash/get/diva2:1448385/FULLTEXT01.pdf

Soares, F. B., Recuero, R., Volcan, T., Fagundes, G., and Sodré, G. (2021). 'Research note: Bolsonaro's firehose: How Covid-19 disinformation on WhatsApp was used to fight a government political crisis in Brazil', *Harvard Kennedy School (HKS) Misinformation Review*, https://doi.org/10.37016/mr-2020-54

Stecula, D. A., and Pickup, M. (2021). 'How populism and conservative media fuel conspiracy beliefs about COVID-19 and what it means for COVID-19 behaviors'. *Research and Politics*, https://doi.org/10.1177/2053168021993979

Van Leeuwen, T. (2007). 'Legitimation in discourse and communication', *Discourse and Communication*, *1*(1), pp. 91–112, https://doi.org/10.1177/1750481307071986

Van Dijk, T. A. (2006). 'Discourse and manipulation', *Discourse and Society*, 17(2), 359-383.

Van Dijk, T. A. (2009). 'Critical Discourse Studies: A Sociocognitive Approach', In: Wodak, R., and Meyer, M. *Methods for Critical Discourse Analysis*. London: Sage, pp. 62-85.

Vaara, E., and Tienar, J. (2008). 'A Discursive Perspective on Legitimation Strategies in Multinational Corporations', *Academy of Management Review*, 33(4), pp. 985–993, https://doi.org/10.5465/amr.2008.34422019

Vosoughi, S., Roy, D., and Aral, S. (2018). 'The spread of true and false news online', *Science*, 359(6380), pp. 1146–1151, https://doi.org/10.1126/science.aap9559

Watmough, S. P (2021). 'Jair Bolsonaro: Far-Right Firebrand and Cheerleader for Dictatorship', *Populism Studies*. Retrieved from https://www.populismstudies.org/jair-bolsonaro-far-right-firebrand-and-cheerleader-for-dictatorship

Wasserman, S., and Faust, K. (1994). *Social Network Analysis*. Cambridge: Cambridge University Press.

Wondreys, J., and Mudde, C. (2020). 'Victims of the Pandemic? European Far-Right Parties and COVID-19', *Nationalities Papers*, pp. 1-18, https://doi.org/10.1017/nps.2020.93

Wodak, R. (2015). *The politics of fear*: What right-wing populist discourse mean. Sage. (257p.) http://dx.doi.org/10.4135/9781446270073

Wodak, R. and Krzyzanowski, M. (2017). 'Right-wing populism in Europe and USA: Contesting politics and discourse beyond 'Orbanism' and 'Trumpism'', *Journal of Language and Politics*, 16(4), http://doi.org/10.1075/jlp.17042.krz

JOURNAL OF DIGITAL SOCIAL RESEARCH

WWW.JDSR.IO ISSN: 2003-1998

VOL. 4, NO. 1, 2022, 98–127

TWEETED ATTITUDES TOWARDS WOMEN PARLIAMENTARY CANDIDATES IN KUWAIT: A SOCIAL DOMINANCE PERSPECTIVE

Hesham Mesbah[a]

ABSTRACT

This study explores how Kuwaitis use Twitter to communicate their attitudes towards women Parliamentary candidates (WCs) in a traditionally male-dominated society, and how these tweeted attitudes are thematically constructed, either negatively or positively. The study also explores how these attitudes differ according to gender and evolve quantitatively and qualitatively over subsequent elections. A total of 1744 tweets about all eight women candidates in 2013 posted for 40 days prior to the Kuwait Parliamentary Election were retrieved and analyzed. The tweets posted about the two women candidates of those eight who continued to run in 2016 and 2020 were also analyzed in terms of length, content, and themes. Gender significantly correlated with attitudes in the first election, but not in the subsequent two elections. Tweeted attitudes turned to be more elaborate, information-based, and longer over consecutive elections. The dominant positive theme was generic, whereas the dominant negative theme was specific, and candidate based. Women candidates were praised for acting like men, whereas they are mocked for looking like men.

Keywords: Women leadership; Twitter; Attitudes; Kuwait; Gender; Social Dominance.

[a] Rollins College, USA

1 INTRODUCTION

Twitter has provided a platform for online political communication for citizens to discuss policies, public events, and politicians. Consequently, the use of Twitter for political purposes has aroused considerable interest among researchers. Compared to other popular social media platforms, content on Twitter has two traits that facilitate large-scale data collection and analysis. Unlike Facebook and blogs, "the majority of Twitter content is explicitly public; Twitter data is encoded in a single format" (Bamman, Eisenstein, & Schnoebelen, 2014, p. 139).

The use of Twitter for sharing political views during either presidential or parliamentary elections is particularly intense (Ceron, Curini, Iacus, & Porro, 2014; Conway, Kenski, & Lang, Cuhna, 2014; 2015; Larsson & Moe, 2011; Minot, Arnold, Alshaabi, Danforth, & Dodds, 2021). A significant body of research on political communication on Twitter has been devoted to analyzing the opinions and attitudes of voters and how specific factors, such as gender, explain those attitudes (Boutet, Kim, & Yoneki, 2011; Ceron et al., 2014; Cohen & Ruths, 2013; Conover, et al., 2011; Gul et al., 2016; Hu and Kearney, 2021; Mejova, Srinivasan, & Boynton, 2013; Rao et al., 2021; Rodríguez-Ibáñez, Gimeno-Blanes, Cuenca-Jiménez, Soguero-Ruiz, & Rojo-Álvarez, 2021; Wong, Tan, Sen, & Chiang, 2013).

Politically marginalized groups, such as women in patriarchal societies, are found to be using Twitter as a "counter-public sphere" to share their voices and gain publicity (Hu and Kearney, 2020). Political leadership in the Arab World is dominated by men. No woman has ever been able to attain the position of head of state or Speaker of Parliament in this part of the world over centuries. In addition, the Arab region has the lowest regional average of women parliamentarians worldwide (IPU, 2021). In these historically patriarchal societies, women must overcome cultural hurdles to claim political leadership roles, even though the first Arab woman minister was appointed in 1962 in Egypt (Bier, 2010).

In the traditionally patriarchal society of Kuwait, women were granted the right to vote and stand in Parliamentary elections in 2005. It was a major political and cultural transformation that posed several questions about how the candidacy and potential leadership of women would be evaluated by both the dominant and marginalized genders in this country. Ever since, the participation of women in the Kuwaiti parliamentary elections has not been representative of the percentage of women in the population. In some instances, women candidates have never been able to win a seat in Parliament. This demonstrates that women do not necessarily vote for women in this part of the world, which raises questions about how voters perceive the strengths and weaknesses of women candidates and whether these perceptions vary according to the gender of the voters. A longitudinal analysis of the attitudes expressed by social media users would provide an unobtrusive tool for measuring these attitudes from one election to another. Using a survey to analyze

the attitudes of voters toward women candidates in a collectivist culture that values face-saving implies the risk of obtaining less reliable responses.

The present study seeks to deconstruct and hence, thematically categorize the attitudes of Twitter users towards women candidates (WCs) in Kuwait over time, not just the general sentiment. The study also seeks to identify whether those tweeted attitudes toward WCs who keep running in consecutive parliamentary elections change affectively and cognitively over time. In addition, the study fills a gap in the literature about the cultural perceptions of WCs in a traditionally patriarchal society that is passing through a phase of both modernization and democratization.

2 BACKGROUND

Kuwait gained its independence from Britain in 1961. A year later, this monarchy wrote its constitution and established the first People's Assembly in the Arabian Peninsula. The new constitution denied women the right to either vote or run for parliamentary elections. Women's suffrage had to wait until 2005 when the parliament honored the Amir (ruler) Decree, which granted women the right to elect and be elected. A unique element of the political environment in Kuwait is that this constitutional democracy has never had any official political parties. Tribes and ideological leanings (Islamic versus liberal) represent the functional alternative to what is called a multiparty system in other modern democracies.

With the initiation of commercial exports of Kuwait oil in 1946, the per capita income and the gross domestic product started to skyrocket. This monarchy was ranked as having the 8th highest per capita income in 2017 (Worldodometer). The foundation of the Kuwaiti economy has been transformed from trading, fishing, and pearling to the extraction and exportation of oil, paving the way for modernization in Kuwait. In recent years, communication technology has become an intrinsic component of the lives of average citizens. Kuwait has become one of the most active countries on Twitter. With the highest number of per capita tweets, Kuwait was known as the "tweetiest" country in the world (Moncanu et al., 2013). In 2016, Kuwait had the most Twitter users per capita in the MENA region (Statista, 2016). Although the population of Kuwait represents only one percent of the population of the Arab world, Kuwaitis were estimated to produce 10% of the tweets in the region (Arab Social Media Report, 2017). In March 2017, Twitter had a penetration rate of 12.6% in Kuwait, which was the second-highest rate after Bahrain (Weedoo, 2017).

In the petroleum era, Kuwait has found itself caught between the two opposing forces of modernization and tradition (Tetreault & Al-Mughni, 1995), which has led to the creation of a political system that encourages the education of women but discourages their civil independence and participation in political leadership. By law, women do not have the same legal rights as men. Kuwaiti women who marry non-Kuwaitis cannot pass on their nationality to their children,

whereas the children of a Kuwaiti male are born Kuwaitis. These newly granted rights to elect and be elected did not lead to a significant surge in the political activism of women. Although women represent most of the electorate in Kuwait, only a handful of WCs to none have managed to win seats in Parliament. As reported by the World Bank (2021), the marginalized political status of Kuwait women is not commensurate with their participation in the labor force compared to men. According to the latest report of the Central Statistical Bureau (CSB, 2019-2020) of Kuwait in 2019, Kuwaiti women slightly outnumber Kuwaiti men in the labor market (188,141 women vs. 154,276 men), even though women represented around 75% of the student body at Kuwait University (the only state university in the country) in the same year (National Academies of Sciences, Engineering, and Medicine, 2020).

3 THEORETICAL FRAMEWORK

In his "resource curse" theory, which seeks to explain gender inequality in the context of the prevailing economic activity in a society, Michael Ross suggests that oil production is the primary reason women lag (Ross, 2008). According to him, oil extraction relies on manual jobs, which are traditionally male-dominated, and that fact lowers the level of participation of women in the labor market. This in turn affects women's identities and self-perceptions negatively, and thus limits their participation in formal political and economic networks. Accordingly, the political influence of women would be diminished because "petroleum perpetuates patriarchy" (Ross, 2008, p. 120).

However, this theory falls short of explaining women's empowerment in other traditional but less industrialized societies. Rwanda is the leading country worldwide in terms of women parliamentary representation, with women holding 61.3% of the seats in the Lower House in 2021 (Inter-Parliamentary Union, 2021). On the other hand, women suffer from gender inequality in non-oil-rich countries such as Jordan and Yemen. Pippa Norris argues that Ross does not factor in how the "public feels about men and women in the workforce, society's moral values towards sexuality, and the strength of religious identity" (Norris, 2011, p. 9). Norris suggests that it is "Mecca" rather than "Oil" that explains gender inequality in Arab societies. In her meta-analysis of a series of social change studies, Norris remarked that egalitarian attitudes towards the election of women are stronger in non-Islamic societies. She concluded that "long-standing religious traditions leave an enduring mark on gendered norms and beliefs" (Norris 2011, p. 17). In her view, it is Islam that determines "the division of labor for men and women in the home, family, and public sphere."

A substantial body of literature suggests that Islam is the main factor that accounts for the disempowerment of women (Abdalla, 1996; Amawi, 2001; Inglehart & Norris, 2003; Jamali, Sidani, & Safieddine, 2005; Khimish, 2014; Pettygrove, 2011; Saleema & Tlaiss, 2011; Sidani, 2005). However, women have

reached the top position (presidency or premiership) in several Muslim nations, such as Bangladesh, Indonesia, Pakistan, Turkey, and Tunisia. Women also have 43% of the seats in the Senegal Parliament, where Islam is the main religion. In contrast, women have 27.6% of the House of Representatives in the U.S. (IPU, 2021). Attitudes towards women in leadership roles are found to be negative in some predominantly Christian societies. In Greece, employed men were not found to hold strong positive beliefs about the ability of women to assume leadership roles (Galanaki, Papalexandris, & Halikias, 2009).

Both the "Resources Curse" and the "Mecca" theories adopt ungeneralizable perspectives by attributing gender inequality in oil-rich Arab countries to historical and ideological factors, respectively. Women claim significant leadership roles in oil-rich, Muslim countries, such as Azerbaijan, which met the world's demand for crude oil in 1890 and gave women the right to vote in 1918 before both the U.K. and the U.S. did (Scott, Dakin, Heller, & Adriana, 2013). The "Resources" and "Mecca" theories lack this cultural/social perspective in explaining women disempowerment, whereas the theory of Social Dominance Orientation (SDO) is more focused on that perspective.

The theory of SDO explains how societal hierarchies and attitudinal orientations cause individuals to either accept or reject the traditional hierarchy of power within their societies. It focuses on "attitudes towards hierarchical relationships between groups and the desire to promote intergroup domination" (Sidanius, Pratto, Laar, & Levin, 2004, p. 858). Individuals who have a high level of SDO tend to adhere to the traditional hierarchy in their society, even if it requires bias to do so (Simmons, Duffy, & Alfraih, 2012). Simmons and his colleagues studied a sample of 89 college students from the U.S. and Kuwait. They found that high levels of SDO were negatively related to favorable attitudes towards women as managers.

The relationship between gender and SDO has been the focus of the Invariance Hypothesis (IH). According to this hypothesis, men tend to have higher levels of SDO than women, regardless of social class and family income (Pratto and Sedanius, 1994). Men, who are historically more dominant politically and economically, will be more reluctant to embrace ideas of egalitarianism and gender equality. They would also use negative, sarcastic attitudes towards WCs as an equality-suppressing and hierarchy-enhancing strategy that serves to legitimize the myth that inequality is normal and, thus, justified.

To the contrary of IH, SDO explains why some men might not overtly oppose women's leadership roles. It predicts that men generally play roles that enhance hierarchy (such as police, military, and business executives), whereas women are "over-represented in roles that accentuate hierarchy, such as teachers and social workers" (Pratto, Sidanius, & Levin, 2006, p. 295). Men will tend to maintain hierarchy if they are in control of the resources of their society. However, men do not tend to show overtly biased behavior towards women, as this will not be necessary when hierarchically driven, long-established oppression within a

society is already suppressing women (Simmons, et al., 2012). For instance, all parties emphasized the role of women during the parliamentary elections of 1997 in Jordan. However, all the parties and tribes—except for the Communist Party—refrained from fielding WCs (Amawi, 2007).

SDO also assumes that the subordinate members of a society will collaborate with the dominant members in order to maintain the hierarchical structure of their society and, hence, participate in their own oppression (Pratto et al. 2006, p. 276). Although the oppressed might tend to support each other emotionally, they will tend to "display greater endorsement of hierarchy-attenuating legitimizing myths" (Pratto et al., p. 276). This might explain why WCs in traditionally hierarchical societies do not receive the majority of the votes, even though women represent the majority of the voters. This suggests that women voters in Kuwait are less likely to be significantly more supportive of WCs than male voters.

Accordingly, gender might predict the variance in the tweeted attitudes towards WCs in Kuwait, especially after the legal changes in 2012 that led candidates to run individually. Accordingly, voters would no longer vote for one of the collaborative groups of candidates but instead select individual candidates. This represented a challenge for WCs, who would have to run individually instead of collaborating with other well-established male candidates.

The literature does not provide empirical evidence that gender will correlate significantly with attitudes towards WCs following this legal change, especially after the inability of such candidates to win more than two seats in 2013. The theory of SDO predicts that the dominant social group tend to communicate more favorably towards the suppressed group as long as they are capable of maintaining the traditional societal hierarchy.

RH1: Gender is expected to correlate more significantly with the tweeted attitudes towards Kuwaiti WCs posted before the 2013 elections than with those posted before the subsequent elections in 2016 and 2020.

Perceptions of gender roles and attitudes towards women leadership are relative constructs that are influenced by a myriad of differing factors. An analysis of such perceptions and attitudes may help in identifying those factors. Kauser and Tlaiss (2011) suggest that negative attitudes towards women managers "reinforce gender role stereotyping that defines women via their domestic and reproductive chores" (p. 38). Similarly, a study analyzing tweets sent to the parliamentary members in Britain found that women members were the targets of gendered abuse and incivility (Southern and Harmer, 2021). In the Arab region, negative attitudes towards working women have been shown to be the biggest obstacle to the career progress of women (Jamali et al., 2005; Mostafa, 2005). In such patriarchal societies, women are perceived primarily as caregivers who are either dependent on or dedicated to the care of others, whereas men are perceived as breadwinners and leaders. In these societies, women rely on the support of their families or tribes in order to succeed in business or politics. Shanahan (2009) reported that the six women who won in the 2005 Parliamentary election in Lebanon succeeded because

of their links with patriarchal political families. Madsen (2010) conducted in-depth interviews with six women business leaders in the UAE. These professional women did not report any dissatisfaction with gender roles in their families; they described their fathers as being well-traveled, open-minded, and supporters of women's education, whereas they referred to their mothers as "kind" and "pretty."

However, the continued involvement of women in leadership roles might be another factor in changing how they are perceived by other members of society. The attitudes and opinions voiced about WCs via public forums might evolve over time if those candidates maintain their exposure and political involvement. According to the elaboration likelihood model, with repeated exposure to messages from WCs, individuals might process those messages more centrally, where they tend to "bring diverse issue-related thoughts and previous experience to evaluate an argument" (Oh & Sundar, 2015, p. 215). Individuals might also use a central route (issues and policies) in processing information about WCs after gaining more information about them from one campaign to the subsequent. Accordingly, the peripheral stimuli about WCs (non-issue-related stimuli, such as gender, age, or appearance) might be processed more attentively by voters who lack prior knowledge about those candidates. Being exposed to more information about WCs, individuals might apply a more logical reasoning in supporting or rejecting them.

Research also shows that Twitter's decision in 2017 to increase the maximum length of tweets from 140 to 280 characters had an impact on the linguistic aspects of tweets (Gligorić, Anderson, & West, 2018). Accordingly, Twitter users are expected to use more characters (quantitative cues) and express more information-driven attitudes (qualitative cues) in their tweets about WCs who run for elections frequently.

RH2: Tweeted attitudes towards WCs in Kuwait will be associated with peripheral cues in the elections of 2013 and central cues in the subsequent elections.

RH3: Tweeted attitudes towards WCs in Kuwait will be significantly longer in 2016 and 2020 compared to 2013.

Identifying the "content" of both negative and positive tweeted attitudes towards WCs in Kuwait is also important to "capture emerging themes" (Barry et al., 2005, p. 239) and uncover how Twitter users react to the candidacy of women. Such thematic analysis is expected to provide a description of the affective and cognitive components of the tweeted opinions and attitudes towards WCs in Kuwait over three consecutive parliamentary elections.

RQ: What are the cognitive and affective components that summarize the main themes of the tweeted attitudes towards WCs in Kuwait before the parliamentary elections in 2013, 2016, and 2020?

4 METHOD

4.1 Data

In 2012, the electoral laws were amended to reduce the number of candidates a Kuwaiti can elect from four to one. This represented a challenge for underrepresented groups, especially women, who had to run individually, not in collaboration with long-established candidates. The present study analyzes the tweeted attitudes towards all WCs who ran after this legal modification. First, the tweets about all eight WCs[1] who ran in 2013 were retrieved and analyzed. The tweets posted about the two WCs who continued running in the subsequent two elections in 2016 and 2020 were also analyzed. Only the tweets posted during the legal time of campaigning in those three elections (40 days before voting started) were retrieved. The call for candidates to declare their candidacy is usually announced 40 days before the election. The tweets were searched and retrieved using Twitter Advanced Search. The names of the WCs were the search keywords. Table 1 shows specific dates of data collection and the ratio of participation of WCs in each election.

Table 1. Timeframe for Retrieving Tweets about WCs Before Three Parliamentary Elections in Kuwait (2013-2020)

Year	No. of tweets retrieved	Data collection start date	Data collection end date	Total candidates	Total WCs	% Of WCs	No. of WCs analyzed
2013	1521	June 17	July 27	310	8	2.6%	8
2016	215	Oct.15	Nov. 25	454	15	3.3%	2
2020	141	Oct. 20	Nov. 30	395	33	8.4%	2
Total	1877			1159	56	4.8%	12

A total of 1877 tweets were retrieved. The tweets posted by either the news media or the WCs themselves were separated from the rest of the tweets and discarded from the analysis. Retweets that included only video and/or news clippings were

[1]These eight WCs were: ***Anwar Alkahtany***, a political activist who ran for parliamentary elections several times but never won; ***Awatef Alkallaf***, a Kuwaiti athlete who won several regional championships in shooting; ***Hya Almutairy***, a psychiatrist who ran in the parliamentary elections in 2013. ***Jawaher Marefy***, a schoolteacher who ran for the first time in 2013; ***Masoma Almubarak***, a professor of political science at Kuwait Univ. and the first woman minister in Kuwait. She was appointed Health Minister in 2005. She was also the first woman elected as parliamentarian in 2009; ***Rabah Alnejada***, a professor of physical education at the General Authority of Applied Education; ***Riham El-jelwy***, a young businesswoman and the sister-in-law of the former Kuwaiti Prime Minister Sheikh Jaber Mubarak Alsobah. Her first parliamentary candidacy was in 2013; ***Safa Alhashem***, a businesswoman who was a parliamentary candidate six times in 2012, December 2012, 2013, 2016, and 2020). She won in 2013 and 2016.

also excluded, and only text messages were included in the analysis. The actual number of tweets included in the analysis is 1744.

4.2 Coding & Analysis

A native Kuwaiti researcher[2] coded 10% of the tweets. A test of inter-coder reliability showed a range of .85 to .95 of Alpha Cronbach's reliability for the categories and themes used in the analysis, indicating an acceptable level of coding reliability.

We coded each tweet according to these categories:

1) Character count (excluding all non-textual elements such as links to videos and websites).

2) Generality or specificity of the tweeted attitude. A tweet was coded as general (Gen) when it generally included a candidate's name, good wishes for a candidate, a general eulogy, such as "the best" or "well-deserving," a pledge to vote for a specific candidate, or ill-wishes, such as, "I wish she wouldn't win." A tweet was coded as "specific" when it included an assessment of a candidate's skills, achievements, opinions, policies, and/or specific reasons for supporting or rejecting a candidate.

3) Peripheral processing in the tweet, which indicates that it is more focused on the personal or demographic characteristics of a candidate rather than her views, policies, and actions. A tweet was coded "peripheral" when it focused on a candidate's gender, caste, appearance, age, tribal affiliation, and social status.

4) Gender of tweeters (identified by reading their profile information). In cases where we did not find information about the gender of the tweeters, we read their tweets to identify whether they used masculine or feminine pronouns (which are easily identified in the Arabic language). Luckily, we did not find any accounts that were accessible only to followers. In previous research, the gender of Twitter users was identified through the user's first name (Bamman, et al. 2014; Burger, et al. 2011; Cunha, Magno, Gonçalves, Cambraia, & Almeida, 2014). However, using only the first name to infer the gender of a Twitter user was not an option for the present study, as it is not uncommon for Kuwaiti Twitter users, especially women, to use pseudonyms as their user names.

5) The general attitude of the tweets towards WCs. Each tweet was coded into three sub-categories: positive, neutral, and negative. The coders read each tweet to determine whether it was favoring a candidate, against her, or just making a neutral reference to her. Neutral references include such things like posting a link to a news story in traditional media, quoting a candidate, and posting an announcement for a candidate's rally or media appearance. Although such tweets could have been indicative of a positive leaning towards a candidate, we focused our analysis only on the explicit, verbalized expressions of the tweeters. Negative tweets were coded according to the existence of any of these aspects: the use of negative descriptions/attributes about any WC (such as liar, hypocrite, greedy, and nepotistic); mockery of the ethnicity, looks, and age of the WCs; criticism directed

towards the views of any WC; expressing a preference for male candidates in referring to a WC; and general rejection of the candidacy of a WC (I won't vote for you). Positive tweets included positive descriptions (such as honorable, experienced, and having a good style), overt support and good wishes, appreciation of past political performance, preferring any WC over male candidates, complimenting a WC's looks and young age, and expressing general advocacy for a WC (such as encouraging tweeters to attend a WC's political rally and praying that she will win).

In addition, an extant textual analysis was applied to all the tweets to determine their main attitudinal themes without using a preexisting coding frame (Braun & Clarke 2006). Extant textual analysis is an unobtrusive method that allows more objectivity in data collection (Charmaz 2007) and inhibits the interference of researchers and, hence, the potential contamination of original opinions. This method of analysis in the present study is adopted to interpret the cultural symbols used by tweeters in expressing their attitudes towards WCs and make connections between such symbols to create themes that summarize and describe the data at hand (the tweets).

We compiled all the tweets verbatim and read them several times. Our aim was to provide a thematic description of the entire data set. Those tweets were coded according to the main theme of the views and attitudes expressed. When a tweet had more than one theme, each theme was coded separately. A data-driven, interpretive thematic analysis was conducted to identify the meanings of attributes and descriptions in Kuwaiti culture and determine the subthemes of each theme. The analysis was conducted in two phases. First, each tweet was coded according to its attitude and the properties of such an attitude. For instance, a tweet that read, "She needs a new face," was coded as a negative attitude towards the physical features of the WC. The general code for this tweet was "Criticism of physical traits." In the second phase, those general codes were recoded into main themes, and each main theme was coded into sub-themes (axial coding) depending on its components or how any WC was mentioned. "Criticism towards physical traits" was thematically coded as "Physical traits/negative" with an axial coding of physical traits such as the face, body weight, and age. Tables 2 and 3 provide an illustration of these two phases of textual analysis.

Table 2. General Codes of Selected Tweets Based on the Attitude Properties of Each Tweet

Examples of Tweets	Properties of attitude	General Codes
She needs a new face	Negative towards looks	Criticism of physical traits
She is the most powerful female candidate ever	Positive towards general past performance	General acceptance
She says she is not anxious to win, when everybody knows that she pays the voters to win their votes	Negative towards political integrity of candidate	Lack of integrity/corrupt
Why don't you stick it with an adhesive? [a comment on a video showing the artificial teeth of a female candidate falling]	Negative towards physical traits, with a mockery component	Physical-based rejection
I swear this oldie coming from hell is the reason for the setbacks of Kuwait	Rejection based on age and negative impact on politics	Physical/personal-based rejection/age
She is so old and still dreaming of the parliamentary chair	Negative towards age of candidate	Demographic-based rejection/age
Has the brother announced her candidacy	Negative towards looks	Looks-based rejection (man-like)
We wish to see her in the House	General positive attitude based on good wishes	General acceptance/wishes
May God give her success	General positive attitude based on good religious wishes	General acceptance/wishes/religion
Please follow this account of the candidate [Account attached]	General positive attitude	Advocating the candidate/garner supporters online
She is active and balanced, and she deserves to win	Acceptance based on evaluating objective, performance-related traits	Skills-related acceptance of candidate
Her style is awesome	Positive attitude towards looks	Physical-based acceptance
I love her when she talks	Positive attitude towards communication skills	Evaluation of the communication skills of WC
I like your political past	Positive attitude towards past political performance	General evaluation of political skills/performance of WC
Let her educate herself first and then she may run	Negative attitude towards the skills/knowledge of candidate	General criticism of aptitude of candidate
She is better than 50 moustaches	Positive attitude based on comparison with male gender	Gender-based acceptance
She is the worst candidate the House has seen	Rejection based on general previous political performance	General rejection of performance/skills
She goes to a Pakistani barber shop	Negative towards looks	Criticism/mockery of physical traits

Table 3. Selective Codes (Themes) and Axial Codes (Subthemes) Based on General Codes

General codes	Selective Codes (themes)	Axial Codes (subthemes)
Advocating the candidate/garner supporters online	Advocacy/positive	Add; Vote; Attend
Gender-based acceptance	Gender-based/positive	Compare to men
General evaluation of political skills/performance of candidate; Skills-related acceptance of candidate	Political skills/positive	Past performance
Physical-based acceptance; evaluation of the skills of candidate	Personal traits/positive	Appearance Cognitive skills
General acceptance/wishes/religion	General acceptance	Good wishes Religion-related wishes
Lack of integrity/corrupt	No-integrity/negative	Corrupt
Demographic-based rejection/age; physical/personal-based rejection/age	Demographics/negative	Gender Age
Looks-based rejection (man-like); criticism/mockery of physical traits	Physical traits/negative	Looks
General criticism of aptitude of candidate; general rejection of performance/skills	Aptitude/negative	General lack of knowledge/poor performance

5 RESULTS

5.1 Research hypotheses

In the election held in 2013, only two out of eight WCs made it to Parliament, although women made up more than 53% of the electorate. These two candidates received the highest number of tweets about the WCs. In 2016, only one WC won a seat out of fifteen WCs. None of the 33 WCs won enough votes to be in Parliament in 2020.

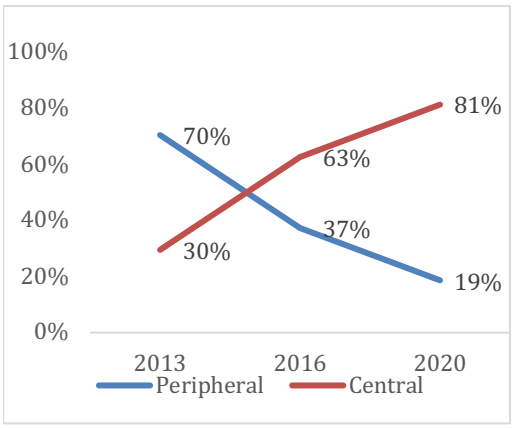
While the negative tweets outnumbered the positive ones (39% negative versus 33% positive), the women tweeters expressed significantly more positive attitudes towards WCs than the male tweeters did in 2013 (X^2 = 0.21, p = 0.001). Table 4 shows that this association was not significant in either the election of 2016 or 2020, even though the percentage of positive tweets about WCs remained higher among women tweeters in those two elections. This explains the significant positive association between gender and the tweeted attitudes towards WCs shown in Table 4. The data in that table supports RH1 and rejects its null hypothesis.

Table (4). Tweeted Attitudes towards Kuwaiti WCs (2013-2020) According to Gender

Year	Attitude	Male		Female		Total	
		N	%	N	%	N	%
2013*	Positive	148	18%	140	36%	288	23%
	Neutral	315	38%	150	38%	465	38%
	Negative	376	45%	103	26%	479	39%
	Total	839	100%	393	100%	1232	100%
2016	Positive	123	75%	86	80%	209	77%
	Neutral	10	6%	8	7%	18	7%
	Negative	32	19%	13	12%	45	17%
	Total	165	100%	107	100%	272	100%
2020	Positive	44	27%	27	35%	71	30%
	Neutral	8	5%	2	3%	10	4%
	Negative	111	68%	48	62%	159	66%
	Total	163	100%	77	100%	240	100%
Total**	Positive	315	27%	253	44%	568	33%
	Neutral	333	29%	160	28%	493	28%
	Negative	519	44%	164	28%	683	39%
	Total	1167	100%	577	100%	1744	100%

**Correlation between gender and attitudes is significant; X^2 =0.22, p< .001*
***Correlation between gender and attitudes is significant; X^2 =0.18, p< .001*

RH2 predicts a significant association between the use of specific/central cues in the tweeted attitudes towards WCs and the year of election. The data in Figure 1 shows a gradual increase in tweeting about the political statements and actions of WCs from one election to the next, while the emphasis on their demographic characteristics took the opposite direction. Nominal by nominal contingency coefficients between use of either central criteria or central cues in tweeted attitudes towards WCs and years were found significant (X^2 = 0.39 and 0.40, respectively, p< 0.001). The data rejects the null hypothesis of RH2.



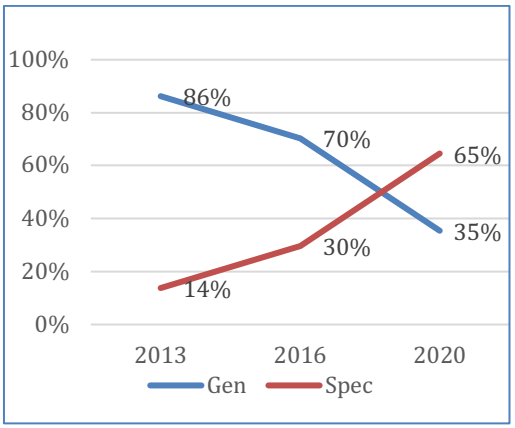


Figure 1. Use of Cues in Tweeted Attitudes towards WCs According to Year of Election

RH3 predicts that the tweeted attitudes towards WCs in Kuwait will be significantly longer (in terms of the number of characters used) in 2016 and 2020 compared to 2013. The Kruskal-Wallis test shows a statistically significant association between the mean number of characters used in the tweeted attitudes towards WCs and the year of election ($H = 29.91$, $df = 2$, $sig < 0.001$). To test the statistical difference between each pair of election years, pairwise comparisons have been computed to test these expected differences. As shown in Table 5, the tweeted attitudes in 2020 were significantly longer than those used in either 2013 or 2016. However, the average number of characters in the tweeted attitudes during the 2016 elections was not statistically different from the average in 2013. This data partially supports RH3.

Table 5. Statistical Significance of Difference between Average No. of Characters of Tweeted Attitudes Posted in 2013, 2016, and 2020

Year	Mean	N	SD	Sample 1-Sample 2	Test Statistic	Sig.	Adj. Sig.
2013	66.5	1232	35.3	2016-2013	0.65	0.42	1.00
2016	67.7	272	41.6	2016-2020	6.15	0.01	0.04
2020	105.3	240	80.4	2013-2020	11.69	0.00	0.00
Total	72.1	1744	47.1				

5.2 RQ: Tweeted qualities/traits of WCs

The data demonstrated five positive and five negative attitudinal themes. The coders counted the frequency of each theme and its subthemes to assess the weight of each attitudinal component. Figures 1 and 2 illustrate, in ascending order, the positive and negative attitudinal themes and their subthemes. These figures were

created using the free engine WordSift (a word visualization and open program that identifies word clusters and visualizes them in the fashion of a Word Cloud).

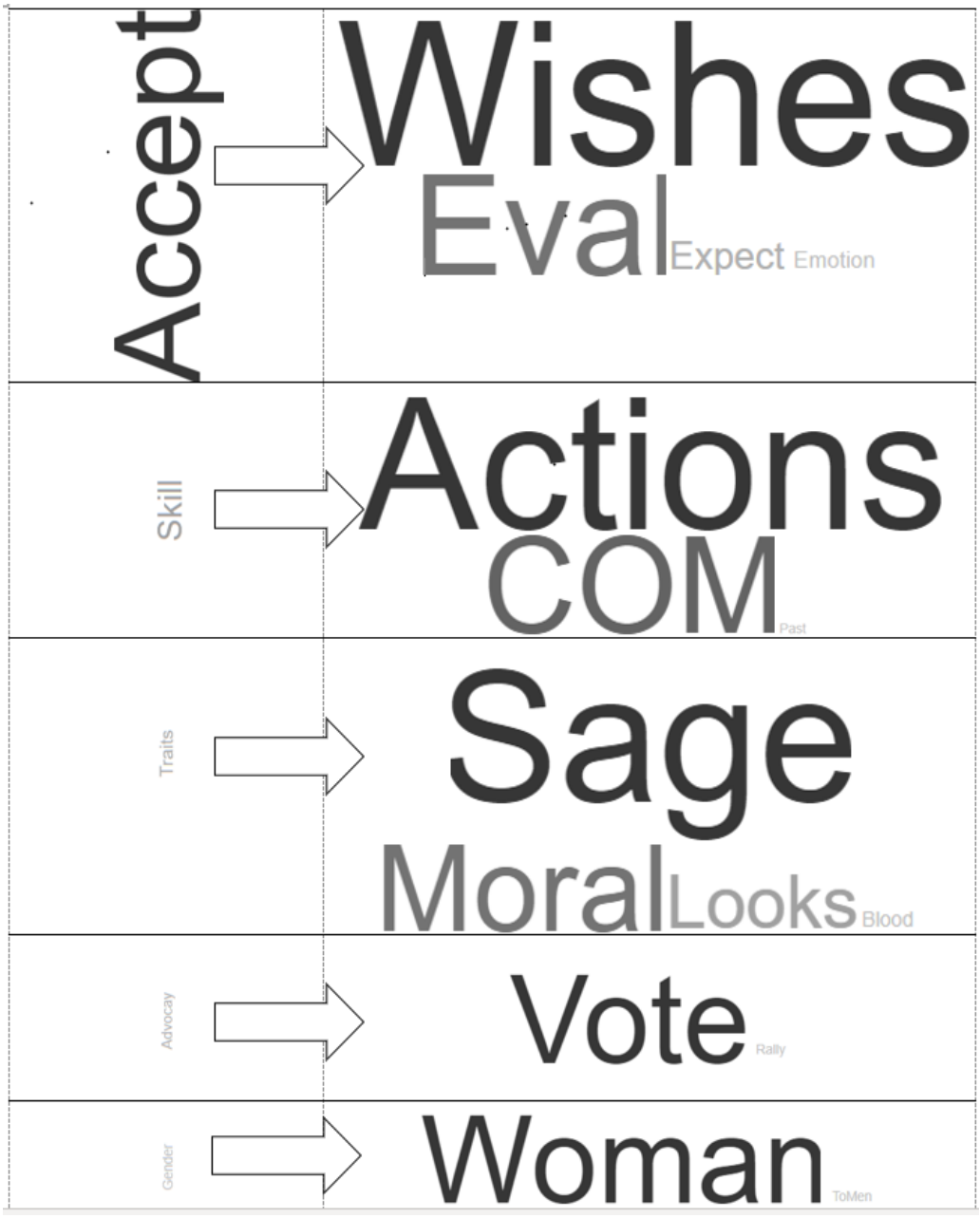


Figure 2. Visual Representation of the Weights of Positive Tweet Themes (Left) and Subthemes (Right)

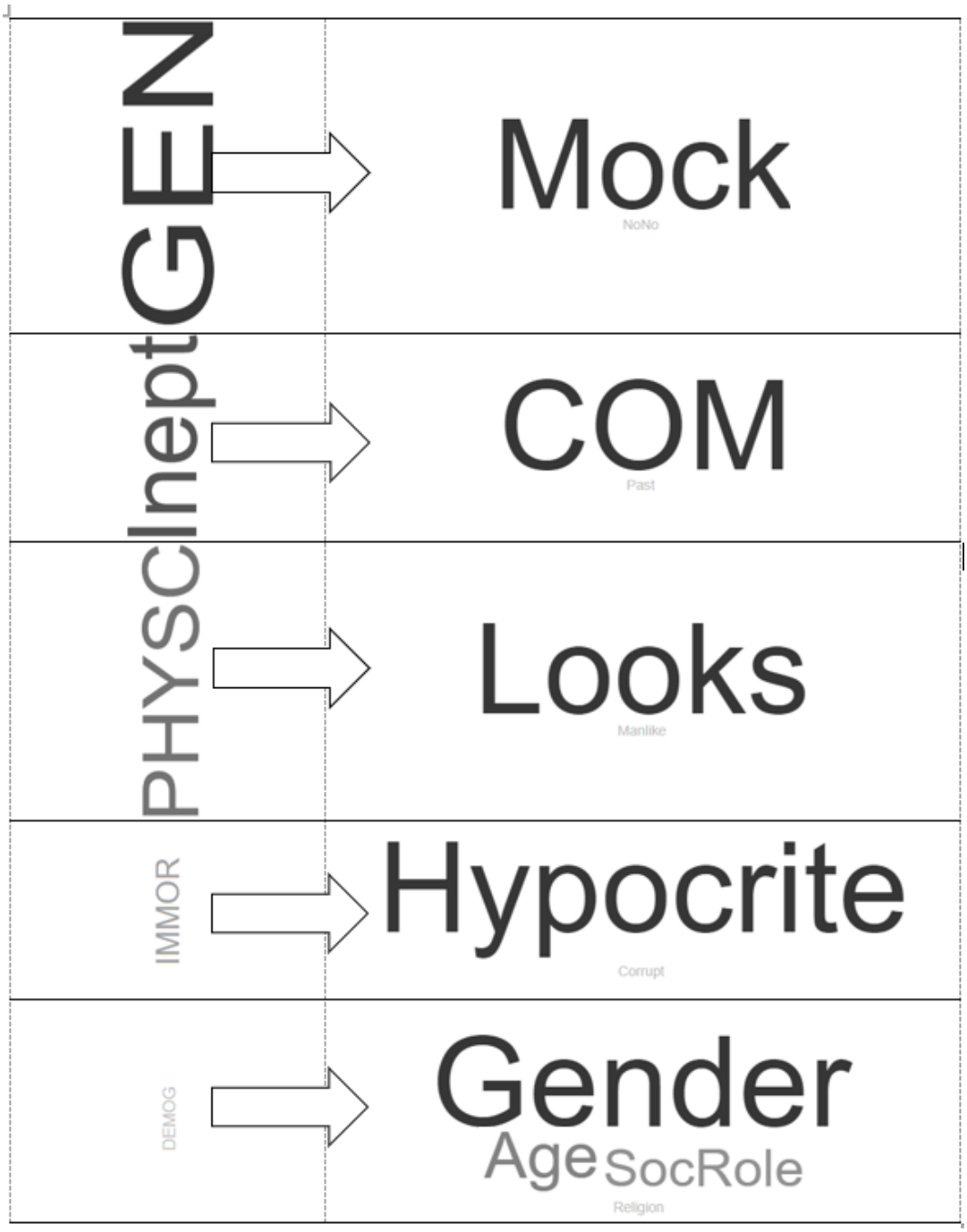


Figure 3. Visual Representation of the Weights of Negative Tweet Themes (Left) and Subthemes (Right)

5.2.1 Positive Themes

The components of the positive attitudes expressed did not differ significantly between men and women, except for the use of emotion. Although the emotional component of the "general acceptance" theme was minimal, women were the main generators of this component. Six out of the seven emotional attitudes towards WCs were posted by women tweeters. This concurs with research done in the West,

which shows that women are more prone on social media to using terms that convey affection, whereas men tend to use terms that assert their power (Bamman et al., 2014; Ottoni, Pesce, Las Casas, Franciscani, & Meira, 2013).

The major theme of the positive attitudes towards WCs is too general, with a minimal cognitive component. Tweeters with positive attitudes primarily expressed their "best wishes" for a candidate by saying, "may Allah grant her victory" (Wish), expecting them to win (Expect), evaluating them in broad strokes, such as "she is the best" or "she deserves to win" (Eval), and expressing an emotional disposition, such as "I love her" and "happy she is running" (Emotion). This dominant positive theme reflects a subjective, tweeter-based general acceptance of a candidate rather than a candidate-based acceptance. Tweeters favored candidates out of love, support, and/or good wishes for them. All these are subjective favorability criteria that are not based on any specific objective traits of the WC.

The second dominant theme, however, did represent an objective, candidate-based acceptance, in which tweeters were favorable toward a candidate because of her political performance (Actions) and communication (COM) skills. Those skills of a WC have been demonstrated ether in a previous political job or during a previous parliamentary term. They were also expected to be demonstrated after the WC won the election. Satisfaction with both past political performance and communication skills were major subthemes of this positive theme. The tweeters placed a special emphasis on the communication skills of the candidates, praising them for their "eloquence," "sophistication as speakers," and "cogent logic while talking to the media." Some tweeters described the speaking style of one candidate as "making the liver of the listener cooler" (a Kuwaiti colloquial expression connoting a good mood and a positive reaction).

The third positive theme relates to how tweeters evaluated the personal traits of WCs. The most prevalent subtheme under this theme was the perceived sagacity and knowledge of a WC (Sage), followed by the moral traits of the WC, such as being "modest," "patriotic," and/or "religious." One tweeter addressed a WC by tweeting, "I wish you an early win of a seat; your mind is astute, and you have solutions and administrative skills." Other tweeters (especially women) favored the good looks of some WCs or praised them for being young and bringing new blood to the House.

The fourth attitudinal theme pertained to general advocacy for WCs, as several tweeters posted vows of voting for a WC or calls urging users to vote for a specific candidate (Rally) and/or to watch for her media and YouTube appearances. This reflects an active, positive attitude towards WCs in the hopes of garnering more support for them and helping them gain wider exposure in the public eye.

Gender-based tweets represent the least prevalent positive theme. Tweeters expressed a positive attitude towards a WC because she is "a daughter of men," "a daughter of Kuwait," "the pride of Kuwaiti women," or "better than many men." One candidate was supported because she was "more honorable than the mustache of her opponent." A man described one candidate as being "better than 50

mustaches" (the mustache used to be a common symbol of manhood in Arab culture). This emphasis on the superiority of WCs was not observed in the tweets posted by women. A smaller percentage of their positive tweets maintained that certain candidates were the best representatives of women.

5.2.2 Negative Themes

For the most part, the negative tweeted attitudes towards WCs employed a general (Gen) theme of sarcasm and vowing not to elect a WC or calling on others not to choose her (NoNo). Some tweeters posted reactionary non-word cues to ridicule a candidate by tweeting a political statement of hers and commenting on it by using an expression that mimicked the sound of laughter. The following tweets represent this theme: "When a TV show host asks candidate Awatef Elkallaf what her goals are if she wins in the elections, she replies, "When I win, I will think about that." Loooool." Other tweeters ridiculed some WCs for looking so helpless or even having hope of winning. One tweeter said mockingly, "Hope in life is sweet, and she [a candidate] will win 50 years after her death."

As illustrated in Figure 3, the second dominant negative theme was directed towards the perceived ineptitude of WCs by using a more specific approach. This included under-evaluating their communication skills (COM) or terming their speeches "too loud," "too dramatic," "too arrogant," or "signifying dumb self-confidence." One tweeter suggested that a WC needs to receive some tutoring to master the Kuwaiti colloquial. Other tweeters doubted the political competency of some WCs because of their "failed" political past. A few tweeters even wished that women's suffrage had never been granted in Kuwait. They also wished that a specific WC would never win the election.

WCs were also the targets of negativity based on their physical traits. This third negative theme had two main subthemes: one that critiqued the general appearance of the WCs (Looks), and another that lashed out at them because they looked like men (Manlike). Accordingly, some WCs were called "ugly," "junk-like," and even "dog-like." One woman wondered why all WCs "were so unattractive!" A video showing the false teeth of a candidate falling out during her televised interview was the object of mockery by some tweeting men. Others joked that the candidate belonged in a junkyard, called her "Ms. Slippers," or inquired about the "Pakistani barber who does her hair" (in Kuwaiti culture, Pakistani barbers are perceived as having the cheapest saloons, catering to clients of lower social status). One candidate was described as "having the mouth of an ugly Indonesian domestic servant" (Asian female domestic workers are a source of cheap labor found in most Kuwaiti households).

In the same vein, some negative tweets referred to WCs as having a man-like appearance or as being "manly" women. One candidate was called "Rambo" or "brother" and was ridiculed for looking like a famous Egyptian actor who is known for his film roles as a bully and a thug. A few tweeters even suggested that the same

candidate needed to "trim her mustache." Other tweeters said that they were "having a problem identifying a WC among men." They suggested that men who vote for one WC will not incite the jealousy of their wives since that candidate can hardly be considered a woman.

The fourth negative theme pertains to some moral evaluations of WCs. A few of these tweets cast doubt on the morality of the actions and political statements of some WCs. Such tweets generally label some candidates as "corrupt" and "lacking integrity." Others referred to them as being hypocrites, inconsistent, and fake, such as, "Never heard her talk about buying out votes in her constituency," "You are just words, no action," and "What about your son getting a job with a tempting salary?"

Notwithstanding, the least common negative theme pertains to the demographics of the WCs: their gender roles, religion, age, and social status. Some tweeters suggested that "grocery shopping," "staying home," "taking care of the kids," and "cooking" should be the more appropriate jobs for these WCs. One male tweeter thought that women make a crucial contribution to the fabric of society "because what they do is just give birth." Another tweeter imagined a winning WC "taking her stove to the House." Some men expressed their negative attitudes towards the ethnicity and religious sects of a specific WC. One Shiite WC was described as being more loyal to (Shiite) Iran than to (Sunni) Kuwait. One tweeter said about one WC that "her name is Safavid," (referring to the Shiite Dynasty that ruled Iran for centuries). A few tweeters criticized a candidate for wearing "non-Islamic" clothes. Others urged unveiled candidates to cover up, and veiled candidates to put on their *hijab* (head cover) properly. Age and social status were also two components of the negative "demographics" theme. Some tweeters, especially men, thought that certain WCs were "old hags" who should give youth a chance. Other tweets read, "I wish my auntie would withdraw and give youth a chance; she is a grandma born in 1946." "You are an oldie from hell and one of the setbacks of Kuwait." Several candidates were also called "Aunties" and were demeaned as "cheap" or coming from a lower social stratum or ridiculed as being "less sophisticated Bedouins."

6 DISCUSSION

The results partially agree with previous research in showing that Twitter users frequently tend to express sarcastic, negative attitudes towards WCs (Mejova et al., 2013) and that gender determines political attitudes (e.g., Sidanius, 1997a; Sidanius & Pratto, 1999; Southern & Harmer, 2021). The analysis shows that women have more positive and less negative attitudes towards WCs; this is also in line with the literature, which shows that women tend to possess more positive attitudes towards the issue of leadership by women than men (Adeyemi-Bello & Tomkiewicz, 1996; Al-Khoori, 2011; Alibeli, 2015; and Mostafa, 2010). However, this association between gender and attitudes was only significant in the election of 2013. In the subsequent two elections, no statistically significant association between gender and

attitudes was detected. The results suggest that gender is not expected to remain a significant factor in predicting the attitudes of Kuwaiti tweeters towards WCs. The results also show the importance of adopting a longitudinal to examine the association between gender and political attitudes in patriarchal societies.

A closer examination of the components of the tweeted attitudes towards WCs in Kuwait over more than one election term shows that these attitudes took a "cognitive" turn, especially in the latest election in 2020. In the first election, both positive and negative tweeted attitudes towards WCs included primarily generic terms, such as expressing good wishes, welcoming the nomination of a candidate, expecting their victory, or vowing not to vote for a WC. In the following two elections, those attitudes turned out to be more elaborate cognitively. Tweeters became more cognizant of the performance of a specific WC in parliament and more critical of both the communication skills of WCs and their political statements. Tweeters also reflected more on the integrity and credibility of WCs in the last election. Even though the credibility of the source has been studied as a peripheral cue in the literature (Wu & Schaffer, 1987; Zhu, Xie, & Gan, 2011), it is an indication that tweeters pay more attention to the objective cues in WCs (such as their political performance) than their subjective cues, such as their looks and appearance. Several tweeters announced their support of specific WCs because of their patriotic political speech that "reveals corruption" and "focuses on development." Tweeters tended to show a growing interest in the message of the WCs over time and a decreasing interest in their physical, and gender-based characteristics. In other words, the longer the WCs maintain their civic engagement, the less likely they will be evaluated according to gender-based criteria or "othered" by the socially dominant groups.

This increasing prevalence of "objective" evaluation of WCs in the tweets explains the significant growth in the length of those tweets during the second and third elections. The results showed that the tweeted attitudes posted in 2020 were significantly longer (in terms of characters) compared to those posted in either 2013 or 2016. Several tweeted attitudes towards WCs in 2013 consisted of a few words or sometimes one word to express either general good wishes or negative descriptions of a WC. Although the tweeted attitudes posted in 2016 were longer than those posted in 2013, this difference was not statistically significant. However, the length of the tweeted attitudes in 2020 was significantly higher than the posts in both 2013 and 2016. This difference is hardly attributed solely to the growing interest of tweeters in the political communication and performance of WCs. This could also be attributed to the decision made by Twitter in November 2017 to increase the maximum allowable length of tweets from 140 to 280 characters. Accordingly, tweeters in 2020 could abandon this "telegraphic" writing style dictated by the maximum limit of 140 characters and ramble more freely after this maximum has been doubled. However, a second examination of the data reveals another factor that might explain the varying lengths of tweets across the three elections. The two WCs who ran in the three elections (Safaa and Anwar) received

unequal lengths of tweets. Figure 4 shows that Safaa, who won in both 2013 and 2016, had a gradual increase in the length of the tweeted attitudes towards her. Anwar never made it to Parliament in any election, and the length of the tweets posted about her took an irregular pattern.

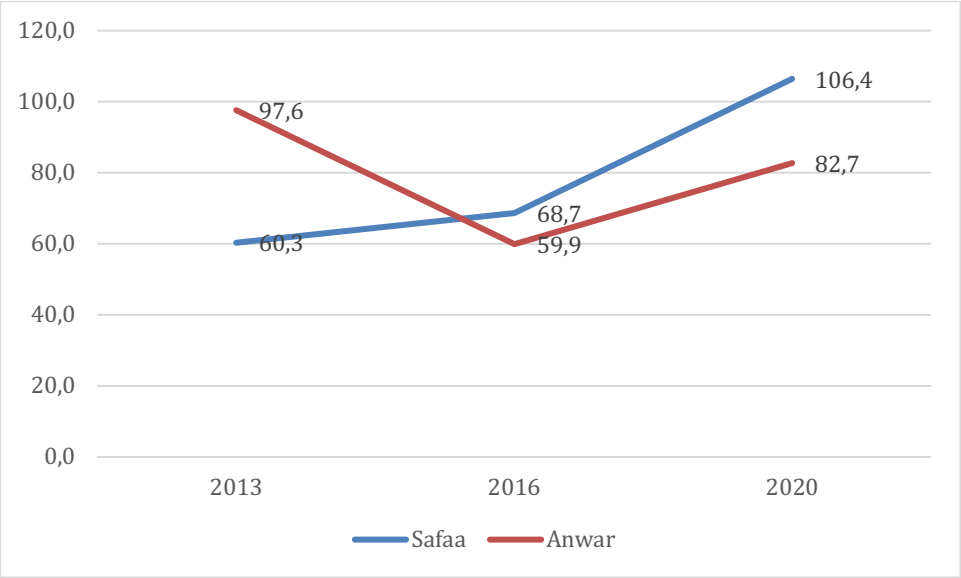


Figure 4. Average No. of Characters in Tweeted Attitudes towards Two WCs in Three Elections

The length of both negative and positive tweeted attitudes towards Safaa (a PM for two terms) in 2016 (one year before Twitter increased the maximum length of tweets from 140 to 280 characters) was significantly larger than in 2013 ($t = 6.0$, adj sig. = 0.04). This seems to demonstrate that the formal, continued civic engagement of this WC (rather than the length constraint of tweets) triggered more central processing of her political communication, which was reflected in the longer, information-based tweeted attitudes about her from one election to the subsequent. The other candidate (Anwar) had intermittent exposure to the public eye, mainly during the election campaigns, which limited the amount of information available about her. On the other hand, research comparing the impact of Twitter's move to increase its length constraint on the linguistic aspects of tweets found that tweets constrained by a 140-character limit contained fewer hashtags, more abbreviations, and more symbols (Gligorić et al., 2018), but the authors did not report any significant increase in the quantity or quality of opinions expressed. This suggests that the continued political engagement of WCs determines both the length and cognitive elaboration of the tweeted attitudes towards them.

Those tweeted attitudes were also thematically rich. The data used for the present study provides limited information about the tweeters. Though there is no reliable way to measure their SDO levels, their tweets reveal how they perceived and assessed WCs. Those WCs were substantially supported based on general,

affective reactions to them. Among the five themes that summarized tweeted positive attitudes, only one pertained to their skills in politics and communication, and yet that one was not the most dominant theme. A few male tweeters supported WCs because they would be "the best to represent the women of Kuwait." This implies that WCs, if elected, will not topple the long-standing societal power hierarchy existing in the country, but will have a "segregated" position of power to represent their fellow female voters effectively. Other men supported specific WCs because they believed that they acted in as "manly" a fashion as their male counterparts. To justify their public support for those candidates, those tweeters had to redefine the WCs to perceive them as acting similarly to men; this suggests that WCs were (subconsciously perhaps) perceived as belonging to the powerful, dominant level of the social hierarchy in order for them to be supported.

This perception might have neutralized the inclination of individuals high in SDO to deny women a higher social status and/or leadership roles. The acceptance of WCs by male tweeters on the basis that those candidates are compared favorably to men suggests a subconscious legitimization of the consensually shared belief about male dominance/superiority in Kuwait. Several male tweeters even expressed their support for a WC because they felt she had been "victimized" in Parliament when she was previously serving in it. Accordingly, it is not just gender, age, clan, ethnicity, and social class that determine SDO, as was suggested by Sidanius and Pratto (1999), but also perceptual and cultural processes that explain how the dominant gender could end up supporting politically the opposite, subordinate gender. In this light, the tweets favoring WCs based on feminism-related ideas were limited even among female tweeters. This finding is not in line with previous research, which shows that women are more likely to vote for WCs who are perceived as "profeminist" (Plutzer and Zipp, 1996; Campbell & Heath, 2017). As some tweeters have suggested, WCs in patriarchal societies are expected to look feminine, but not to adopt a feminist agenda. To the contrary, they are praised on the basis that they act more "manly" than their male counterparts. This dilemma facing women leaders seems to be universal. Masculinity and power seem to be perceived as concomitant. Mavin (2009) argued that women leaders who seek to be credible will not identify with the gender stereotype of women who are stereotypically linked to both appearance and materialistic artifacts, and not to ideas. The media reconstructions of women politicians significantly criticize their appearance and sometimes lash out at them for looking "too sexy" or "too feminine" (Mavin, Bryans, & Cunningham, 2010).

The data also shows that tweeters tend to reject WCs in Kuwait because of who they are (subjective cues) or how they act or communicate (objective reasoning). In return, tweeters accept WCs mainly because of how they feel about those candidates (subjective reasoning). This seemingly unjustified, disengaged public support represents "low-commitment active citizenship" (Stern, Dietz, Abel, Guagnano, & Kalof, 1999, p. 82). Given this weak commitment, the chances for attitudes to be reflected in behavior are not high. With less motivation and less

elaborate cognitive processing of the "qualities" of WCs, attitudes are less expected to be translated into behavior, viz., voting.

The "social role" component of the "demographics" negative theme is supported by the stipulations of SDO. Some tweeters identified the role of family care-taker as an alternative role for WCs. Although women in patriarchal societies are still oppressed by such stereotypical depictions, some tweeting women take part in this societal oppression. Voters associated care-giving attributes, nurturing roles, and appearance-related traits with the WCs. In doing so, they have subconsciously dissociated WCs from the attributes of competence and leadership and tended to maintain the long-standing hierarchy in society. This unspoken agreement between superior and inferior genders over the social and occupational roles of women represents another legitimizing force for the "stereotypes and cultural ideologies" (Pratto et al., 2006, p. 275) in Kuwaiti society.

The negative theme of demographics also had a religious component. A few tweeters rejected some candidates based on their affiliation with the minority religious caste in the country. Other tweeters criticized WCs for either not wearing the Islamic head covering, or not wearing it properly. Kuwait falls within the Muslim "patriarchal belt" that extends from North Africa through the Muslim Middle East and to South and East Asia (Offenhauer, 2005). While gender segregation in all stages of education is legally enforced in Kuwait, wearing a veil is not obligatory. A few tweeters rejected the potential political leadership of these WCs based on religious teachings that might deny such a leadership. The Islamic religion has been suggested to be the determining factor in the division of labor and the acceptance of gendered norms and beliefs (Norris, 2011). However, this argument ignores the social and cultural individuality found among various Islamic populations. For instance, anthropologist Mounira Charrad (2001) analyzed the status of women in the Arab nations of Morocco, Algeria, and Tunisia and reported that women's legal rights and political participation in these countries are determined more by kinship relations than by the Islamic faith. Yet, this is not to say that religion does not remain a potentially significant factor in the determination of leadership and power relationships within traditional, patriarchal societies. To examine this issue more fully, we need a more detailed empirical test of the relationship between religiosity and SDO.

Overall, the political tweets about WCs in Kuwait reveal the complexity of the attitudes towards them and explain some of the gender differences and similarities in accepting or rejecting them. Those political tweets have also shown that they might be used as a predictor of power in Kuwait, as the two WCs who won the election in 2013 and the WC who won in 2016 also received the highest number of mentions on Twitter, compared to the other WCs. This replicates some findings in the previous literature (O'Connor et al., 2010; Tumasjan et al., 2010).

7 CONCLUSION & IMPLICATIONS FOR FUTURE RESEARCH

The present study shows that gender partially determines political and social attitudes and explains how dominant groups may undergo perceptual modifications that will allow them to support the political leadership of the subordinate group. The political longivity and civic engagement of women in this patriarchal culture is expected to normalize their acceptance as candidates rather than "other" them based on their gender. The results also underscore the important role of attitudes in explaining gender inequality. The dominant perception of WCs within this tech-savvy, educated Kuwaiti citizenry is that these candidates are mainly women whose first obligation is their family and whose most important trait is demographic, rather than intellectual. Consequently, women leaders in patriarchal societies are praised for acting like men. Such long-established, basic inequality in this traditional society has maintained the distribution of power and reinforced its acceptance, even by less powerful groups. In order for WCs to be accepted by those voters who have higher levels of SDO, they must identify with the dominant hierarchy rather than present themselves as advocates for feminist rights and gender equality. If they wish to seek leadership roles, they should not act as feminists.

More components of SDO could be explored by studying transitional societies that are passing through modernization processes while still maintaining their long-standing cultural and religious foundations. Such components might prove more consequential in predicting and explaining ingroup inequality and power distribution. These components include, but are not limited to, the use of information technology, study abroad, interaction with different cultures through travel, and the adoption of liberal ideas. An understanding of the impact of such dynamics will contribute to the degree to which the social dominance theory may be generalized and applied to understanding intergroup relations within a dynamic, modernizing society.

8 LIMITATIONS

As with any content analysis study, the research point, variables, and analysis are restricted by the content of the tweets in the present study. Innately, tweets are "telegraphic" forms of expression that are limited by a specific number of characters, which further constrains the depth of information and reflection shared. This data-related limitation makes it inevitable for the researcher to analyze the tweeted attitude in isolation from the demographic and psychographic traits of the tweeters. Significant variables that might predict attitudes towards the leadership of women in Kuwait are tribal and religious affiliations. It was hard to decipher such variables from the data used in the present study, which has limited the researcher's ability to discuss and explain the results from a broader perspective. Even though the theory of social dominance was adopted as a theoretical framework, the data inherently did not provide information that could be used in measuring the

tweeters' levels of social dominance, which has represented another limitation for the study. Moreover, the lack of previous studies in the same field conducted in Kuwait restricted the scope of the present study.

REFERENCES

Abdalla, I. A. (1996). Attitudes towards women in the Arabian Gulf region. *Women in Management Review, 11*(1), 29-39.

Adeyemi-Bello, T., & Tomkiewicz, J. M. (1996). The attitudes of Nigerians toward women managers. *Journal of Social Behavior and Personality, 11*(5), 133-139.

Alibeli, M. A. (2015). Gender and attitudes toward women in the United Arab Emirates. *Perspectives on Global Development and Technology, 14*, 109-125. DOI: 10.1163/15691497-12341335

Al-Khoori, M. A. (2011, 12 5). Attitudes towars Females in the UAE. Dubai, UAE. Retrieved April 1, 2020, from http://bspace.buid.ac.ae/bitstream/1234/337/1/80037.pdf

Amawi, A. (2007). Against all odds: Women candidates in Jordan's 1997 elections. In V. M. Moghadam, *From patriarchy to empowerment: Women's participation, movements, and rights in the Middle East, North Africa, and South Asia* (pp. 40-57). Syracuse, New York: Syracuse University Press.

Amawi, A. M. (2001). *Against all odds: Jordanian women, elections and political empowerment.* Amman: Al Kutbah Institue of Human Development and Konrad-Adenauer Stiftung.

Arab Social Media Report. (2017). *The 7th Arab Social Media Report: social media and the Internet of Things - Towards Data-Driven Policymaking in the Arab World: Potential, Limits, and Concerns.* Dubai School of government, Dubai. Retrieved September 1, 2021, from https://www.arabsocialmediareport.com/Twitter/LineChart.aspx?&PriMenuID=18&CatID=25&mnu=Cat

Bamman, D., Eisenstein, J., & Schnoebelen, T. (2014). Gender identity and lexical variation in social media. *Journal of Sociolinguistics, 18*(2), 135–160. https://doi.org/10.1111/josl.12080

Bier, L. (2010). "The Family Is a Factory": Gender, Citizenship, and the Regulation of Reproduction in Postwar Egypt. *Feminist Studies*, 36(2), 404-432. DOI:10.2307/27919108

Boutet, A., Kim, H., & Yoneki, E. (2012). *What's in your Tweets? I know who you supported in the UK 2010 general election*. Dublin, Ireland: Proceedings of Sixth International AAAI Conference on Weblogs and Social Media (ICWSM-12).

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*, 77–101. https://doi.org/10.1191/1478088706qp063oa

Burger, J. D., Henderson, J., Kim, G., & Zarrella G, G. (2011). Discriminating gender on Twitter. Pennsylvania , USA: Proceedings of the Conference on Empirical Methods in Natural Language Processing.

Campbell, R., & Heath, O. (2017). Do women vote for women? Attitudes towards descriptive representation and voting behavior in the 2010 British election. *Politics & Gender, 13*, 209 –231. DOI: 10.1017/S1743923X16000672

CARE International Policy Report (2013). *Arab Spring or Arab Autumn? Women's political participation in the uprising and beyond.* Retrieved December 12, 2020, from https://www.care.org/wp-content/uploads/2020/05/report_women-arab-spring_english-2013.pdf

Central Statistical Bureau (CSB) (2019-2020). Annual statistical abstract, Edition 54. Retrieved on November 30, 2021, from https://www.csb.gov.kw/Pages/Statistics?ID=18&ParentCatID=2

Ceron, A., Curini, L., Iacus, S. M., & Porro, G. (2014). Every tweet counts? How sentiment analysis of social media can improve our knowledge of citizens' political preferences with an application to Italy and France. *New Media & Society, 16*(2), 340–358. https://doi.org/10.1177/1461444813480466

Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis.* London: Sage Publications.

Charrad, M. M. (2001). *States and Women's Rights: The Making of Postcolonial Tunisia, Algeria, and Morocco.* Berkeley: University of California Press.

Cohen, R., & Ruths, D. (2013). Classifying political orientation on Twitter: It's not easy. Boston, USA: Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media (ICWSM-13). Retrieved August 28, 2015, from http://ssc.sagepub.com.ezproxy.rollins.edu:2048/content/31/6/649.full.pdf+html

Conover, M. D., Ratkiewicz, J., Francisco, M., Goncalves, B., Flammini, A., & Menczer, F. (2011). *Political polarization on Twitter. Barcelona, Spain*: Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media (ICWSM-11).

Conway, B. A., Kenski, K., & Wang, D. (2015). The rise of Twitter in the political campaign. Journal of Computer-Mediated Communication, 20, 363-380. Doi:10.1111/jcc4.12124

Cunha, E., Magno, G., Gonçalves, M. A., Cambraia, C., & Almeida, V. (2014). He votes or she votes? Female and male discursive strategies in Twitter political hashtags. *PloS One, 9*(1). Retrieved June 18, 2021, from http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0087041

Galanaki, E., Papalexandris, N., & Halikias, J. (2009). Revisiting leadership styles and attitudes towards women as managers in Greece. *Gender in Management, 24*(7), 484-504. DOI:10.1108/17542410910991782

Gallant, M., & Pounder, J. S. (2008). The Employment of Female nationals in the United Arab Emirates (UAE): An analysis of opportunities and barriers. *Education, Business and Society: Contemporary Middle Eastern Issues, 1*(1), 26-33.

Gligorić, K., Anderson, A., & West, R. (2018). How Constraints Affect Content: The Case of Twitter's Switch from 140 to 280 Characters. *Proceedings of the International AAAI Conference on Web and Social Media*, *12*(1). Retrieved October10, 2021, from https://ojs.aaai.org/index.php/ICWSM/article/view/15079

Gul, S., Mahajan, I., Nisa, N. T., Shah, T. A., Asifa, J., & Ahmad, S. (2016). Tweets speak louder than leaders and masses: An analysis of tweets about the Jammu and Kashmir elections 2014. *Online Information Review, 40*(7), 900-912. doi: http://dx.doi.org.ezproxy.rollins.edu:2048/10.1108/OIR-10-2015-0330

Hu, L., & Kearney, M. W. (2021). Gendered Tweets: Computational text analysis of gender differences in political discussion on Twitter. *Journal of Language and Social Psychology*, *40*(4), 482–503. https://doi.org/10.1177/0261927X20969752

Inglehart, R., & Norris, P. (2003). *Rising tide: Gender equality and cultural change around the world.* New York and Cambridge: Cambridge University.

Inter-Parliamentary Union (IPU). (2015, March 10). Sluggish progress on women in politics will hamper development. Geneva: IPU. Retrieved March 11, 2021, from http://www.ipu.org/press-e/pressrelease201503101.htm

Inter-Parliamentary Union. (2021). *Global and regional averages of* w*omen in national parliaments*. Retrieved September 24, 2021, from https://data.ipu.org/women-averages.

Kauser, S., & Tlaiss, H. (2011). Middle Eastern women managers: Participation, barriers, and future prospects. *Journal of International Business and Economy, 12*(1), 35-56.

Khimish, H. A. (2014). The impact of religion on Arab women. *International Journal of Business and Social Science, 5*(3), 132-142.

Larsson, A. O., & Moe, H. (2011). Studying political microblogging: Twitter users in the 2010 Swedish election campaign. *New Media & Society, 14*(5), 729–747. https://doi.org/10.1177/1461444811422894

Madsen, S. R. (2010). The experiences of UAE women leaders in developing leadership early in life. *Feminist Formations, 22*(3), 75-95. DOI:10.1353/ff.2010.0014

Martin, G. (2013, May). Instrument or structure? Investigating the potential uses of Twitter in Kuwait. Guelph, Ontario: The University of Guelph.

Retrieved November 10, 2014, from https://atrium.lib.uoguelph.ca/xmlui/bitstream/handle/10214/6676/Martin_Geoff_201305_MA.pdf?sequence=1

Mavin, S. (2009). *Gender stereotypes and assumptions: popular culture constructions of women leaders*. Gdansk, Poland: The 10th International Conference, HRD Development, Research and Practice across Europe: Complexity and Imperfection in Practice.

Mavin, S., Bryans, P., & Cunningham, R. (2010). Fed-up with Blair's babes, Gordon's gals, Cameron's cuties, Nick's nymphets: Challenging gendered media representations of women political leaders. *Gender in Management, 25*(7), 550-569. DOI:10.1108/17542411011081365

Mejova, Y., Srinivasan, P., & Boynton, B. (2013). GOP primary season on Twitter: 'Popular' political sentiment in social media. Rome, Italy: WSDM'13. Retrieved August 22, 2021, from https://dollar.biz.uiowa.edu/~street/wsdm73-mejova.pdf

Minot JR, Arnold M. V., Alshaabi, T., Danforth C. M., Dodds, P. S. (2021) Rationing the President: An exploration of public engagement with Obama and Trump on Twitter. *PLoS ONE*, 16(4): e0248880. https://doi.org/10.1371/journal. pone.0248880

Moncanu, D., Baronchelli, A., Perra, N., Gonçalves, B., Zhang, Q., & Vespignani, A. (2013, April 18). The Twitter of Babel: Mapping world languages through microblogging platforms. *8*. PLoS *ONE*. Retrieved September 10, 2021, from http://www.plosone.org/article/fetchObject.action?uri=info%3Adoi%2F10.1371%2Fjournal.pone.0061981&representation=PDF

Mostafa, M. M. (2005). Attitudes towards women managers in the United Arab Emirates: The effects of patriarchy, age, and sex differences. *Journal of Managerial Psychology, 20*(6), 522-540. https://doi.org/10.1108/02683940510615451

Mostafa, M.M. (2005). Attitudes towards women managers in the United Arab Emirates The effects of patriarchy, age, and sex differences. *Journal of Managerial Psychology*, 20(6), 522-540. Doi:10.1108/02683940510615451

National Academies of Sciences, Engineering, and Medicine (2020*). The Inclusion of Women in STEM in Kuwait and the United States*: Proceedings of a Workshop. Washington, DC: The National Academies Press. https://doi.org/10.17226/25820.

Norris, Pippa. (2011). Mecca or oil? Why Arab states lag in gender equality. Irvine, California, USA: Harvard University. Retrieved February 20, 2021, from http://www.hks.harvard.edu/fs/pnorris/Acrobat/Irvine%20Inglehart%20Festschrift%20Mecca%20or%20Oil%20%20Paper%20Norris.pdf

Offenhauer, P. (2005). *Women in Islamic societies: A selected review of social scientific literature.* Washington D.C.: Library of Congress. Retrieved May 17, 2021, from https://www.loc.gov/rr/frd/pdf-files/Women_Islamic_Societies.pdf

Oh, J., & Sundar, S. S. (2015). How Does Interactivity Persuade? An Experimental Test of Interactivity on Cognitive Absorption, Elaboration, and Attitudes. Journal of Communication, 65(2), 213–236. https://doi-org.ezproxy.rollins.edu/10.1111/jcom.12147

Ottoni, R., Pesce, J. P., Las Casas, D., Franciscani, G., & Meira, W. (2013). *Ladies first: Analyzing gender roles and behaviors in Pinterest*. Boston, USA: Proceedings of the 7th International AAAI Conference on Weblogs and Social Media.

Pettygrove, M. W. (2006, May 11). *Obstacles to women's political empowerment in Jordan: Family, Islam, and patriarchal gender roles*. Jordan, Jordan: School for International Training. Retrieved March 1, 2021, from http://digitalcollections.sit.edu/cgi/viewcontent.cgi?article=1362&context=isp_collection

Plutzer, E., & Zipp, J. (1996). Identity politics, partisanship, and voting for women candidates. *Public Opinion Quarterly, 60*(1), 30–57. https://doi.org/10.1086/297738

Pratto, F., Sidanius, J., & Levin, S. (2006). Social dominance theory and the dynamics of intergroup relations: Taking stock and looking forward. *European Journal of Social Psychology, 17*, 271 – 320. https://doi.org/10.1080/10463280601055772

Rao, A., Morstatter, F., Hu, M., Chen, E., Burghardt, K., Ferrara, E., & Lerman, K. (2021). Political Partisanship and Anti-science attitudes in online discussions about COVID-19: Twitter content analysis. *Journal of Medical Internet Research, 23(6),* e26692. Doi: 10.2196/26692

Rodríguez-Ibáñez, M., Gimeno-Blanes, F. J., Cuenca-Jiménez, P. M., Soguero-Ruiz, C. & Rojo-Álvarez, J. L. (2021) Sentiment analysis of political tweets from the 2019 Spanish elections. *IEEE Access*, 9, 101847-101862. Doi: 10.1109/ACCESS.2021.3097492.

Ross, M. (2008). Oil, Islam, and Women. *American Political Science Review, 102*(1), 107-123. https://doi.org/10.1017/S0003055408080040

Scott, J., Dakin, R., Heller, K., & Adriana, E. (2013). *Extracting lessons on gender in the oil and gas sector.* The World Bank. Retrieved June 10, 2021, from http://documents.worldbank.org/curated/en/266311468161347063/pdf/798940NWP0280E0Box0379795B00PUBLIC0.pdf

Shanahan, R. (2009). *Women in Arab Politics-Part 2.* Retrieved October 1, 2021, from https://archive.lowyinstitute.org/the-interpreter/women-arab-politics-part-2

Sidani, Y. (2005). Women, work, and Islam in Arab societies. *Women in Management Review, 20*(7), 498 - 512. http://dx.doi.org/10.1108/09649420510624738

Sidanius, J., Pratto, F., Laar, C., & Levin, S. (2004). Social dominance theory: Its agenda and method. *Political Psychology, 25*, 845 – 880. https://psycnet.apa.org/doi/10.1111/j.1467-9221.2004.00401.x

Simmons, A. L., Duffy, J. A., & Alfraih, H. S. (2012). Attitudes toward women managers: The influence of social dominance orientation and power distance on men in college. *Gender in Management, 27*(7), 482-498. https://doi.org/10.1108/17542411211273469

Statista (2016). *Distribution of tweets per user per day from the Middle East and North Africa in March 2016, by country*. Retrieved September 11, 2021, from https://www.statista.com/statistics/729693/mena-tweets-per-user-per-day-by-country/

Stern, P. C., Dietz, T., Abel, T., Guagnano, G., & Kalof, L. (1999). A value-belief-norm theory of support for social movements: The case of environmentalism. *Research in Human Ecology, 6*(2), 81-97.

Southern, R. & Harmer, E. (2021). Twitter, incivility, and "everyday" gendered othering: An analysis of tweets sent to UK members of Parliament. *Social Science Computer Review*, 39(2), 259-275. https://doi.org/10.1177%2F0894439319865519

Tetreault, M. A., & Al-Mughni, H. (1995). Modernization and its discontents: State and gender in Kuwait. *The Middle East Journal, 49*(3), 403-420.

Weedoo. 2017. *Twitter Arab World.* February. Retrieved June 2, 2021, from: https://weedoo.tech/twitter-arab-world-statistics-feb-2017/

Wong, F. M., Tan, C. W., Sen, S., & Chiang, M. (2013). *Quantifying political leaning from tweets and retweets. Boston, USA*: Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM-13).

World Bank. (2021). Labor force, female (% of total labor force) – Kuwait. Retrieved September 24, 2021, from https://data.worldbank.org/indicator/SL.TLF.TOTL.FE.ZS?locations=KW

Worldometer (2017). *GDP per Capita*. Retrieved May 18, 2021, from https://www.worldometers.info/gdp/gdp-per-capita/

Wu, C., & Shaffer, D. R. (1987). Susceptibility to persuasive appeals as a function of source credibility and prior experience with the attitude object. *Journal of Personality and Social Psychology*, *52*(4), 677–688. Doi: 10.1037/0022-3514.52.4.677

Zhu, D., Xie, X., & Gan, Y. (2011). Information source and valence: How information credibility influences earthquake risk perception. *Journal of Environmental Psychology*, *31*(2), 129–136. Doi: 10.1016/j.jenvp.2010.09.005

JOURNAL OF DIGITAL SOCIAL RESEARCH

WWW.JDSR.IO ISSN: 2003-1998

VOL. 4, NO. 1, 2022, 128–151

EXPLORING SIRI'S CONTENT DIVERSITY USING A CROWDSOURCED AUDIT

Tim Glaesener[a]

ABSTRACT

This study aims to describe the content diversity of Siri's search results in the polarized context of US politics. To do so, a crowdsourced audit was conducted. A diverse sample of 170 US-based Siri users between the ages of 18-64 performed five identical queries about politically controversial issues. The data were analyzed using the concept of algorithmic bias. The results suggest that Siri's search algorithm produces a long tail distribution of search results: Forty-two percent of the participants received the six most frequent answers, while 22% of the users received unique answers. These statistics indicate that Siri's search algorithm causes moderate concentration and low fragmentation. The age and, surprisingly, the political orientation of users, do not seem to be driving either concentration or fragmentation. However, the users' gender and location appear to cause low concentration.

Keywords: Siri; content diversity; crowdsourced audit; voice assistants; US politics; politically controversial issues; algorithmic bias; search algorithms; search results; long tail distribution; concentration; fragmentation.

[a] School of Arts and Communication, Malmö University, Sweden.

1 INTRODUCTION

The algorithms of search engines are powerful[1], as they act as gatekeepers to information (Diakopoulos et al., 2018, p. 321). In their function as gatekeepers, they help users to navigate through the web by deciding what information is most relevant to them (Gillespie, 2014, p. 167). By ranking results, these algorithms strongly impact users' attention (Trielli & Diakopoulos, 2019, p. 1), since users tend to click on top results more often than on lower ones (Agichtein et al., 2006, p. 3).

While it is largely unknown how algorithms actually work (Pasquale, 2015, p. 3), it is known that they function in context (Kitchin, 2017, p. 25). For example, algorithms depend on personal data to create situated outcomes (Sandvig et al., 2014, p. 10). As a consequence, personalized ranking algorithms of search engines can provide two users with different results to the same query (Bozdag, 2013, p. 209; Gillespie, 2014, p. 188).

Due to their power to provide users with different information, the algorithms of search engines can be linked to a larger discussion around the role of audience fragmentation in democracies. Some degree of fragmentation is welcome, since it represents a diverse media environment (Fletcher & Nielsen, 2017, p. 477), which helps citizens to form their political opinions (Sunstein, 2002, p. 9). In well-functioning democracies, citizens need to share information in order to engage with it and debate it (Fletcher & Nielsen, 2017, p. 477).

This is in line with the concept of deliberate democracy, which is a type of democracy in which citizens and their representatives are expected to justify relevant decisions. The most important characteristic is that both parties are expected to give reasons for their decisions as well as to respond to the reasons given by others (Gutmann & Thompson, 2004, p. 3). According to Habermas, to participate in deliberative democracy, people need first and foremost to be able to take part in discourse (Olson, 2011, p. 140). A high degree of audience fragmentation is problematic because it might people might end up only being exposed only to political opinions they already agree with. This could, in effect, end up shielding people from other viewpoints, perhaps encouraging people to hold more rigid and extreme positions. In turn, this could "hinder consensus-building in society" (Vike-Freiberga et al., 2013, pp. 27-28).

The issue of algorithmic content selection has become increasingly relevant, so it is not surprising that in recent years, many researchers have been exploring it (Zuiderveen Borgesisus et al., 2016, p. 9) - often relying on Pariser's (2011) disputed concept of 'filter bubbles'. The term 'filter bubble' refers to the idea that personalization of content caused by algorithms can lead to selective exposure (Pariser, 2011). Concerns about such personalization often focus on the possible

[1] There are numerous definitions of the concept of algorithm. In this paper, I use Diakopoulos' (2015) definition that considers an algorithm to be a series of steps undertake to solve a problem or accomplish an outcome (p. 400).

negative effects of filter bubbles on democracy (Zuiderveen Borgesisus et al., 2016, p. 4).

Scholars have come to different and, at times, contradictory findings. Most of them have challenged the idea that search algorithms cause fragmentation. For example, the findings of meta-researchers, i.e., scholars who study the research of others, state that filter bubbles do not constitute a serious danger. With regard to news websites, Zuiderveen Borgesisus et al. (2016) conclude that "there is little empirical evidence that warrants any worries about filter bubbles" (p. 1). More recently, Bruns (2019), who has critically reviewed the idea of the filter bubble, points to empirical evidence that the users of search engines and social media tend to be exposed to more diverse information that non-users (p. 1).

However, it is not only meta-research that has challenged the notion that algorithms cause fragmentation; empirical studies have done so as well. For example, after having analyzed the search results of 187 US Google users (p. 148), Robertson et al. (2018) conclude that there is little evidence of the existence of filter bubbles. Similarly, Nechusthai and Lewis (2019) found that "users with different political leanings from different states were recommended very similar news" (p. 298). Their study involved 168 participants who used their personal Google accounts (Nechusthai & Lewis, 2019, p. 298). Also in the US, Feezell et al. (2021) conducted a nationally representative survey of young adults and one of the general population and found that "neither non-algorithmic nor algorithmically determined news contribute to higher levels of partisan polarization" (p. 1).

Outside the US, scholars have come to similar conclusions. For example, Krafft et al. (2019) have analyzed the search results of more than 4,000 German Google users and discovered hardly any evidence of filter bubbles (p. 1). Similarly, Haim et al. (2018) has researched the personalization of the "content and source diversity of Google News" (p. 330) in Germany. The researchers found minor effects of personalization, concluding that their results indicate that the danger of algorithmic filter bubbles might be exaggerated (Haim et al., 2018, p. 330).

In Belgium, the research findings of Curtois et al. (2018) did not support the existence of filter bubbles as far as social and political information were concerned (p. 2006). However, the scholars coded (in other words: categorized) search results before comparing them (Curtois et al., 2018, p. 2010). In Denmark, Bechmann and Nielbo (2018) researched the Facebook News Feed of 1,000 people. By analyzing link sources and content semantics, they found that roughly 10% to 28% of the sample size was exposed to content that did not overlap. The researchers use this finding as the basis for their critique of the existence of filter bubbles. (Bechmann & Nielbo, 2018, p. 999).

In contrast, other researchers have confirmed the existence of filter bubbles. For example, Barker (2018) has determined that art directors and copywriters are exposed to personalized search results on Google Search (p. 85). His qualitative study does not allow for generalizations, however, as it involved only 18 Australian participants (Barker, 2018, p. 85). Cho et al. (2020) carried out a lab experiment

that included manipulating "user search/watch history" (Cho et al., 2020, p. 150). The researchers prove that "political self-reinforcement" (Cho et al., 2020, p. 150) and "affective polarization are heightened by political videos – selected by the YouTube recommender algorithm – based on the participants' own search preferences" (Cho et al., 2020, p. 150). Theoretically analyzing "twelve different information filtering scenarios" (Geschke et al., 2019, p. 129), Geschke et al. (2019) found that "even without any social or technological filters, echo chambers emerge as a consequence of cognitive mechanisms, such as confirmation bias" (Geschke et al., 2019, p. 129). They further claim that social and technological filtering enhances echo chambers (Geschke et al., 2019, p. 129). The research objects of these studies that confirm the existence of filter bubbles differ greatly from those that refute it.

While the answers of algorithms of graphical and textual search engines have increasingly been researched, the replies of algorithms of voice assistants such as Apple's Siri, Amazon Alexa and Google Assistant have not yet been studied extensively. However, I argue that especially these search algorithms have the potential to shape users' attention. When voice assistants are used to retrieve information from the internet, they typically provide users with only one result to their queries (Dambanemuya & Diakopoulos, 2020, p. 1). In contrast, textual and graphical search engines offer a plurality of results (Natale & Cooke, 2021, p. 1003). In other words, graphical and textual search engines allow users to make a choice among the several results. In contrast, in the case of voice assistants, the choices have been pre-selected for the users, who are presented with only one result. Due to their unique affordances, it seems that the search algorithms of voice assistants are even more powerful gatekeepers than those of other web interfaces.

1.1 Siri in the USA

To address this research gap, I conducted a crowdsourced audit of the search algorithm of the voice assistant Siri[2] in the context of US politics. I have chosen to focus on the USA since voice assistants were used by a critical mass of 51% of the population in 2020 (Valishery, 2021). In addition, there is a high degree of political polarization in the US. (Jurkowitz et al., 2020).

Siri was chosen as a research object for two reasons. Firstly, data suggest that it remains more popular than Google Assistant in the USA (Wagner, 2018; Kinsella, 2020). Although Amazon Alexa is even more popular (Dellinger, 2019), I decided not to focus on this device since smart speakers are communal media (Boothby, 2018), which are typically used in households by more than one user and I focus on content diversity on an individual level. Secondly, while the search

[2] In this study, I use the term 'search algorithm' for the sake of simplicity. I acknowledge, though, that in fact there is not a single search algorithm but only a system of algorithms (Natale & Cooke, 2021, p. 1009). Similarly, I speak of Siri even though this voice assistant is not a real entity but rather a complex system (Seaver, 2013, p. 416).

algorithms of neither Siri nor Google Assistant have been researched sufficiently in relation to content diversity, the search algorithm of Google's textual and graphical search engine has already been investigated (e. g. Haim et al., 2018; Nechusthai & Lewis, 2019). Some of these findings might provide some insight into the workings of Google's voice-based search engine. For this reason, I maintain that there is more need to investigate Siri's search algorithm.

The purpose of this study is to *explore* and *describe* whether Siri answers politically controversial questions of US-based users differently. To do so, I seek to answer the following research question: To what extent does Siri's search algorithm provide 18 to 64-year-old US-based users with different results when asked the same politically controversial questions?

Throughout this study, I refer to the previously mentioned 'different results' as 'unique' and 'personalized' search results. Both terms are used interchangeably to indicate that only one individual has received a certain reply.

Inspired by Fletcher et al. (2020, p. 181), in this study, I refer to the case when users receive non-overlapping content as 'fragmentation' and the case when they receive the overlapping as 'concentration'.

My selection of a diverse group of Siri users lowers the generalizability of this study's results. However, I believe that content diversity, which I define in this study as differences in content as do Krafft et al. (2019, p. 330), can best be studied by using a heterogeneous data set. This view was inspired by Bechmann and Nielbo (2018), who warn scholars about relying on homogenous data sets to investigate differences and similarities in content (p. 994).

1.2 Algorithmic bias

To analyze Siri's search results, I draw on Friedman and Nissenbaum's (1996) concept of 'algorithmic bias'. They argue that computer systems are systematically biased in three ways: pre-existing biases which exist in society (e. g. personal values of software developers) and affect the design process; technical biases which result from the influence of the technology on how computers work; and emergent biases which manifest themselves during the use of a particular software after its release (e. g. through feedback loops) (Friedman & Nissenbaum, 1996, pp. 334–335). As a consequence of these biases, search engines can end up providing users with an "increasingly distorted and limited view of the web" (Introna & Nissenbaum, 2000, p. 54).

While the concept of algorithmic bias seems to be suitable to analyze the outcome of Siri's search algorithm, in one aspect it needs to be adjusted to fit to this study. Friedman and Nissenbaum (1996) distinguish between three types of biases: pre-existing, technical, and emergent (pp. 334-335). The technical bias of Siri's search algorithm has already been mentioned: as is the case with all voice assistants, Siri typically offers users only one answer to their queries (Natale & Cooke, 2021, p. 1003).

For this reason, each search result of Siri can be considered technical biased. To identify, however, whether the search results are subject to pre-existing or emergent biases goes beyond the scope of this study. I only aim to identify whether a bias can be observed, but not if is a pre-existing or emergent one. For this reason, in this study, I do not differentiate between the two types of biases, but only refer to a bias.

2 METHODS

2.1 Participants

The participants included 170 US-based Siri users[3] who were between 18 and 64 years old and of different genders, locations, and political leanings. Ethical approval was obtained from Malmö University's Ethic Council before I began recruiting Siri users. This study was advertised exclusively on the crowdsourcing platform Amazon Mechanical Turk on April 21, 2021. The assignment was visible to all workers on the platform. However, to participate, users needed to be based in the US. Apart from this criterion, I did not use any filters such as age or location because I wanted to reach a heterogeneous group of Siri users. Participants were self-selected and compensated 2 USD for their time. Informed consent was obtained from all participants. Within roughly eight hours, 170 workers had completed the task.

However, for the final data analysis, I filtered out 36 users who occasionally or systematically did not report back Siri's replies as was intended – whether intentionally or not is unknown. Some participants answered the questions themselves, others summarized or interpreted the search results or filled the answer boxes with placeholder text. Even though this quality control increases this study's internal validity to some extent, it could not be done rigorously: I was able to identify whether participants made other queries than the ones they were supposed to and reported back answers to those searches. The quality control reduced the data set.

As table 1 shows, the final sample size consisted of 134 users. Each user was asked to report Siri's answers to five questions. However, not everyone reported back the answers to all five questions sufficiently. This explains why the total number of replies is 631, rather than 670 (this would have been the case, had all 134 users reported all five answers).

Despite the reduction of the size of the data set, the sample remained heterogeneous (see Table A1). However, the vast majority of participants belonged to the age of group 25–34. In addition, more males than females and more liberals than conservatives participated in the audit.

[3] This sample size was inspired by other studies which focus on content diversity of search results using a crowdsourced audit, (e. g. Nechusthai & Lewis, 2019; Cho et al., 2020).

Table 1. Data set

Sample size	134
Answers to the question "Should there be stricter gun laws?"	122
Answers to the question Should immigration be limited?	129
Answers to the question "Should the death penalty be abolished?"	130
Answers to the question "Should taxes be lowered? "	123
Answers to the question "Should abortion be legal?"	127
Total replies	631

2.2 Material and procedure

This study used a crowdsourced audit as proposed by Sandvig et al. (2014, p. 15). Every participant individually completed the assignment "Ask Siri Five Questions" on Amazon Mechanical Turk. After having read the assignment's description, participants signed a consent form containing the description of the assignment, which data would be collected and how it would be handled. Upon agreeing to the terms of this study, I collected basic demographic information (this included two multiple choice questions regarding the participants' gender and political leaning as well as two short answer questions regarding the participants' age and location). Next, participants were asked to use their smartphone to ask Siri and note down the first search results to each of the following politically controversial question (Chams, 2020; Najle & Jones, 2019):

- Should there be stricter gun laws?
- Should immigration be limited?
- Should the death penalty be abolished?
- Should taxes be lowered?
- Should abortion be legal?

Finally, the participants were thanked for the completing the assignment and paid through Amazon Payments.

Before the data could be analyzed, it had to be reduced. To do so, I replaced the actual answers that the participants had reported by numbers. For example, the first answer to the question whether guns laws should be stricter became "1", the second was turned into "2" and so on. Identical answers were labeled with the same number. This method reduced the amount of data significantly.

To measure the diversity of search results, I used univariate and bivariate descriptive statistics. The former focuses on a single variable and can be used to report a sample's distribution. To do so, summary measures such as frequency counts can be used. The latter can be used to analyze two variables (e. g. political

orientation and frequency of results) (Blaikie & Priest, 2019, p. 205). In the context of this study, using these two descriptive statistics was the logical choice, because I was only concerned with the frequencies of Siri's search results across a diverse set of users.

Inspired by Webster and Ksiazek (2012, p. 43) and Trielli and Diakopoulos (2019, p. 7), I use the Gini coefficient to calculate the concentration of Siri's replies per query. A Gini coefficient can range from 0 (most equal) to 1 (most unequal). In the field of media and communications studies, defining whether the level of concentration is low, medium, or high is challenging due to a lack of established normative standards. While scholars who study audience concentration use descriptions such as 'high degrees of fragmentation', they do not explain where they see the boundaries between low, medium, and high degrees of concentration (e. g. Webster & Ksiazek, 2012; Fletcher & Nielsen, 2017; Trielli & Diakopoulos, 2019). Because of this, I have chosen to borrow definitions from the field of economics. In the context of this study, a value of 0-0.3 indicates an equal frequency distribution of many search results and therefore low concentration. A Gini coefficient of 0.31-0.7 indicates moderate concentration and a value of 0.71-1 suggests a high degree of concentration, in which few search results dominate (Lambert, 2001, p. 31).

As numerous formulas can be used to calculate the Gini coefficient (Yitzhaki & Schechtman, 2013, p. 11), in the following I present the one that I used to increase this study's reliability (Lehn et al., 2000, p. 61):

$$G = \sum_{i=0}^{n} (H_i + H_{i-1}) \times q_i - 1$$

Once I had calculated the first version of the Gini coefficient for each query's result, I calculated a second and final version of the Gini coefficient that indicates the degree of concentration, independent of the sample size (Lehn et al., 2000, p. 61). To do so, I used the following formula:

$$G* = \left(\frac{n}{n-1}\right) \times G$$

I chose these formulas as they are widely used and less complex than other ones (e. g. Dorfman, 1979; Milanovic, 1997).

To calculate fragmentation, I drew inspiration from Bechmann and Nielbo (2018, p. 994). In this study, I define fragmentation as a lack of overlapping content. This was the case whenever a Siri's search algorithm exposed only one user to a specific search result.

As with concentration, when it comes to fragmentation, there is a lack of established normative standards. Based on my measurement of the degree of concentration among Siri's answers, I decided that a fragmentation of 30% or lower was low, a value of 31%-70% was moderate and one of 71% or higher was high.

For the second step of the analysis, I utilized bivariate descriptive statistics. Using this method of analysis, I analyzed the frequency of search results (dependent

variable) in relation to political orientation, location, gender, and age of users (independent variable).

3 RESULTS

The five queries led to a variety of search results. There are similarities across the searches, though (see Table A2). Firstly, Siri's search algorithm provided multiple answers to each question. The number of replies range from 23 to 46. Three questions led to more than 40 search results. The question whether taxes should be lowered received the lowest number of answers: 23.

Secondly, for each query there are one or two frequent replies. These answers reached 39 to 47 users (first question: 47; second question: 45; third question: 46; fourth question: 41, 45; fifth question: 39). The question whether taxes should be lowered is the only one that led to two frequent replies.

Thirdly, most of the answers to each question were personalized ones. The number of these unique replies range from 19 to 33. The question whether taxes should be lowered has the lowest value: 19.

Analyzing the data using univariate descriptive statistics reveals the sizes of the frequent and personalized answers in relation to all the answers (see Table 2). The few top results reached 35% to 70% of the participants, depending on the search. However, it needs to be kept in mind that questions about taxes, the answers of which were heard by 70% of the participants, had two common replies, while the other questions had one.

Fifteen to twenty-five percent of the participants were given unique replies. Four times these answers reached 20% or more of the users. The query about taxes stands out by having the lowest value: personalized replies were heard by 15% of the participants.

Table 2. Percentages of most frequent and personalized search results

Question	Should there be stricter gun laws?	Should immigration be limited?	Should the death penalty be abolished?	Should taxes be lowered?	Should abortion be illegal?	Total
Most frequent answers	39%	35%	35%	70%	42%	44%
Personalized answers	22%	25%	25%	15%	20%	22%

When the units of analysis are arranged horizontally according to the number of its total users, it becomes obvious that Siri's search algorithm has produced a long tail distribution of search results (see Figure 1). Siri's replies to each of the queries are unevenly distributed. A pattern is obvious across the five queries: The few top search results are at the head of the tail. Then, after a sharp decline, the replies that belong to the second category appear along the tail. The long part of the tail consists of the many personalized answers.

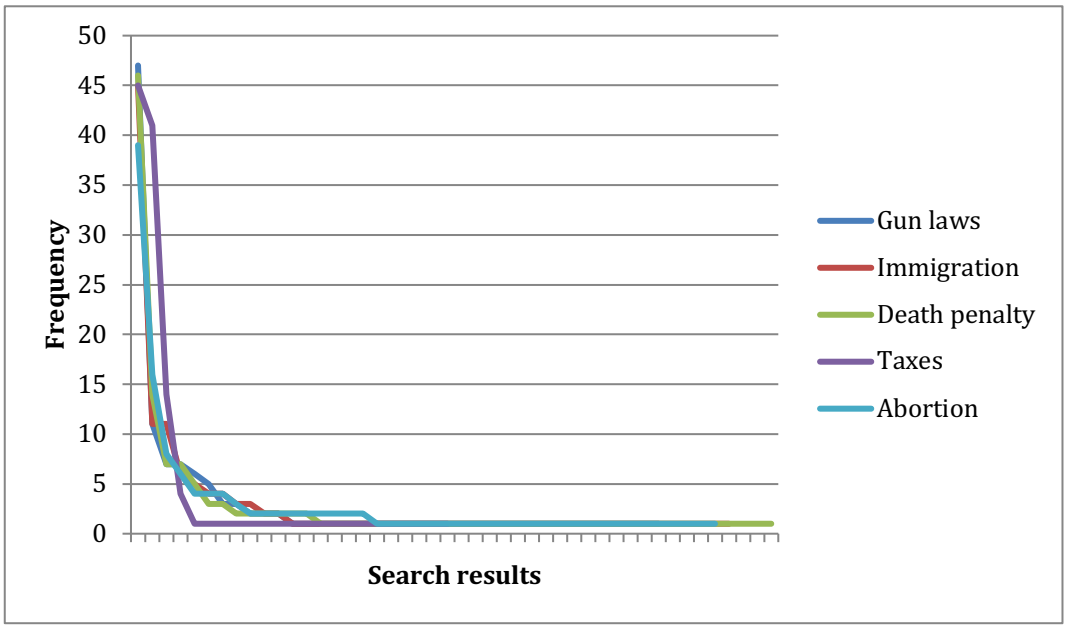


Figure 1. Long tail distribution of search results

To measure the concentration among the search results, I use the Gini index, which showed that the concentration among the search results is mostly moderate. As Table 3 shows, the Gini coefficients of the queries range from 0.588 to 0.758. Four values are below the threshold of 0.71, indicating moderate concentration. The Gini coefficient of the replies to the question about taxes, however, is above 0.71. This value suggests high concentration.

Table 3. Gini coefficients of queries

Query	Gini coefficient
Should there be stricter gun laws?	0.630
Should immigration be limited?	0.611
Should the death penalty be abolished?	0.599
Should taxes be lowered?	0.758
Should abortion be illegal?	0.588

To measure the fragmentation among the search results, I examined the percentage of non-overlapping content as previously explained. To do so, I turned to the unique answers. Table 2 already provided an overview of the percentages of these

replies for each query. They range from 15% to 25%. In total, 22% of the users received a personalized answer.

According to previously defined criteria, the fragmentation of search results is low. For every query, and therefore in total as well, the percentage of non-overlapping content is below 31%. Even though the data suggest that Siri's search algorithm causes only low fragmentation, it needs to be emphasized *that* is caused fragmentation throughout the searches at all.

So far the analysis has shown that Siri's search algorithm produced a long tail of replies and consequently caused moderate concentration and low fragmentation. This finding raises the question whether the identified concentration or fragmentation was driven by the political orientation, gender, age, or location of users.

It appears that Siri does not provide one group of politically like-minded users disproportionally frequently with top results or unique answers. The most frequent answers to each of the five queries were given to liberals, conservatives, people with other political views, and users who have no political leaning. Unique answers also reached a politically diverse audience.

The data suggest Siri's search algorithm is biased to some extent towards the gender of users. Siri's search algorithm provided female and male participants equally with unique search results. However, it caused concentration based on gender in relation to the most common answers: Siri provided men with the most frequent replies disproportionally often (see Table 4). Seventy-eight to eighty percent of the users who received five of the six replies were male, although they only made up 57% of the sample. The other reply represents an outlier: 76% of the participants who were provided with this answer were female, although women made up only 43% of the sample.

Table 4. Gender of users who received the most common search results

Question	Participants	Gender
Should there be stricter gun laws?	47	Female: 10 (21%) Male: 37 (79%)
Should immigration be limited?	45	Female: 9 (20%) Male 36 (80%)
Should the death penalty be abolished?	46	Female: 10 (22%) Male: 36 (78%)
Should taxes be lowered?	45	Female: 10 (22%) Male: 35 (78%)
	41	Female: 31 (76%) Male: 8 (19%) Other: 2 (5%)
Should abortion be illegal?	39	Female: 8 (21%) Male: 31 (79%)

It seems that Siri's search algorithm is biased towards location. The data indicate that Siri's search algorithm does not cause fragmentation, as unique answers were given to a diverse set of participants, but concentration based on the location of users. Siri appears to provide users from Pennsylvania more often with the top answers than others (see Table 5). The percentages of Pennsylvanians who were exposed to the most frequent answers (13% to 16%) is in five out of six cases roughly three times its percentages in the sample (5%) (see Table A1). However, due to the small number of users from different states who received the most frequent answers (see Table A3) (Tversky and Kahneman (1971) caution against relying on small numbers) I consider this finding to be highly tentative.

Table 5. Percentage of users from Pennsylvania who received the most common search results

Question	Should there be stricter gun laws?	Should immigration be limited?	Should the death penalty be abolished?	Should taxes be lowered?	Should abortion be illegal?
Percentage of users based in PA	15%	13%	15%	16%	13%

The data suggest that Siri's search algorithm is not biased towards age. It seems that Siri's top search results do not reach one age group disproportionally frequently. In addition, Siri provided users of different ages with personalized answers. The results indicate that Siri's search algorithm does neither cause concentration nor fragmentation based on the age of users.

4 DISCUSSION

This study sought to discover to what extent Siri's search algorithm provides 18 to 64-year-old US-based users with different answers to the same politically controversial questions. The main finding is that Siri's search algorithm seems to produce a long tail distribution of search results. Across the audit's five questions, it provided 42% of the users, disproportionally many men, with six answers. At the same time, Siri presented 22% of the participants with personalized search results. In the context of this study, these data indicate that Siri's search algorithm causes moderate concentration and low fragmentation. However, the thresholds for concentration and fragmentation were set rather arbitrarily due to a lack of established ones. If different thresholds had been chosen, different degrees of concentration and fragmentation would have been noticed.

4.1 Evaluation of findings

The finding that Siri's search algorithm causes both moderate concentration and low fragmentation is important in three ways. Firstly, the finding that the voice assistant has caused fragmentation, even though to a relatively low degree, is relevant in itself. Several scholars have claimed that democracies benefit from the presentation of a variety of voices and viewpoints of public life in diverse media content (e. g. McQuail, 1992; Benson, 2013; Anderson, 2016). Diversity is important when it comes to political information (Diakopoulos et al., 2018, p. 337–338). However, personalized search results can undermine the diversity of political dialogue (Gillespie, 2014, p. 188). If users received strongly biased answers to the audit's politically controversial questions which correspond to their own world view, this would undermine information diversity. However, due to this study's focus, the content of the search results was not analyzed. It is therefore unknown to what extent ideological diversity or topical diversity among the results exists.

Secondly, the concentration of search results, even though it was only mostly moderate, is an interesting finding. Scholars who are concerned with fragmentation claim that a unified body of relevant information should be accessible to the entire society (e. g. Gitlin, 1998; Sunstein, 2002). I mostly agree – especially when it comes to politically controversial questions. However, it is problematic that Siri's search algorithm decides how to answer these questions. Even if this study had analyzed the content of the six answers that 42% of the users heard, it would remain unclear why they were selected. The question arises whether the top search results managed to become "algorithmically recognizable" (Gillespie, 2017, p. 63); in other words, were these replies optimized to be found by Siri's search algorithm? Or were these results "privileged sources" (Diakopoulos et al., 2018, p. 331) selected by Siri? If so, their dominance might be based on some bias of the voice assistant (Trielli & Diakopoulos, 2019, p. 12). Due to the lack of transparency, these questions remain unanswered; in fact, they could not have been answered even if this study had had a different focus. Since Siri's search algorithm gave some answers significantly more often than others, it is clear, however, that Siri promotes some search results over others, just as other algorithms do (see Gillespie, 2015, p. 1). By doing so, it "shapes the things people encounter" (Beer, 2009, p. 1000). Because these *things* are answers to politically controversial questions, Siri's search algorithm actively influences the political information environment.

The finding that male participants more often received the most frequent results worrying, though not surprising. Gender biases in algorithms have long been known, and well documented, (e. g. see Bolukbasi et al., 2016; Fabris et al., 2020). This finding can be linked to the general issue of 'fairness' in searches (Ekstrand et al., 2019) as well as to 'group fairness' in particular (Fabris et al., 2020, p. 5). One might wonder why men received the frequent replies more often than women did. Although a more in-depth answer to this question would go beyond the scope of this study, Friedman and Nissenbaum's (1996) concept of algorithm bias suggests

two possible explanations. The identified gender bias could either be a pre-existing bias resulting from widespread social values and, hence, of the programmers themselves. Another explanation could be the or an emergent bias resulting from feedback loops caused by the users' behavior (Friedman & Nissenbaum, 1996, pp. 334-335). The former seems to be more likely, since biases in algorithms tend to reflect biases in society (Bolukbasi et al., 2016, p. 8).

4.2 Relation to previous research

The third way in which the finding that Siri's search causes both moderate concentration and low fragmentation is important becomes obvious when relating it to the outcome of other research.

At first sight, the findings of this study challenge some of the research that was discussed in the introduction. This study's data have shown that Siri's search algorithm provides roughly 22% of users with unique search results. This result suggests that Siri's search algorithm causes low fragmentation. In contrast, the empirical research of Haim et al. (2018), Curtois et al. (2018), Robertson et al. (2018), Bechmann and Nielbo (2018), Nechusthai and Lewis (2019) and Krafft et al. (2019) claims that there is little evidence of filter bubbles. Despite their different geographical foci and methodological approaches, these scholars found no evidence of fragmentation. This study's results seem to contradict them.

However, upon closer examination, the data of this study and the results of some of the previously mentioned academic works are similar. For example, Krafft et al. (2019) found that, on average, around two to four results out of ten were personalized, although this depended on the search term (p. 1). Similarly, Bechmann and Nielbo (2018) reported that roughly 10% to 28% (depending on the methodological approach) of users were presented with different content (p. 990).

While the data of this study as well as the studies reviewed here are comparable, the scholars have interpreted their results differently. One reason might be that they focused exclusively on fragmentation. Because of this, their perspective tends to be quite narrow: for them fragmentation, often in the form of filter bubbles, either exists or not. This study offers a different, more nuanced, view by focusing on *both* fragmentation and concentration. Its finding proves that fragmentation and concentration can coexist when it comes to search results given by Siri's search algorithm. This wider approach can provide a clearer understanding of content diversity and therefore provide new insights into the frequency distribution of search results.

The finding that female users received less often the most frequent results adds to the existing literature on gender bias. So far, studies have found out that search results reinforce gender stereotypes (e. g. Kay et al., 2015; Otterbacher et al., 2017; Bolukbasi et al., 2016; Fabirs, et al., 2020). Taking another perspective, this study has indicated that algorithms of search engines, in this case Siri, can also be biased towards gender by providing female and male users with different results.

4.3 Limitations

Some limitations exist. First and foremost, the findings of this study are not representative. I tested Siri's search algorithm using only 170 Amazon participants. None of these Siri users were found by using probability sampling. As a result, I cannot generalize from this sample to the larger population of US-based Siri users. However, it was never the aim of this research to make claims about a population. Instead, this study has sought to provide a mere indication to what extent Siri provides 18 to 64-year-old US-based users with different answers to the same questions.

Secondly, a high internal validity of the data could not be guaranteed. The study's participants made the queries themselves. As I was not present during the data collection, I cannot assure that the Siri users reported back the replies to the queries they were supposed to make.

Thirdly, this study's results are necessarily constrained by its context. The five queries that were used in the audit have resulted in concentration and fragmentation alike. It is unknown whether other, perhaps less controversial, searches (or, for that matter, different participants) would yield answers that have a different frequency distribution. Similarly, the findings here are US-based. Search algorithms might work differently in different countries; in fact, there is evidence that they produce localized outcomes (Kitchin, 2017, p. 25). It is possible that users who are based outside the USA would receive different answers if they asked Siri the same questions.

Fourthly, this study's focus is highly limited. I decided to research only the extent to which Siri's search results differ. The contents of the search results, including ideological or topical differences, have been disregarded. As a consequence, the information that Siri provided users with was not analyzed. In addition, due to the focus of this study and especially because of its theoretical framework I excluded the perspective of the users themselves. Even though scholars such as Gillespie (2014) and Bucher (2017) are increasingly focusing on users in order to understand algorithms more holistically, I decided not to do this. This limitation is not problematic, however. My research question only focuses on the outcomes of Siri's search results and the concept of algorithmic bias enables me to analyze them.

Fifthly, the results might not be replicable. As algorithms are constantly tested (e. g. through a/b testing), developed further and changed, their outcomes are likely to vary over time (Gillespie, 2014, p. 178; Diakopoulos et al., 2018, p. 322). For this reason, the results of the audit, which was conducted on April 21, 2021, are time-based. If the same users had performed the same searches two weeks later, they might have received different replies.

4.4 Future research needed

Due to this study's limitations, further research is needed. The small number of participants and the potentially low internal validity of the data were identified as the main limitations. Therefore, research with representative sample sizes and a more closely controlled environment (e. g. by being in the same place with participants or being virtually connected via a program such as Zoom) would have to be conducted to support or challenge the finding that Siri's search algorithm causes both concentration, especially among male users, as well as fragmentation. Such studies could use different search terms. However, they should not focus on a different location than the US because Siri's algorithm is likely to work differently in another country.

If future studies are able to disprove that Siri's search algorithm produces a long tail of search results, more diverse research would be needed. To avoid focusing only on the role of technology, the role of Siri's users would need to be considered. This approach would be in line with the growing number of scholars who see algorithms as the entanglement between the technical and social (e. g. Seaver, 2013; Bozdag, 2013; Just & Latzer, 2017). For example, inspired by Bucher (2017), one could ask how users feel about receiving unique search results or the same as other users. Or: how do users react to the replies they receive? In addition to a crowdsourced audit, interviews with users could be conducted. By doing so, not only the frequency distribution of results, but also their impact, could be studied.

FUNDING STATEMENT AND ACKNOWLEDGMENTS

The author would like to thank the two anonymous reviewers for their valuable comments and suggestions for improving the quality of this paper.

REFERENCES

Agichtein, E., Brill, E., Dumais, S., & Ragno, R. (2006). Learning user interaction models for predicting web search result preferences. *SIGIR '06*: *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 3–10. https://doi.org/10.1145/1148170.1148175

Anderson, C. (2016). Assembling publics, assembling routines, assembling values: Journalistic self-conception and the crisis in journalism. In J. Alexander, E. Breese, & M. Luengo (Eds.), *The Crisis of Journalism Reconsidered: Democratic Culture, Professional Codes, Digital Future* (pp. 153-169). Cambridge University Press. https://doi.org/10.1017/CBO9781316050774.010

Barker, R. (2018). Trapped in the Filter Bubble? Exploring the Influence of Google Search on the Creative Process. *Journal of Interactive Advertising*, *18*(2), 85–95. https://doi.org/10.1080/15252019.2018.1487810

Bechmann, A., & Nielbo, K. L. (2018). Are We Exposed to the Same "News" in the News Feed? An empirical analysis of filter bubbles as information similarity for Danish Facebook users. *Digital Journalism*, 6(8), 990–1002. https://doi.org/10.1080/21670811.2018.1510741

Beer, D. (2009). Power through the algorithm? Participatory web cultures and the technological unconscious. *New Media & Society*, 11(6), 985–1002. https://doi.org/10.1177/1461444809336551

Benson, R. (2013). *Shaping immigration news: A French-American comparison*. Cambridge University Press. https://doi.org/10.1017/CBO9781139034326

Blaikie, N., & Priest, J. (2019). *Designing Social Research*. Polity Press.

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? Debiasing word embeddings. *Advances in Neural Information Processing Systems 29*, 4349–4357.

Boothby, H. (2018, October 15). *Radio Production Autumn 2018* [Lecture recording]. Canvas@Malmö University. https://mau.se/canvas/

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, *15*(3), 209–227. https://doi.org/10.1007/s10676-013-9321-6

Bruns, A. (2019). Filter bubble. *Internet Policy Review*, *8*(4), 1–14. https://doi.org/10.14763/2019.4.1426

Bucher, T. (2017). The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. *Information, Communication & Society*, *20*(1), 30–44. https://doi.org/10.1080/1369118x.2016.1154086

Chams, Z. (2020, October 15). *Republicans vs Democrats: Where do the two main US political parties stand on key issues?*. ABC News. https://www.abc.net.au/news/2020-10-15/us-election-political-parties-explained-democrats-vs-republicans/12708296

Cho, J., Ahmed, S., Hilbert, M., Liu, B., & Luu, J. (2020). Do Search Algorithms Endanger Democracy? An Experimental Investigation of Algorithm Effects on Political Polarization. *Journal of Broadcasting & Electronic Media*, *64*(2), 150–172. https://doi.org/10.1080/08838151.2020.1757365

Curtois, C., Slechten, L., & Coenen, L. (2018). Challenging Google Search filter bubbles in social and political information: Disconforming evidence from a digital methods case study. *Telematics Informatics*, 35(7), 2006–2015. https://doi.org/10.1016/j.tele.2018.07.004

Dambanemuya, H. J., & Diakopoulos, N. (2020, March 20-21). *"Alexa, what is going on with the impeachment?" Evaluating smart speakers for news quality* [Paper presentation]. Computation and Journalism Symposium, Boston, MA, USA. https://cj2020.northeastern.edu/research-papers/

Dellinger, A. J. (2019, April 30). *Survey says that Siri and Google Assistant are the most used voice assistants*. Digitaltrends.

https://www.digitaltrends.com/home/siri-google-asistant-most-used-voice-assistants-alexa/

Diakopoulos, N. (2015). Algorithmic Accountability. *Digital Journalism*, 3(3), 398–415. https://doi.org/10.1080/21670811.2014.976411

Diakopoulos, N., Trielli, D., Stark, J., & Mussenden, S. (2018). I vote for – How search informs our choice of candidate. In M. Moore & T. Tambini (Eds.), *Digital Dominance: The Power of Google, Amazon, Facebook, and Apple* (pp. 320–341). Oxford University Press.

Dorfman, R. (1979). A Formula for the Gini Coefficient. *The Review of Economics and Statistics*, 61(1), 146–149. https://doi.org/10.2307/1924845

Ekstrand, M. D., Burke, R., & Diaz, F. (2019). Fairness and discrimination in retrieval and recommendation. *RecSys '19: Proceedings of the 13th ACM Conference on Recommender Systems*, 576-577. https://doi.org/10.1145/3298689.3346964

Fabris, A., Purpura, A., Silvello, G., & Susto, G. A. (2020). Gender stereotype reinforcement: Measuring the gender bias conveyed by ranking algorithms. *Information Processing & Management*, *57*(6), 102377. https://doi.org/10.1016/j.ipm.2020.102377

Feezell, J. T., Wagner, J. K., & Conroy, M. (2021). Exploring the effects of algorithm-driven news sources on political behavior and polarization. *Computers in Human Behavior*, 116, 1–11. https://doi.org/10.1016/j.chb.2020.106626

Fletcher, R., & Nielsen, R. K. (2017). Are News Audiences Increasingly Fragmented? A Cross-National Comparative Analysis of Cross-Platform News Audience Fragmentation and Duplication. *Journal of Communication*, 67(4), 476–498. https://doi.org/10.1111/jcom.12315

Fletcher, R., Cornia, A., & Nielsen, R. K. (2020). How Polarized Are Online and Offline News Audiences? A Comparative Analysis of Twelve Countries. *The International Journal of Press/Politics,* 25(2), 169–196. https://doi.org/10.1177/1940161219892768

Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems*, 14(3), 330–347. https://doi.org/10.1145/230538.230561

Geschke, D., Lorenz, J., & Holtz, P. (2019). The triple-filter bubble: Using agent- based modelling to test a meta-theoretical framework for the emergence of filter bubbles and echo chambers. *British Journal of Social Psychology*, 58(1), 129–149. https://doi.org/10.1111/bjso.12286

Gillespie, T. (2014). The relevance of algorithms. In: T. Gillespie. T., P. Boczkowski & K. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society* (pp. 167–194). MIT Press. https://doi.org/10.7551/mitpress/9780262525374.003.0009

Gillespie, T. (2015). Platforms Intervene. *Social Media & Society*, 1–2. https://doi.org/10.1177/2056305115580479

Gillespie, T. (2017). Algorithmically recognizable: Santorum's Google problem, and Google's Santorum problem. *Information, Communication & Society*, 20(1), 63–90. https://doi.org/10.1080/1369118X.2016.1199721

Gitlin, T. (1998). Public sphere or public sphericules? In J. Curran & T. Liebes, (Eds.). *Media, ritual and identity* (pp. 168–174). Routledge. https://doi.org/10.4324/9780203019122

Gutmann, A., & Thompson, D. F. (2004). *Why Deliberative Democracy?*. Princeton University Press.

Haim, M., Graefe, A., & Brosius, H. (2018). Burst of the Filter Bubble?. *Digital Journalism*, 6(3), 330–343. https://doi.org/10.1080/21670811.2017.1338145

Introna, L. D., & Nissenbaum, H. (2000). Shaping the web: Why the politics of search engines matters. *The Information Society*, 16(3), 169–185. https://doi.org/10.1080/01972240050133634

Jurkowitz, M., Mitchell, A., Shearer, E., & Walker, M. (2020). *U.S. Media Polarization and the 2020 Election: A Nation Divided: Deep partisan divisions exist in the news sources Americans trust, distrust and rely on*. Pew Research Center. https://www.journalism.org/2020/01/24/u-s- media-polarization-and-the-2020-election-a-nation-divided/

Just, N., & Latzer, M. (2017). Governance by algorithms: reality construction by algorithmic selection on the Internet. *Media, Culture & Society*, 39(2), 238–258. https://doi.org/10.1177/0163443716643157

Kay, M., Matuszek, C., & Munson, S. A. (2015). Unequal representation and gender stereotypes in image search results for occupations. *CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 3819–3828. https://doi.org/10.1145/2702123.2702520

Kinsella, B. (2020, November 5). *Voice Assistant Use on Smartphones Rise, Siri Maintains Top Spot for Total Users in the U.S.*. Voicebot. https://voicebot.ai/2020/11/05/voice-assistant-use-on-smartphones-rise-siri-maintains-top-spot-for-total-users-in-the-u-s/

Kitchin, R. (2017). Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 14–29. https://doi.org/10.1080/1369118X.2016.1154087

Krafft, T. D., Gamer, M., & Zweig, K. A. (2019). What did you see? Personalization, regionalization and the question of the filter bubble in Google's search engine. EPJ Data Science, 8(38), 1–23. https://doi.org/10.1140/epjds/s13688-019-0217-5

Lambert, P. J. (2001). *The Distribution and Redistribution of Income.* Manchester University Press.

Lehn J., Müller-Gronbach T., & Rettig S. (2000). Konzentrationsmaße. In J. Lehn, T. Müller-Gronbach & S. Rettig (Eds.), *Einführung in die Deskriptive Statistik. Teubner Studienbücher Mathematik* (pp. 50-64). Vieweg & Teubner Verlag. https://doi.org/10.1007/978-3-322-80099-2_3

McQuail, D. (1992). *Media performance: Mass communication and the public interest*. Sage.

Milanovic, B. (1997). A simple way to calculate the Gini coefficient, and some implications. *Economics Letters*, 56(1), 45-49. https://doi.org/10.1016/S0165-1765(97)00101-8

Najle, M., & Jones, R. P. (2019, February 19). *American Democracy in Crisis: The Fate of Pluralism in a Divided Nation*. PRRI. https://www.prri.org/research/american-democracy-in-crisis-the-fate-of-pluralism-in-a-divided-nation/

Natale, S., & Cooke, H. (2021). Browsing with Alexa: Interrogating the impact of voice assistants as web interfaces. *Media, Culture and Society*, 43(6), 1000–1016. https://doi.org/10.1177/0163443720983295

Nechusthai, E., & Lewis, S. C. (2019). What kind of news gatekeepers do we want machines to be? Filter bubbles, fragmentation, and the normative dimensions of algorithmic recommendations. *Computers in Human Behavior*, 90, 298–307. https://doi.org/10.1016/j.chb.2018.07.043

Olson, K. (2011). Deliberative democracy. In B. Fultner (Ed.), *Jürgen Habermas: Key Concepts* (pp. 140–155). Acumen Publishing. https://doi.org/10.1017/UPO9781844654741.008

Otterbacher, J., Bates, J., & Clough, P. D. (2017). Competent Men and Warm Women: Gender Stereotypes and Backlash in Image Search Results. *CHI '17: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 6620–6631. https://doi.org/10.1145/3025453.3025727

Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press.

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms that Control Money and Information*. Harvard University Press.

Robertson, R. E., Jiang, S., Joseph, K., Friedland, L., Lazer, D., & Wilson, C. (2018). Auditing Partisan Audience Bias within Google Search. *Proceedings of the ACM on Human-Computer Interaction*, 2, 1–22. https://doi.org/10.1145/3274417

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2014, May 22). *Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms*. [Paper presentation]. 64th Annual Meeting of the International Communication Association. Seattle, WA, USA.

Seaver, N. (2013). Knowing Algorithms. *Media in Transition*, 8, 412–422.

Sunstein, C. R. (2002). *Republic.com*. Princeton University Press.

Trielli, D., & Diakopoulos, N. (2019). Search as News Curator: The Role of Google in Shaping Attention to News Information. *CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3290605.3300683

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, 76(2), 105–110. https://doi.org/10.1037/h0031322

Valishery, L. S. (2021, January 22). *Share of voice assistant users in the U.S. 2020 by device*. Statista. https://www.statista.com/statistics/1171363/share-of-voice-assistant-users-in-the-us-by-device/

Vike-Freiberga, V., Däubler-Gmelin, H., Hammersley, B., & Pessoa Maduro, L. M. P. (2013). *A free and pluralistic media to sustain European democracy*. High Level Group on Media Freedom and Pluralism. http://ec.europa.eu/digital-agenda/sites/digital-agenda/files/HLG%20Final%20Report.pdf

Wagner, P. (2018, June 29). *Siri Remains The Most Used Mobile Voice Assistant*. Statista. https://www.statista.com/chart/14505/market-share-of-voice-assistants-in-the-us/

Webster, G. J., & Ksiazek, T. B. (2012). The Dynamics of Audience Fragmentation: Public Attention in an Age of Digital Media. *Journal of Communication*, 62(1), 39–56. https://doi.org/10.1111/j.1460-2466.2011.01616.x

Yitzhaki S., & Schechtman E. (2013). More Than a Dozen Alternative Ways of Spelling Gini. In S. Yitzhaki & E. Schechtman (Eds.), *The Gini Methodology:* A Primer on a Statistical Methodology (pp. 11–31). Springer. https://doi.org/10.1007/978-1-4614-4720-7_2

Zuiderveen Borgesius, F. J., Trilling, D., Möller, J., Bodó, B., de Vreese, C. H., & Helberger, N. (2016). Should we worry about filter bubbles? Internet Policy Review, 5(1), 1–16. https://doi.org/10.14763/2016.1.401

APPENDIX

Table A1. Characteristics of the sample (N = 134)

Age	18-24: 8 25-34: 87 35-44: 14 45-54: 16 55-64: 8
Gender	Female: 57 Male: 74 Other: 3
State	Alabama: 3 Arkansas: 1 Arizona: 2 California: 9 Connecticut: 8 Delaware: 1 Florida: 8 Georgia: 3 Idaho: 1 Illinois: 8 Indiana: 2 Kansas: 1 Kentucky: 2 Michigan: 5 Minnesota: 2 Mississippi: 3 Nevada: 3 New Jersey: 8 New York: 26 Pennsylvania: 7 Rhode Island: 2 South Carolina: 1 Tennessee: 2 Texas: 7 Utah: 2 Virginia: 4 Washington: 3 Wisconsin: 1 West Virginia: 1 Only indicating USA: 8
Political inclination	Liberal: 61 Conservative: 43 Other: 4 No political leaning: 26

Table A2. Number and frequencies of search results

	Questions				
Number of replies	Should there be stricter gun laws?	Should immigration be limited?	Should the death penalty be abolished?	Should taxes be lowered?	Should abortion be illegal?
1	11	1	46	4	1
2	1	1	14	1	1
3	1	1	1	1	1
4	1	1	1	1	1
5	5	1	1	1	1
6	1	1	1	1	6
7	1	1	1	1	2
8	1	2	1	1	16
9	7	4	1	1	1
10	2	1	3	1	2
11	1	1	1	1	2
12	1	1	1	45	1
13	1	1	2	1	1
14	1	1	1	1	1
15	1	1	2	41	2
16	1	1	2	1	1
17	2	4	1	1	2
18	1	1	1	1	1
19	1	1	1	1	1
20	1	1	1	1	1
21	1	1	1	1	4
22	1	6	1	14	1
23	1	1	1	1	1
24	7	1	1		1
25	2	1	1		1
26	47	11	1		2
27	1	1	5		4
28	3	1	1		39
29	1	1	1		1
30	1	2	2		1
31	1	1	1		8
32	1	3	1		1
33	1	1	7		4
34	1	45	1		2
35	3	5	2		1
36	6	11	1		1
37	1	1	1		1
38	1	1	1		3
39		1	1		2
40		3	1		1
41		1	1		1
42		1	1		2
43		1	3		
44			7		
45			1		
46			2		

Table A3. Location of users who received the most common search results

Question	Should there be stricter gun laws?	Should immigration be limited?	Should the death penalty be abolished?	Should taxes be lowered?		Should abortion be illegal?
Participants	47	45	46	45	41	39
Location	AR: 1	AR: 1	AR: 1	AZ: 1	AL: 1	AR: 1
	AZ: 1	AZ: 1	CT: 1	CT: 1	AR: 1	AZ: 1
	CT: 1	CT: 1	FL: 4	FL: 3	CA: 2	CT: 1
	FL: 5	FL: 3	GA: 3	GA: 3	CT: 7	FL: 3
	GA: 3	GA: 3	ID: 1	ID: 1	FL: 1	GA: 3
	ID: 1	ID: 1	IN: 1	IN: 1	IL: 7	ID: 1
	IN: 1	IN: 1	KS: 1	KS: 1	MI: 2	KS: 1
	KS: 1	KS: 1	KY: 1	KY: 1	MS: 1	KY: 1
	KY: 1	KY: 1	MI: 1	MI: 1	NV: 1	MI: 1
	MN: 2	MI: 1	MN: 1	MN: 2	NJ: 3	MN: 1
	MS: 1	MN: 2	MS: 2	MS: 2	NY: 6	MS: 1
	NV: 1	MS: 2	NJ: 3	NJ: 2	TX: 1	NJ: 1
	NJ: 2	NJ: 2	NY: 8	NY: 8	VA: 2	NY: 8
	NY: 7	NY: 8	PA: 7	PA: 7	WA: 2	PA: 5
	PA: 7	PA: 6	RI: 2	RI: 2	USA: 4	RI: 2
	RI: 2	RI: 2	SC: 1	SC: 1		SC: 1
	SC: 1	SC: 1	TN: 1	TN: 2		TN: 2
	TN: 2	TN: 2	TX: 1	TX: 2		TX: 2
	TX: 3	TX: 3	UT: 2	UT 2		UT: 1
	UT: 2	UT: 1	VA: 1	VA: 1		WA: 1
	VA: 2	VA: 1	WA: 1	WA: 1		
	WA: 1	WA: 1				