

**Keepin' it Real: Performing authenticity on Twitter
disinformation accounts**

Aurora Goodwin

Presented in partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Language and Communication

2022

School of English, Communication and Philosophy
Cardiff University

Abstract

Disinformation is information that is shared with the intention to mislead (Fallis 2015). In 2017, it was revealed that online personas who gained sizeable Twitter followings were part of a large-scale disinformation campaign that aimed to undermine public trust in American democratic systems, specifically the 2016 Presidential Election. According to the United States Department of Justice (2018:15), content was evaluated to ensure it ‘appeared authentic- as if operated by U.S. persons’. This thesis adopts a discourse analytic perspective and examines how disinformation accounts associated with the campaign constructed authentic personas on Twitter.

The data used in this thesis draws on a publicly accessible dataset of accounts and tweets that were part of the aforementioned campaign. Although some discussions of disinformation prioritise an approach to authenticity as an essential quality, this study approaches authenticity as a discursive performance. Performances of authenticity are analysed from linguistic, visual, and narrative analytic perspectives to address the following questions: How are handles, usernames, and profile pictures employed for authenticating identity performance on Twitter disinformation accounts? How do disinformation accounts manage (co-)tellership to (dis)identify with different social groups when storying a particular news event? This thesis reveals that across handles, usernames, and profile pictures, authenticity is achieved by appropriating established norms and practices primarily associated with two types of accounts: those belonging to individuals and those that share news. At the same time, the accounts’ narration of a high-profile news story reinforces the identity aspects presented in naming strategies and profile pictures, contributing to valid representations of particular personas. I argue that the disinformation accounts perform a particular type of authenticity, which can be referred to as: *authenticity at-a-glance*. I suggest that authenticity at-a-glance draws on the adoption of established practices of identification on Twitter that are readily accessible to viewers (via metadata units and tweet content) and orients to pre-existing social and political stereotypes.

Acknowledgements

Firstly, I am eternally grateful to have been blessed with Dr. Tereza Spilioti as my supervisor. Not just through my PhD, but during my whole time at Cardiff University, Tereza has been unwaveringly supportive, encouraging, and patient. Thank you for never failing to pick up on BS. Thank you for hearing me out and for inspiring me to have faith in my own ideas. And thank you for being constantly compassionate and kind. I could not have wished for anyone better to guide me through this journey.

I'd also like to say a special thank you to Dr. Lisa El Refaie for all of the thoughtful discussions during my reviews, to Dr. Chris Heffer for his kindness and support as my personal tutor during a fraught MA year, and to Professor Gerard O'Grady for being the first person to make me believe I might stand a chance at getting PhD funding... "Someone's gonna get it, why not let it be you?" are certainly words to live by. Thank you also to Rhian and the rest of the Cardiff University ENCAP team for the diligent responses to panicky emails and for always finding an answer, no matter the question.

To the Crime and Security Research Institute at Cardiff University, particularly Professor Martin Innes, a huge thank you for trusting me with fascinating data, for allowing me the freedom to explore it as I wished, and for including me in interesting disinformation discussions and events. I'd also like to say a big thank you to Andy Dawson for patiently responding to my technophobic questions about Twitter datasets.

Thank you to the Economic and Social Research Council for funding this work and for providing so many avenues for growth. I am forever grateful for being given the opportunity to give this a try.

To my PhD pals, both old and new, thank you for welcoming me into the fold and for reassuring me that I'm not alone in the PhD craziness. A special thank you to Alex, Axel, Ellie, and Emma for making the office more than just a workspace.

An enormous thank you to my friends outside of the PhD for continually asking "how's uni?" and taking an interest in what I'm doing, even though my answers are normally a garbled mess. Thank you for always being there for me even if we're physically far away and thank you for making me laugh every day. I am the luckiest person to have such wonderful mates.

Finally, to my family, thank you for loving me so loudly that sometimes I get embarrassed. To Nina, thank you for knowing me like nobody else and for always coming to my rescue when I

was a silly fresher. Cardiff was immediately home because I had my sister. Mum and Dad, where would I possibly begin? This is dedicated to you. Thank you will never be enough.

Table of Contents

Chapter 1 Introduction	11
1.1 Research questions	13
1.2 Thesis structure	14
Chapter 2 Research Background: Authenticity and Disinformation	16
2.1 Introduction	16
2.2 Authenticity as essential qualities	16
2.2.1 Authentic objects.....	16
2.2.2 Authentic people	17
2.3 Authenticity as discursive performance	19
2.3.1 Authenticity as performance on social media	22
2.3.2 Authentic performances of race and gender on social media.....	27
2.3.3 Authenticity as valid representation of realness.....	30
2.4 Disinformation on social media	31
2.4.1 Defining disinformation on social media	31
2.4.2 Disinformation research on social media	34
2.5 Social media discourse and platform design	40
2.6 Conclusion.....	43
Chapter 3 Researching Twitter Disinformation Accounts: Data and Methodology	45
3.1 Introduction	45
3.2 Digital discourse analysis.....	45
3.3 The data: Twitter disinformation accounts.....	47
3.3.1 Context	47
3.3.2 The datasets	48
3.3.3 Using pre-existing datasets: opportunities and challenges.....	52
3.3.4 The ethics of researching Twitter disinformation accounts	54
3.4 The methods: Investigating authenticity on Twitter disinformation accounts.....	57
3.4.1 Collecting and analysing handles and usernames	57

3.4.2 Collecting and analysing profile pictures.....	59
3.4.3 Collecting and analysing tweets from a narrative perspective	61
3.5 Conclusion.....	65
Chapter 4 Authenticating Identity Performance on Twitter Disinformation Accounts: Handles and Usernames.....	68
4.1 Introduction	68
4.2 Defining Twitter profile handles and usernames	68
4.3 Identification and authentication through Twitter handles	70
4.3.1 Eponyms and pseudonyms	70
4.3.2 First names indexing gender and race	71
4.3.3 Pseudonyms indexing various identity aspects	73
4.3.4 Identification through Twitter handles: A qualitative analysis of the five main accounts	74
4.3.5 Identification through handles: Using descriptive statistics to analyse the wider sample.....	78
4.4 Identification and authentication through Twitter usernames.....	81
4.4.1 Identification through usernames: A qualitative analysis of the five main accounts	81
4.4.2 Identification through usernames: Using descriptive statistics to analyse the wider sample.....	85
4.5 Discussion	89
4.5.1 News accounts.....	90
4.5.2 Personal accounts	92
4.6 Conclusion.....	95
Chapter 5 Authenticating Identity Performance on Twitter Disinformation Accounts: Profile Pictures	97
5.1 Introduction	97
5.2 Visual analysis and authenticity	98
5.3 Identification in profile pictures: visual authenticity on Twitter.....	99

5.3.1 Kress and van Leeuwen: visual modality.....	99
5.3.2 Beyond visual modality.....	101
5.3.3 The visual analysis of Twitter profile pictures.....	102
5.4 Identification through profile pictures: A qualitative analysis of the five main accounts	103
5.4.1 @TEN_GOP	103
5.4.2 @Pamela_Moore13.....	107
5.4.3 @wokeluisa.....	110
5.4.4 @SouthLoneStar	113
5.4.5 @USA_GunSlinger.....	115
5.5 Identification through profile pictures: Using descriptive statistics to analyse the wider sample.....	120
5.5.1 Profile pictures and identity in more and less popular accounts	122
5.5.2 Identity performance across metadata units in the more popular sample	125
5.5.3 Identity performance across metadata units in the less popular sample	128
5.6 Discussion	130
5.6.1 Personal accounts	130
5.6.2 News accounts.....	132
5.7 Conclusion.....	134
Chapter 6 Authenticating Identity Performance on Twitter Disinformation Accounts: Co-telling News Stories	136
6.1 Introduction	136
6.2 Storytelling and authenticity on social media	137
6.3 The story: The Unite the Right Rally, Charlottesville	138
6.3.1 Background and context.....	139
6.3.2 The aftermath	140
6.4 Telling stories on social media: Small stories heuristics.....	141
6.4.1 Tellers.....	142
6.4.2 Sites	144

6.4.3 Ways of telling	148
6.5 (Co-)tellership and dis/identification in the storying of Charlottesville	149
6.5.1 Tellership formats used in the storying of Charlottesville	149
6.5.2 Co-tellers involved in the storying of Charlottesville	151
6.6 Tweet formats and dis/identification in the storying of Charlottesville.....	152
6.7 Ways of telling and dis/identification in the storying of Charlottesville	156
6.7.1 The crash	157
6.7.2 Terrorist incident in Barcelona.....	159
6.7.3 Trump’s response to Charlottesville	160
6.7.4 Boston Free Speech Rally	166
6.8 Summary	167
6.9 Concluding discussion.....	168
Chapter 7 Concluding Discussion.....	172
7.1 Introduction	172
7.2 Summary of findings.....	172
7.2.1 Handles and usernames as means for authentication	172
7.2.2 Profile pictures as means for authentication	175
7.2.3 (Co-)tellership in storytelling as means for authentication	177
7.3 Methodological contribution	179
7.4 Theoretical contribution	184
7.4.1 Identifying disinformation.....	184
7.4.2 Keepin’ it real on social media.....	186
7.4.3 Authenticity at-a-glance	188
7.5 Critical reflection: limitations and avenues for future research	190
7.6 Conclusion.....	194
References	195
Appendix	222

List of tables

Table 1: Summary of tweet numbers for the five key accounts.....	50
Table 2: Summary of engagement rates for the five key accounts	50
Table 3: Distribution of disinformation accounts according to number of followers.....	58
Table 4: Summary of sample used in analysis of the storying of Charlottesville.....	62
Table 5: Identity aspects signalled in five account handles	77
Table 6: Naming strategies used in handles across both samples	79
Table 7: Identity aspects presented in handles for more popular accounts	79
Table 8: Identity aspects presented in handles for less popular accounts	80
Table 9: Identity aspects signalled in five account usernames.....	84
Table 10: Naming strategies used in more popular accounts.....	85
Table 11: Naming strategies used in less popular accounts	85
Table 12: Aspects of identity presented in more popular accounts.....	86
Table 13: Aspects of identity presented in less popular accounts.....	86
Table 14: Shared news naming strategies in the more popular sample	91
Table 15: Identity aspects signalled in five accounts profile pictures.....	120
Table 16: Identity aspects presented in profile pictures in more popular dataset	122
Table 17: Identity aspects presented in profile pictures in less popular dataset.....	124
Table 18: Broad categories of co-tellers	143
Table 19: A summary of tweet formats and their associated narrative potentials	146
Table 20: Single and multiple tellership tweets across the five disinformation accounts.....	149
Table 21: Types of co-tellers involved in storying Charlottesville across the five accounts.	151
Table 22: Distribution of tweet formats across the five disinformation accounts	152
Table 23: Examples of tweet formats used when engaging with right-wing co-tellers	153
Table 24: Examples of tweet formats used when engaging with left-wing co-tellers	154
Table 25: Coding categories for handles and usernames	224

Table of figures

Figure 1: Layout of a Twitter profile	68
Figure 2: A tweet as it appears in a home timeline	68
Figure 3: Continuum of naming strategies in relation to authenticity.....	95
Figure 4: A default Twitter profile picture	97
Figure 5: Cartoon as profile picture	97
Figure 6: Logo as profile picture.....	97
Figure 7: Photograph of person as profile picture	97
Figure 8: @TEN_GOP profile picture (August 2016- May 2017)	104
Figure 9: @TEN_GOP profile picture from June 2017 onwards	104
Figure 10: State seal of Tennessee	104
Figure 11: Confederate flag.....	104
Figure 12: American flag	104
Figure 13: @Pamela_Moore13 profile picture throughout activation	108
Figure 14: @wokeluisa Profile picture throughout activation	110
Figure 15: @SouthLoneStar Profile picture throughout activation	113
Figure 16: @USA_GunSlinger profile picture (July 2016-June 2017)	115
Figure 17: @USA_GunSlinger profile picture from June 2017 onwards.....	115
Figure 18: Word-based profile picture indexing location	121
Figure 19: Image-based profile picture indexing location	121
Figure 20: The 'today' group	123
Figure 21: The 'online' group	123
Figure 22: The 'daily1' group	123
Figure 23: The 'daily2' group	123
Figure 24: The 'top news' group.....	123
Figure 25: Profile picture that replicates username, St. Louis Online	128
Figure 26: Profile picture used by verified BBC News Twitter account	133
Figure 27: Profile picture used by @StLouisOnline from the 'more popular' data sample of disinformation accounts	133
Figure 28: Verified CBS News account, @CBSChicago	133
Figure 29: Verified CBS account, @CBSNewYork	133
Figure 30: @MinneapolisON, Minneapolis Online	134
Figure 31: @NewOrleansON, New Orleans Online	134
Figure 32: @PhiladelphiaON, Philadelphia Online	134

Figure 33: An example of an original post with single tellership	144
Figure 34: An example of a retweet, where @TEN_GOP retweets Donald Trump.....	145
Figure 35: An example of a quote tweet, where @USA_GunSlinger reports another user's tweet and adds their own commentary	145
Figure 36: An example of a reply, where @TEN_GOP replies to another user.....	145
Figure 37: Image accompanying Table 24.....	154
Figure 38: Chronological timeline of Charlottesville events	157
Figure 39: Image accompanying Example 10.....	163
Figure 40: First image accompanying Example 12.....	165
Figure 41: Image accompanying Example 12.....	165
Figure 42: Image accompanying Example 12.....	165

Chapter 1 Introduction

On social media, the idea of “being real” is ‘consistently embraced as a kind of meta-value’ (Whitehead 2015:126), where authenticity is understood as a ‘presumed correspondence of representation and reality, wherein the contents of something are what they claim to be’ (Whitehead 2015:125). This drive to “being real” draws on the presupposition that there are “non-real” or “fake” people and/or situations online, an assumption that is reinforced by earlier claims about internet anonymity whereby people can be “anyone” online. Sztompka (1999:82) for example, suggests that for online users to trust one another, anonymity must be reduced to ‘the levels acceptable for any “normal” transactions’. These early debates about anonymity, trust, and deception online continue to be couched in discourses of “being real”- throughout the duration of internet communication, there has been some fascination with the idea that people can be anyone online. We, as a society, are enthralled with the idea of outing and unmasking people, demonstrated through the popularity of television phenomena like *Catfish*, where the programme scrutinises online relationships to find out if both romantic partners are “real”, i.e., they are indeed who they claim to be.

While (in)authenticity online has primarily been explored in interest-based chatrooms, internet dating, and personal blogging, in more recent years, there has been a shift of attention to (in)authentic activity in political contexts. For example, bots, shills, and trolls astroturf, hijack and target topics and individuals in relation to political debates and other social issues. These terms are often discussed in relation to the concept of disinformation, i.e., the intentional spread of misleading information (Fallis 2015). It is important to acknowledge that disinformation is not solely defined by the factuality of content, but it also depends on the context within which information is generated and shared. For example, the manufactured, organised dissemination of information for specific purposes (e.g., to undermine public trust in political systems) can be classified as disinformation, regardless of whether the information being shared is factually true or false (Fallis 2015:418).

The concept of disinformation is, of course, not new. The practice of sharing information that smears opponents, skews public opinion, and could potentially harm citizens can be traced back to the Roman times (Posetti and Matthews 2018:1). However, following the exposure of Russian disinformation campaigns across social media platforms in relation to the 2016 US Presidential election (Donald Trump vs. Hillary Clinton), issues relating to organised efforts to disrupt democracy using social media have risen in public consciousness.

This thesis explores a set of disinformation accounts that have been exposed as being operated by the Internet Research Agency (henceforth IRA), a Russian disinformation organisation that were described by the United States Department of Justice (2018:2) as ‘engaged in operations to interfere with elections and political processes’. According to the indictment of individuals engaged in these activities, ‘defendants and their co-conspirators [...] regularly evaluated the content posted by specialists (sometimes referred to as “content analysis”) to ensure they appeared authentic—as if operated by U.S. persons’ (US Department of Justice 2018:15). The notion of ‘authenticity’, thus, as the result of consistent effort and management by specific individuals is explicitly associated with the specific types of accounts. In fact, as discussed in Chapter 3, mainstream media, such as the BBC and British tabloids, unknowingly reproduced content of these accounts as evidence of public opinion, indicating that the accounts were believed to be valid representations of who/what they claimed to be (cf. Whitehead 2015:125). Largely, these accounts (re)presented US citizens who displayed polarised, partisan, positions, supporting either left-wing or right-wing politics. Through the performances of these personas, the accounts ultimately aimed to ‘sow discord’ (Linville and Warren 2020:447) and exacerbate social tensions as a means of disrupting democratic processes in the US.

This study aims to understand how the accounts activated by the IRA constructed authentic personas for the purposes of disinformation. In this thesis, authenticity is understood and approached as a dynamic process that is performed and constructed through discourse, rather than an essentialised concept of unquestionable truth. In other words, my research aims to understand how the accounts activated by the IRA constructed authentic identity performances. Twitter allows and encourages users to present aspects of their identity in both textual and visual modes. While features facilitated by the platform of Twitter, like usernames and profile pictures, have been found to influence the way that identity performances are perceived (Andrews et al 2016; Morris et al 2012), relatively little research pays attention to how these features work together with tweet content and contribute to the identity performances within IRA accounts. As a result, a key contribution of this research is that identity performance is analysed by paying attention both to features facilitated by the platform (like handles, usernames and profile pictures) that are typically analysed from computer science perspectives, and to tweet content, that tends to take precedence in linguistic studies of Twitter. Overall, this research also aims to contribute to popular discussion about disinformation and demonstrate that disinformation accounts use a range of practices to perform personas that may not

immediately appear suspicious based on the criteria or features of disinformation that are often mentioned in mainstream media.

1.1 Research questions

As evident in the above discussion, my research aims to answer the following main question: how do disinformation accounts construct authenticity through identity performance on Twitter?

According to Jones (2009, 2013, 2020), the technological architecture of social media platforms influences the way users share information about themselves, particularly if we approach social media as sites of display, i.e., specific contexts in which different modes and media work together to enable particular social actions. Therefore, as well as exploring how users utilise language to perform identity online, researchers should consider how the platform itself influences the opportunities available to users when performing ‘the self’. This thesis follows this line of research and investigates how identity performances are situated within the sites on which they are displayed. More specifically, my study will investigate the following three areas which correspond to some of the key features of Twitter’s platform affordances:

1. How are handles and usernames employed for authenticating particular personas on Twitter disinformation accounts?
2. How are profile pictures used for authenticating particular personas on Twitter disinformation accounts?
3. How do disinformation accounts manage (co-)tellership to (dis)identify with different social groups when storying news events, such as the Charlottesville Unite the Right Rally?

Each of the above questions addresses an area of identity performance that is linked with the architecture of the Twitter social media platform. Firstly, all Twitter accounts must have a handle and a username, as this is part of the process of setting up a Twitter account and neither of the two features can be removed at any time. Secondly, all Twitter accounts must have a profile picture, as even users who decide not to upload their own personal image will end up with the default profile picture provided by Twitter. Finally, Twitter provides users with several options for interacting with other users and co-telling stories. Twitter’s features of retweeting, replying, and quoting other users are part of the platform’s design, and thus, enable particular social actions that can contribute to the performance of collective identities.

1.2 Thesis structure

Having briefly outlined the key concepts and aims of this study in this chapter, I will continue with defining authenticity and disinformation in chapter 2. More specifically, I will review previous research into authenticity performances both in the broader context of social media communication and in the more specific context of disinformation. I will critically evaluate previous approaches to authenticity and adopt an approach that views authenticity as a valid representation of “realness” that is both constructed and performed through discourse. The chapter ends with a brief discussion of social media discourse and platform design, with a specific focus on platform-facilitated options for identification and associated discussions of authenticity.

Chapter 3 presents the data that have been collected for the purposes of this thesis, as well as the specific samples and methods that will be used in each analysis chapter. I also critically reflect on digital discourse analysis as a research framework and explain how the platform architecture and its associated affordances can be investigated alongside language use by orienting to Jones’ (2009, 2020) view of digital discourses as facilitated by sites of display. I also discuss the sociocultural context within which this research takes place, as well as the ethical considerations associated with the design and execution of this project.

Chapters 4-6 represent the main body of the thesis with the three analytical chapters that address the three questions introduced in section 1.1. In chapter 4, I explore what identity aspects are presented in handles and usernames, with a specific focus on how authentic personas are performed through naming strategies (e.g., eponyms, pseudonyms, etc). I firstly examine the use of handles and usernames in the context of five popular IRA disinformation accounts and qualitatively analyse the identity aspects invoked through their naming strategies. I then use descriptive statistics to code the naming strategies of eighty-one (81) IRA accounts and explore any wider patterns in the types of identity aspects invoked in the data.

In chapter 5, I explore how profile pictures are also used to perform different aspects of identity. I analyse how different kinds of images (such as photos of people in comparison with cartoons or logos) connote different meanings and are associated with different social media norms. I also explore the extent to which identity aspects presented in profile pictures correspond with those shown in naming strategies and consider how the presentation of similar identity aspects across metadata units (i.e., handles, usernames, and profile pictures) may relate to practices of authenticity online. Like chapter 4, I initially carry out a qualitative analysis of the profile pictures used by the five main accounts, where I use existing visual analysis models

to investigate image composition and social meanings. Based on the labels generated in the qualitative analysis, I use descriptive statistics when analysing profile pictures used by the wider sample of 81 disinformation accounts.

Chapter 6 focusses on how Twitter's platform affordances, such as the ability to reply to, retweet, or quote other users, can be used by disinformation accounts to (dis)identify with other users and social groups when engaging with news stories. In this chapter, I explore how the five accounts focussed on throughout the thesis use storytelling to reinforce collective identity aspects (e.g., political affiliation) which are also indexed in each account's metadata. To do this, I use narrative analysis to analyse the telling of a contentious, high-profile breaking news story: the Charlottesville Unite the Right rally. With a focus on sites, tellers, and ways of telling, I investigate how the five disinformation accounts (dis)identify with other users in the narration of the event.

In chapter 7, I summarise the findings of each analysis chapter and I discuss this thesis's methodological and theoretical contribution to existing research on authenticity and disinformation. Furthermore, I discuss how my approach to authenticity 'at-a-glance' can be applied to the study of authenticity in other social media platforms. Finally, I acknowledge the limitations of this study and highlight avenues for future research.

Chapter 2 Research Background: Authenticity and Disinformation

2.1 Introduction

In this chapter, I discuss different approaches to authenticity and define the term as it will be used in this thesis. I then move to the concept of disinformation, a term that has also been debated in the literature, and review the ways in which disinformation on social media has been investigated from computer science and discourse analysis perspectives. Considering the focus of this thesis is on the discourse of online disinformation, I also discuss key aspects of social media discourse that have been foregrounded in the literature, including the concepts of metadata and platform affordances that are central to this thesis. Finally, I conclude this chapter by demonstrating how the research questions that this thesis aims to answer are motivated by the discussions within the chapter.

Authenticity is not a concept developed within- or for- linguistic research, such as notions like systemic functional grammar, speech acts or mood. Rather, similar to other concepts like politeness, authenticity can be understood in folk terms, especially when used in everyday contexts such as the branding of holiday experiences, restaurants, and art as “authentic”. As a result, attempts to define the term have been described as ‘confusing and contradictory’ (Taylor 1994:1). Nevertheless, there does appear to be some degree of consensus across folk and academic definitions of authenticity. For example, according to Beverland and Farrelly (2010:839) ‘what is consistent across the literature is that authenticity encapsulates what is genuine, real, and/or true’. A closer examination of different approaches to authenticity, however, reveals that not all studies define realness, genuineness, and truth in the same way. Instead, as Lindholm (2008:2) suggests, these terms are either approached from the perspective of origin (e.g., an item’s historical context) or content (e.g., in relation to identity or correspondence; whether something/someone is what/who they claim to be).

2.2 Authenticity as essential qualities

2.2.1 Authentic objects

According to Trilling (1972:93), “authenticity” has been initially and primarily used in the context of art and, more specifically, in relation to an artefact’s worth. This kind of approach to authenticity as a measure of the monetary and cultural value of objects is prevalent in everyday discourse. We commonly refer to authentic food, authentic paintings, authentic diamonds, or

authentic watches based on the provenance of the items and the historical or cultural context in which they originated (Lindholm 2008:1; Newman and Smith 2016:609).

Similarly, van Leeuwen's (2001:392-393) discussion of authenticity as genuineness and/or authorisation echoes this view of authenticity as a concept that is established from the (e)valuation of objects. 'Authenticity as genuineness' refers to the judgement of an item as authentic because it is genuine, i.e., because scientific procedures have been used to establish that an item is not an imitation or a copy, while 'authenticity as authorisation' refers to the quality of authenticity being established by an authority, i.e., an item receiving a stamp or seal of approval (van Leeuwen 2001:392-393). Establishing authenticity through evaluating objects in relation to a fixed origin (or ideal) reveals that the concept is viewed as a static quality that is inherent in the object. Based on the presence and/or absence of observable features, an item can be declared (in)authentic. For example, a painting is either an authentic Picasso, (because it has been assessed as such by an art curator) or it is not. While this view has been applied to the (e)valuation of objects, an approach to authenticity as an essential, binary, quality also permeates everyday experiences of people, particularly within the realm of identification.

2.2.2 Authentic people

Official identification documents like passports, driving licenses, and birth certificates can all be viewed as authorisations of a person's identity. Without these kinds of documents to authenticate our claims about who we are, we cannot legally travel overseas, receive paid employment, or enrol at University. In a similar vein to the authentication of objects, therefore, the authentication of a person's identity can also be understood in static terms when viewed through the lens of legal identification.

While such systems of identity authentication are widely used in our everyday 'offline' life, what happens on the internet where anonymity can be easily achieved has been the subject of popular debate. The potential for people to "be anyone" on the internet has become central to public discussions of authenticity and identification, encapsulated by popular phenomena like the television show, *Catfish*. Catfishing, as defined by the Cambridge Dictionary (2022), is 'the practice of pretending on social media to be someone different, in order to trick or attract another person'. In the television show, people in online relationships seek guidance in trying to find out if their partner is a "catfish", usually resulting in the unmasking of someone whose online identity is revealed as fake and dishonest because they do not orient to the morally correct practice of showing an 'almost seamless continuation of one's "real" offline self' (Lovelock 2016:206). According to Lovelock (2016:204), programmes like *Catfish* are structured to

convey the catfish as the ‘epitome of “bad” social media use’. In the context of mainstream media, thus, authenticity is also understood as an essential quality; users are evaluated within the binary categories of “real” or “fake” and “good” or “bad”, based on the degree of observable likeness between a person’s online identity and offline physical appearance.

Similarly, a “real” and “fake” distinction tends to be prioritised within discussions of nefarious online activity targeted towards the general public. In online guides designed to help users identify inauthentic activity, inauthenticity is defined in terms of whether the activity is automated or not. Users are encouraged to look out for features that indicate that the activity is automatically generated by computerised algorithms rather than being posted by human beings (Barojan 2018; Calabrese 2020; Dotto and Cubbon 2019). Authenticity is therefore again understood as an essential quality: a user is either a bot or a human, real or fake, authentic or inauthentic, and can be evaluated as such based on procedures that can detect and establish genuine authorship (cf. van Leeuwen 2001).

Furthermore, features of social media platforms also propagate an understanding of online authentic identity as an essential quality. On Twitter, for example, verified status is currently used to inform users that ‘an account of public interest is authentic’ (Twitter 2019). When an account is verified, a blue tick appears next to the user’s name, acting as a stamp of authority from Twitter that the account is officially affiliated with the person in question. To gain verified status, users must provide evidence that supports their identification, such as government-issued ID (McLean 2022)¹. In the context of the discourse of social media verification then, authenticity is approached as authorisation (cf. van Leeuwen 2001) and thus, an essential quality.

Beyond authorisation, authenticity has also been viewed in relation to the character of an individual, specifically, how true to themselves a person appears to be (Dyer 1991; van Leeuwen 2001). Across media contexts, this kind of authenticity has been a central focus of public discussion; in reality TV shows like *Big Brother*, ‘audiences look for the moment of authenticity when real people are “really” themselves in an unreal environment (Hill 2002:324), while chat shows that feature celebrities prioritise revealing something ‘real’ about the public figure (Langer 1981:361). Similarly, on social media, being ‘true to oneself’ is prioritised by

¹ It is important to acknowledge that at the time of writing, there are proposed changes to this system following a change in ownership. Plans include making it possible for any user to purchase verified status. It is currently unclear how identity will be verified in these cases, i.e., whether government-issued ID will still be required, and how this information will be stored.

users when assessing the authenticity of social media influencers and politicians (Lee and Eastin 2021; Luebke and Englemann 2022).

To make these assessments of truth to the self, there is an assumption that the authentic self is a ‘constant and unified’ entity (van Leeuwen 2001:393). For example, chat show hosts try to link their guests’ work with their lives, and generally attempt to get the guests to ‘speak as themselves’ with the ‘mask slipping’ (van Leeuwen 2001:393). Similarly, artists and singers must adopt one style and abandon all others if they want to be viewed as having ‘their own authentic style or voice’ (van Leeuwen 2001:393). The idea that the authentic self is constant is also highlighted in research on political authenticity, where citizens use existing knowledge about politicians to assess how ‘real’ they are and ultimately judge politicians who act similarly across different contexts as being true to themselves (Luebke 2021; Pillow et al 2018). Within these contexts, the “real” or “true” self is thus evaluated based on preconceived beliefs about what people are (or should be) like; we decide whether a politician, reality TV contestant, or social media influencer is being “real” or “fake” based on an external ideal or stereotypical persona that we have in mind about what the “true self” (or the person behind the mask) is like. An approach to authenticity that views identity as constant, unified, and thus, fixed, demonstrates that authenticity as truth to the self is largely understood as an essential quality even in everyday discourse.

2.3 Authenticity as discursive performance

In response to views of (authentic) identity as an essential quality, previous research has argued that identity can be approached as a fluid and dynamic performance. For example, rather than each person having one fixed identity, Bucholtz and Hall (2005:587) suggest that identity is a ‘discursive construct that emerges in interaction’, thus being situated within particular contexts. As a result, when authenticity is viewed through the lens of identity (Lindholm 2008:2), research has pointed to alternative views of authenticity as a dynamic performance constructed through discourse.

According to Coupland (2001), the concept of authenticity can be reconceptualised in terms of authenticities that are realised at different levels in language and interaction. Coupland (2001:416) uses the context of a TV news review to exemplify different levels of authenticity through language and suggests that authenticity can be constructed both in the discourse of institutions (e.g., the orientation to genre-specific norms such as dramatic theme tunes or contemporary studio sets of news programmes) and the discourse of news presenters (e.g., use of standardised language and formal dress). Of these different foci of authenticity, Coupland

(2001:415) highlights that personal authenticity can be understood ‘in accordance with the lay prescription that people should “be themselves”’ and suggests that styles of linguistic performance can index this kind of personal authenticity.

Research into celebrity identity constructions further exemplifies how authenticity can be understood as a discursive performance of what celebrities are ‘really like’ (Tolson 2006:150). According to Dyer (1991:228), the construction of the celebrity’s ‘real self’ is achieved through a rhetoric of authenticity, where lack of control, as well as lack of premeditation and privacy, are viewed as markers of authentic talk in media texts. Although van Leeuwen (2001:394) acknowledges that certain forms of talk (i.e., spontaneous) are often privileged as ‘authentic’, he points out that our judgements of authenticity are primarily governed by cultural norms and personal beliefs. Regardless of whether talk is produced spontaneously or not, judgements of authenticity are ultimately governed by what we accept as ‘valid’ in any given situation: ‘authentic talk is talk that can be accepted as a source of truth, beauty, and so on’ (van Leeuwen 2001:396). Based on this view then, authenticity is a concept that refers to what can be/is accepted as authentic, rather than whether something ‘is’ inherently authentic or not. Consequently, rather than asking whether something is authentic or not, van Leeuwen (2001:397) argues that analysts should instead ask ‘who takes this as authentic and who does not?’ and ‘on the basis of what visible or audible cues are those judgements made?’

This approach to authenticity can be seen in Tolson’s (2001) study of the identity performance of Geri Halliwell in a ‘behind-the-scenes’ documentary film. The study demonstrates how Halliwell constructed a valid and thus authentic construction of an ‘ordinary’ person through discourse. Tolson (2001:445) argues that ‘being yourself’ must be understood as a performance that is not meant to be perceived as acting but projects a general ‘aura of authenticity’. In Halliwell’s case, she projected an aura of authenticity as an ordinary person by, for example, comparing an upcoming job as a spokesperson for the United Nations Population Fund to learning how to use a photocopier in any, “ordinary”, new job. In other words, Halliwell validated her identity performance as an ordinary person by orienting to stereotypical associations with ordinary people and jobs.

The link between authenticity and ordinariness is also underlined in Thornborrow’s (2001:459) research into public participation media, a context that has given ‘new and enhanced status to the “authentic” voice of lay members of the public’. Compared to ‘experts’ (e.g., media professionals, politicians, scientists) who tend to be identified based on their credentials that legitimise their position and warrant their participation in discussions of a particular topic, Thornborrow (2001) argues that ordinary, or ‘lay’, people need to do extra discursive work to

validate their participation in public debates. In broadcast media, ordinary people tend to ‘discursively construct authentic positions for their own knowledgeable participation’ in discussions and debates by ‘doing’ authenticating talk (Thornborrow 2001:459). Thornborrow’s (2001) work exemplifies various self-categorising strategies that speakers use to authenticate their talk: in their initial turns, lay people mention previous actions relevant to the topic and/or invoke membership of social or professional categories, with expressions like “I’m a nurse”, “I’m the parent of two boys”, “I’m a diabetic”, etc. In other words, it appears that authenticity is ‘done’ differently depending on the communicative aim of the speaker; celebrities can construct authentic performances as ordinary people by ‘doing’ ordinariness (cf. Thornborrow 2001), while ordinary people can validate and authenticate particular positions in mediated debates by drawing on collective identity aspects.

The role that social categorisation can play in performing authenticity is also argued by Cutler (2003) who studied the construction of authenticity in identity performances of white middle-class hip-hop fans in New York. In her study of how white hip-hoppers attempted to construct belonging within the hip-hop community, Cutler (2003) highlights the wider cultural significance of the concept of authenticity for the specific community: to ‘do’ hip-hop is to ‘keep it real’, i.e., to be true to oneself, not put up a front, and not pretend to be something one is not. Cutler (2003:215) found that middle-class white hip-hoppers on the periphery of hip-hop culture (i.e., those that did not take part in activities like MCing and DJing) oriented to authenticity as being situated specifically in the strength of a person’s connection to the ‘street or urban ghetto’. To establish credibility and ‘keep it real’ within the hip-hop community, some white hip-hoppers thus obscured the racial and class boundaries that detached them from a community typically viewed as authentically hip-hop (i.e., the urban African American community) by positioning themselves ‘semiotically closer to the “urban ghetto”’ (Cutler 2003:215). One of the ways that white hip-hoppers achieved this was by adopting speech markers stereotypically associated with African American Vernacular English (AAVE), such as *ain’t*, *habitual be*, and *multiple negation* (Cutler 2003:216). In a similar vein, Coupland (2003) also argues that certain linguistic features can encode personal or cultural authenticities. For example, Coupland (2003:424) suggests that in Wales, speakers who use ‘monophthongal (ou) and (ei) as in /ho:m/ and /we:lz/, are regularly stereotyped as being more “truly Welsh” than others’. Similarly, he suggests that different regions of Wales can be considered ‘more Welsh’ than others based on factors like historical Anglicisation. Both Cutler’s (2003) and Coupland’s (2003) research thus highlight the role that normative linguistic and cultural stereotypes play in performances of authenticity, particularly in relation to collective identity.

Their studies highlight that certain language features are adopted in an attempt to construct a self-presentation that can be taken as a valid representation of realness by peers, thus signalling authentic community membership.

In all cases discussed above, authenticity is ultimately viewed as a performance that emerges in interaction and is constructed through discourse. Repeatedly, studies make reference to authenticity as a particular type of “realness” that is accepted as valid by particular groups. Across these contexts, speakers attempt to construct realness by adopting recurrent and shared practices that are established as valid or normative within particular communities of practice: talking about ‘ordinary’ jobs in Halliwell’s case (Tolson 2001), adopting AAVE features in hip-hop (Cutler 2003), or using linguistic features that sound ‘more Welsh’ in Wales (Coupland 2003). Due to judgements of authenticity being wrapped up in personal beliefs and cultural norms, authenticity cannot be analysed as an ‘all or none’ concept (van Leeuwen 2001:395); what may seem authentic to one person or community, may seem inauthentic to another. While an approach to authenticity as an essential quality may prioritise establishing whether something **is** authentic, an approach to authenticity as performance instead encourages us to focus on who may have interpreted something as authentic, and on the basis of what cues, as recommended by van Leeuwen (2001:397).

2.3.1 Authenticity as performance on social media

As acknowledged in section 2.2, authenticity on social media can be approached as an essential quality, particularly as evident in the discourse of social media platforms (cf. verified accounts). In academic research of social media discourse, however, authenticity has been mainly studied as a discursive performance, with a focus on how language is used to identify and authenticate users as belonging to particular communities, particularly in an environment where information relating to the corporeal self is relatively unknown or inaccessible.

Research on online blogging has criticised views of (in)authenticity that limit the concept in terms of sharing false or true information (such as in a “catfishing” context) and demonstrates that criteria for authenticity are related to the extent to which there is adherence to community norms. When approached as a kind of online personal diary, blogging audiences assume a ‘kind of faith or trust in the truthful representation of the blogger and her story’ (Whitehead 2015:122). In the fashion blogging context, specifically, authenticity is not just desirable, but ‘positioned as an invaluable yet ineffable quality’ which distinguishes “good” blogs from “bad” blogs (Marwick 2013:1). Although interviews with bloggers and their readers suggest there is no clear consensus about what authenticity is, recurrent themes emerge that associate the term

with honesty, transparency, and realness: as summarised by one interviewee, when looking at fashion blogs, they often think ‘is this for real?! Is this authentic?!’ (Marwick 2013:6). Research across various online communities shows that judgements of authenticity and realness emerge and are assessed from below: they are established and assessed by one’s peers using subject-specific knowledge (Varis and Wang 2015:220; Arnesson 2022; Hunter 2016; Marwick 2013; McRae 2017; Whitehead 2015; Williams 2006). In the Christian mommy blogging community, for example, knowledge of motherhood allowed users to easily spot inconsistencies within the pregnancy/birth story of a blogger who fabricated her pregnancy and was thus ‘outed’ from the group for being inauthentic (Whitehead 2015:133). Similarly, in the fashion blogging community, knowledge of trends, brands, and prices is used by readers to assess whether a blogger can feasibly afford to wear the clothes they are showing on their blog. In cases where this is deemed unrealistic, these bloggers are viewed as inauthentic (Marwick 2013:5).

While authenticity is achieved through effective performance of community-specific subject knowledge (such as knowledge of motherhood and pregnancy in mommy blogging, or trends, brands, and prices in fashion blogging) across contexts, personal disclosure is also understood as an important strategy for constructing authenticity. Similar to research on celebrities that discusses the unveiling of the ‘real’ self in a chat show context (Langer 1981; van Leeuwen 2001), readers of fashion blogs report that bloggers are deemed authentic when the readers feel as though they “really” know them. This feeling is based on how much they know about the bloggers’ “real” lives i.e., their lives outside of the immediate context of fashion blogging (Marwick 2013:4). Similarly, motherhood blogs promise ‘an unedited, entertaining picture of motherhood as it exists in real life, not on TV or in glossy magazines’ (Whitehead 2015:129).

The construction of “real” life appears consolidated by displays of imperfection. For example, when speaking about the authenticity of fashion bloggers, one reader said ‘I find it hard to read or even know people where everything seems perfect all the time. I’m kind of like, “What’s really going on that you have to put up this farce?” Real life isn’t perfect all the time’ (Marwick 2013:5). Goffman’s (1959) metaphors of ‘frontstage’ and ‘backstage’ can be used to differentiate between these kinds of perfect (and assumed performed) presentations and imperfect (and assumed “real”) presentations. Goffman (1959) uses the metaphor ‘frontstage’ to refer to behaviours that he describes as ‘performances’, i.e., those that are polished, rehearsed and aim to achieve specific goals. In contrast, he uses ‘backstage’ to describe when a person can ‘step out of character’ and behave in a way that is not constrained by particular institutions or goals (Goffman 1959:112). Research into community formation online suggests that an identity performance is assessed as authentic when a blogger constructs a (re)presentation of

backstage self that the readers take as valid. In other words, they show aspects of their lives that would not typically be revealed in ‘frontstage’ performances on glossy magazines, for example.

Ultimately, research into blogging shows that authenticity can be understood as a ‘feeling’ of authenticity, i.e., bloggers construct performances that seem ‘true to themselves’ (Marwick 2013:4). Considering the overall social media ecology that affords users’ presence in multiple and interconnected platforms, Marwick (2013) argues that the consistency of a blogger’s self-presentation across media plays a key role in constructing authenticity on social media. Differences between a blogger’s “blog voice” and the voice they use on other platforms can indicate inauthenticity. As summarised by one interviewee, ‘[I unfollowed a blog that has] nice outfits, but the blogger’s voice on the blog doesn’t match their twitter account AT ALL. They’re so nice on their blog, but rude and mean on their twitter. DISLIKE!’ (Marwick 2013:4). Such findings suggest that there is a presupposition that bloggers will exist in various spaces and that lay users understand “real” parts of the self as those that remain consistent across contexts.

Research also draws attention to the fact that the performance of authenticity in blogging is negotiated and constructed in interaction between bloggers and audiences. For example, trust can be questioned and lost if the shared practices and values that bind bloggers and readers together appear compromised (Arnesson 2022; Whitehead 2015). Similarly, Marwick (2013:5) suggests that ‘genuine, frequent interactions with readers establish authenticity better than anything else’, with bloggers who do not engage with their readers being considered less “real” than those who do. It therefore appears that authenticity is also intimately related to performances of closeness that are not only achieved via the sharing of personal information, but also by the way in which bloggers navigate the communicative context of participatory media and interact with their audiences.

Based on research on blogging, authenticity appears a central value to which the identity performances of successful bloggers orient. According to Marwick (2013:7), authenticity ‘requires consistent labour to achieve’. As mentioned above, this discursive labour can take the form of effective performances of community-specific subject knowledge, self-disclosure practices, consistent management of the voice across different platforms and, last but not least, engagement with the bloggers’ audience where trust and authenticity are negotiated and co-constructed in interaction.

Beyond blogging, authenticity as a discursive performance achieved through orientation to community-specific norms has also been examined in research on celebrity discourse and identity on social media. For example, Kytölä and Westinen (2015) studied how the Finnish

footballer, Mikael Forssell, discursively constructed his identity as a hip-hop fan on Twitter. Similar to the white middle-class youth Cutler (2003) examined in her research, Forssell was found to draw on linguistic features typically associated with AAVE and to make explicit references to hip-hop culture (i.e., through clothing, ‘attitude’, and reference to shared knowledge). While such strategies were deemed as contributing to ‘keepin’ it real’ for some white middle-class hip-hop fans in Cutler’s (2003) study, Kytölä and Westinen (2015) found that online discussion forums labelled Forssell’s online performance ‘inauthentic’ due to a deviation from normative expectations related to his wider celebrity persona. The purist normativities (according to which a white, middle-class footballer cannot be a “gangsta”) were at odds with -and consequently caused friction with- the identities that stylisations of ‘gangsta English’ evoke (i.e., Black youth) (Kytölä and Westinen 2015:17). As a result, Forssell’s performance as a hip-hopper was not taken as valid by other online users. Kytölä and Westinen’s (2015:7) work therefore further foregrounds a view of authenticity as ‘a discursive process “through which people can make claims about their own or others’ statuses as authentic or inauthentic members of social groups”’ (Coupland 2010:105). The increased participatory affordances of social media arguably result in increased opportunities for contesting, negotiating, and co-managing authenticity in these contexts, shedding light on how identity performances are ‘taken’ (cf. van Leeuwen 2001) by others. The findings also foreground the issue of consistency of authentic performances across platforms, as well as across online and offline contexts. Rather than authenticity being essentialised, inherent, and thus ‘unquestionable realness’ (Kytölä and Westinen 2015:7), users utilise pre-existing knowledge of Forssell’s identity in the physical world (i.e., a white middle-class footballer) to negotiate and contest the validity of his Twitter identity performance within football fan forums.

The approach to authenticity as discursively performed rather than as unquestionable realness is also illustrated in Enli’s (2017) investigation of Twitter communication by Donald Trump and Hillary Clinton. Her comparative study of tweets produced by Trump and Clinton during the 2016 US Presidential Election focusses on how their online performances adhered to normative expectations related to political discourse, on the one hand, and expectations related to their established personas, on the other. In Enli’s (2017) study, authenticity was examined in terms of ‘authenticity markers’, i.e., discourse features indicating that the tweets were written by the account holders themselves rather than members of their publicity teams. These features, prevalent within Trump’s tweets, included capitalisation, impoliteness, and rants. These features were deemed unmediated, deviated from expectations related to political discourse, corresponded with Trump’s established persona as a political outsider, and therefore

suggested that the tweets were authored by Trump himself. In contrast, Clinton's Twitter identity performance as 'tech-trendy' and 'digitally-savvy' (Enli 2017:57) (e.g., by inviting users to play online games and communicate via emojis) appeared distant from her general image as a self-confessed 'less-than-tech-savvy candidate' (Rodham Clinton 2017:70). This suggested that the tweets were probably not written by her, ultimately leading to her online performance seeming less authentic than Trump's. Enli's (2017) study thus demonstrates how authentic performances are validated in relation to both norms imposed from above (e.g., institutional discourse of political communication) and norms negotiated from below (e.g., by the speaker and one's peers). Rather than certain features being more normative and thus more authentic than others, Enli's (2017) research demonstrates how negotiations of normativity and authenticity are 'continually evolving, emergent, and influenced by norms of specific communities and cultures' (Leppänen and Piirainen-Marsh 2009:261).

The above discussion demonstrates how authenticity can be approached as a performance when examining interactions on social media. The literature shows that authentic performances are judged against norms of communities of practice and are challenged and negotiated in interaction with social media audiences. Ultimately, audiences do not appear to assess authenticity in terms of whether the performance on the account "is" the person being portrayed, but whether the performance seems like a valid representation of a persona (i.e., as a hip-hop fan, a mother etc.). The literature shows that language plays a role in establishing authenticity online; users assess and react to how people use language features that align or are (in)consistent with the persona being represented. Consistency is repeatedly drawn on as an indicator of (in)authenticity, not only in terms of whether a performance appears (in)consistent with a particular persona (i.e., the use of AAVE by a white middle-class footballer), but whether a performance appears (in)consistent across platforms and contexts (i.e., showing the same personality in different spaces). In each of the aforementioned cases, interactions are examined at a micro-social level—that is, the focus is on identities that are made relevant within a 'consolidated genre and community of practice' (Coupland 2007:113) such as the mommy blogging community or football fan forums. While such research provides an insight into how interpersonal authenticity can be approached as a performance online, it is also necessary to acknowledge how socio-cultural authentication can be understood when viewed through a macro-social frame, that is, in relation to broader pre-established social ecologies (Coupland 2007:113). The next section explores such an approach to authenticity in more detail.

2.3.2 Authentic performances of race and gender on social media

The research discussed in section 2.3.1 focusses on performances of authenticity in relation to specific, known individuals. The research highlights the role that collective identity aspects like race or class can play in establishing an authentic persona. These broad social categories also remain relevant when analysed in relation to authentic identity performances by relatively unknown users. In particular, this section will discuss how research into identity performance on social media has focussed on the authentication of the macro-social categories of race and gender.

Coupland (2007:113) suggests that social categories like social class, gender, sexuality, age, and ethnicity can be interpreted as ‘pre-understood’ and that people can position themselves in relation to these categories when performing acts of identity. The macro-social category of race has been widely researched in relation to Twitter and, specifically, in the context of the so-called “Black Twitter”. “Black Twitter” is a term that gained prominence when large networks of Black² Twitter users utilised the hashtag function creatively to dominate trending topics on the platform (Brock 2012:545). There are two ways in which Black cultural identities have been found to be indexed in the trending topics: (i) through the inclusion of slang terms associated with Black American youth, such as #howyouballing, or (ii) through the use of predominantly Black avatars within contributions to a trending topic (Florini 2014:226). Overall, “Black Twitter” does not seek to homogenise all Black users. Instead, it foregrounds the significance of performances of Blackness online and challenges assumptions about a ‘generalised’ white Twitter user (Florini 2014:225).

Similar to the studies discussed in section 2.3, Maragh’s (2018) study of racial authenticity amongst users of Black Twitter approaches racial authenticity as a stylistic performance of racial identity. Maragh (2018) draws on social identity theory (Tajfel 1982) according to which group membership can be used to navigate the performance of one’s own identity. Tajfel (1982) drew on the idea of ‘in groups’ and ‘out groups’ and argued that emphasis of difference can

² I use the capitalised ‘Black’ throughout this work when referring to people who may also define themselves as African American. I acknowledge that such labels are often bound up in personal preference but use this term in line with popular Black media outlets (Glover 2020) and recommendations from scholars. As stated by Watkins-Hayes (2020) in Adams (2020) ‘the move you see now towards black is really to recognise the global nature of blackness. So I think that is the more universal term’. The capitalisation of the word pays homage to a ‘history with a very particular kind of political engagement’ (Cooper 2020) ‘puts respect to [racial] identities’ (Pirtle 2020) and ‘honours Black experiences’ (Fleming 2020) (see Eligon 2020). In contrast, ‘white’ will not be capitalised, in line with decisions by scholars of African American studies, due to whiteness being a social construct that ‘does not describe a group with a sense of common experiences or kinship outside of colonization and terror’ (Dumas 2016:13).

create distance from an out-group, on the one hand, and secure or strengthen one's membership to an in-group, on the other. In Maragh's (2018) work, Black Twitter discourse is shaped by speech norms and practices that are associated with a particular social group. As a result, authenticity is constructed through the extent to which users' talk corresponds with expectations arising from established norms.

Based on interviews with users of Black Twitter, Maragh (2018) revealed that certain types of language form (e.g., typing things "too properly"), content (e.g., not discussing "Black events"), and stance (e.g., having views different to the "black majority") become associated with particular racial identity performances (e.g., 'acting white' and indexing a distance from belongingness to Black culture). Maragh's (2018) research therefore points to the fact that normative associations between discourse and social categories are not only invoked in authenticity claims about performances of public figures and established personas (Enli 2017; Marwick 2013; Kytölä and Westinen 2015; Whitehead 2015), but they are also mobilised in relation to identity performances of ordinary Twitter users.

Based on the above, racial authenticity on social media is achieved through adherence to the expectations and norms associated with the racial group in question. In addition to language form, content, and stance encoded within the tweets, hashtags referred to as 'Blacktags', also 'play a role in constituting the "Black Twitter" identities' and increase the chances of 'being raced' online (Sharma 2013:46). For example, hashtags such as #onlyintheghetto and #ifsantawasblack are used to share anti-racist humour and engage in social commentary. Sharma (2013:46) states that such collective behaviours 'are grasped as emergent aggregations, materialized through the contagious social relations produced by the networked propagation of "Blacktags"'. In this context, we witness how features borne from technological affordances of sites like Twitter (e.g., hashtags) work together with pre-existing language norms and practices and create new repertoires of norms that become associated with particular identities or index social group membership.

The process of authenticating identities/personas through the mobilisation of discursive features that are typically associated with a specific social category has also been discussed in the context of gender performances online. In the context of online gaming, where 'sex swapping' (i.e., the practice of performing a gender identity other than what the player may identify with offline) is possible, Jenson et al (2015:866) argue that gendered identities are enacted not only through the selection of a male or female avatar but also through the 'players' naming practices, interactions with other players, and ongoing customisation options and play styles'. In many cases, gamers use characters in ways that orient to stereotypical gendered

patterns: female characters are given healing roles, while male characters are used during combat (Jenson et al 2015:864; Yee et al 2011:775). It appears that stereotypes also play a role within players' motivations for engaging in sex swapping. For males, performing as female has been found to be used as a way of getting help from other players, particularly in an environment largely designed by and for males, while women have reported that they adopt male personas in order to be taken more seriously and to avoid harassment (Hussain and Griffiths 2008; Lehdonvirta et al 2012; Herring and Stoerger 2014:578).

An orientation to gendered stereotypes within identity performances in a gaming context is also demonstrated in Herring's (2004a) research into the Turing Game³. Herring (2004a) investigated performances of gender and found that the primary performance strategy was the production of stereotypically gendered content. For example, players discussed 'topics males or females are conventionally expected to talk about' and gave 'content-appropriate answers to questions asked by the judge or moderator' (Herring 2004a:441). Herring's (2004a) study revealed that assessments of authenticity primarily relied on this kind of stereotyped content; contestants who produced stereotypical male content were rated as more authentic than those who did not. It therefore appears that within the online gaming context, gendered identity performances can be performed and authenticated both through the technological discourses that afford certain types of identification (i.e., avatar gender selection) and through the discourses utilised by users during their game play (i.e., the player's own behaviours).

Beyond gaming, research on online communication suggests that judgements of gendered performances draw on the linguistic features used in writing. For example, Thomson and Murachver (2001) found that readers were able to accurately identify the gender of the author of anonymous written emails. Readers' judgements were based on gender-linked language differences, according to which males were viewed as more likely to convey opinions, make insults, and produce more content, while it was assumed that females were more likely to make reference to emotion, provide more personal information, and modify the certainty of their statements with modals. Although the aforementioned studies exemplify early work within CMC and presuppose potentially problematic and out-of-date approaches to gender (i.e., presupposing there are two genders that predetermine speech patterns, focussing on what the gender differences are rather than how gender may make a difference (McElhinny 2003)), the

³ The Turing Game is a text-based chat environment where users play "tell the truth" style games, i.e., players identify themselves via text while others ask questions to try and assess the authenticity of their performances. Within this context, gender games are particularly popular (Herring 2004).

idea that CMC users orient to normative linguistic expectations to perform and assess gendered performances remains relevant to the study of authenticity and identity performance online.

In the context of social media, research suggests that gender stereotypes can influence how men and women are perceived: males are more often trusted as credible online information sources (Armstrong and McAdams 2009), while females tend to have more social connections and have higher social capital online (Magdy et al 2017). More specifically, women gain increased levels of audience attention and engagement online when the visual content that they share orients to conventional standards of attractiveness, femininity, and sexualisation (Drenten et al 2019; Duffy 2017). In the context of popular female Instagram influencers, language used in captions can contribute to sexualised performances of femininity. For example, cliché phrases, sexualised innuendos, and performances of playfulness have been found to characterise the discourse of female online influencers who create a ‘porn chic’ aesthetic (Drenten et al 2019:58). Although cues like cliché phrases and sexualised innuendo may create valid representations of ‘porn chic’ women, such strategies would arguably not lead to valid representations of women in other contexts, like mommy blogging, for example.

It therefore appears that across various online contexts, gender stereotypes are oriented to, performed, and exploited for certain results. The associated benefits of these performances (e.g., increased attention by other users), as well as the strategies used for (re)presenting gendered performances that can be ‘taken’ as valid (cf. Van Leeuwen 2001), vary depending on the context, aim, and audience of the interaction. Overall, the studies discussed in this section suggest that racial and gendered authenticities are co-constructed, negotiated and legitimated by orienting to normative collective stereotypes and, more importantly, by gaining normative acceptance from below.

2.3.3 Authenticity as valid representation of realness

Having reviewed research on authenticity, it appears that the concept can be approached either as an essential quality (e.g., something/someone either “is” or “is not” authentic based on an objectively observable feature) or as a performance where authenticity appears as a representation of a range of characteristics that are dynamically constructed through discourse and validated by – as well as in interaction with - others. In this thesis, I will approach authenticity as a valid representation of realness that is constructed through discourse; realness can take different forms in its realisation within specific contexts and communities of practice e.g., “real Black person”, “real man”, “real Republican” etc. The literature reviewed throughout this chapter demonstrates that both on and offline, people take representations of realness as

valid or not based on various cues. These cues can signal both authenticity in relation to an established persona (e.g., someone with a well-known presence both on and/or offline such as celebrities or popular bloggers) or authenticity in relation to group categorisation and membership.

In the context of Twitter disinformation accounts which are examined in this thesis, an essentialist perspective would consider these accounts as inauthentic as they were not from the U.S. as they claimed to be. However, at the time of their activation, the accounts were taken as authentic, i.e., as valid representations of U.S. personas. The fact that these accounts were deemed as authentic and valid is evident in the repeated use of their content as representative of US political views and beliefs by mainstream sources such as the BBC, The New York Times, The Daily Mail, and The Sun (BBC News 2017a; Shane 2017; Robertson 2017; Harvey 2017) and the societal shock at the outing of the accounts. As a result, the study of authenticity in this thesis does not focus on ‘how authentic are these accounts’, but on ‘which visible and audible cues’ these accounts and their content may have been judged as authentic, resulting in their appearance in mainstream news and in being followed and retweeted by thousands of Twitter users (cf. van Leeuwen 2001:396).

2.4 Disinformation on social media

2.4.1 Defining disinformation on social media

The term ‘disinformation’ refers to the intentional sharing of false, inaccurate, or misleading information to manipulate public opinion and/or cause public harm (High Level Expert Group on Fake News and Disinformation 2018:3). While the concept of disinformation is not a new one, discussions surrounding Russian interference in the 2016 US election brought the topic into sharper focus within public consciousness and increased scholarly attention to disinformation on social media (Tandoc et al 2018). In this section, I will discuss how disinformation works on social media, with a focus on the social and technological affordances that facilitate disinformation online and particularly in the case of Twitter.

According to Freelon and Wells (2020:146), social media has played a special role in exacerbating political polarisation during a period when public trust in institutional news media has steadily declined since the 1970s. Social media, particularly Twitter, provides citizens with an alternative news source that some may view as being more trustworthy than the ‘one sided’ mainstream media (Andersen et al 2021; Golovchenko et al 2018:981). When using social media as a news source, research suggests that users prioritise the social identity of other users

when determining the credibility of the information they share (Freelon and Wells 2020:147). Given the emphasis that users appear to place on the social identities of those who share news online, digital disinformation messages tend to trade on appeals of collective identity and group belongingness rather than informed discourse, particularly in relation to specific social groups or subcultures (Freelon and Wells 2020:147). Consequently, a decline of public trust in the mainstream media, combined with the opportunity for users to seek out “news” from people they affiliate with, leads to an environment where what people accept as news is shifting ‘toward a belief and emotion-based market’, where facts and evidence are replaced by emotion and personal belief (Rochlin 2017:386).

The spread of disinformation as situated within a ‘crisis of public communication’ (Freelon and Wells 2020:146) is exacerbated by a range of platform affordances that facilitate quick and easy information dissemination. As discussed in section 2.2.2, there is a lack of identity verification on social media. The ability for users to remain anonymous while using social media reduces the kinds of consequences that users may face in their offline lives, resulting in decreased inhibitions and the potential for the distribution of inaccurate and potentially harmful information (Giachanou and Rosso 2020; Lapidot-Lefler and Barak 2012). The lack of information gatekeeping and fact checking also means that users are free to share whatever they like, with very little censorship. Consequently, users can share information that may be factually inaccurate and harmful, without facing the kind of accountability and scrutiny that may characterise other media contexts.

In addition to the users’ opportunity to create inaccurate information in a context of potential anonymity, social media as a space for sharing allow the rapid spreading of information across large networks of users before platforms have time to remove messages that could lead to harm. On Twitter, for example, users create their own networks by selectively following other users. Content posted by these users will then appear in their followers’ home timelines. This content does not only include messages authored by the individual followed, but also content authored by other users and reposted on the individual’s profile. As a result, content can be shared and viewed by audiences beyond the context of the original post and profile. In terms of the networks that users create for themselves on social media, research suggests that users tend to connect with people that they agree with, and unfollow, block, or unfriend those that offend them (Sasahara et al 2021; Tagg et al 2017). Tagg et al (2017:46) argue that these kinds of negotiations may limit ‘the sort of mutual engagement and understanding across cultural, political, and social boundaries which was imagined in the early days of the Internet’. The tendency to avoid disagreement and seek out views that match with one’s own result in ‘echo

chambers', where oppositional viewpoints are scarce and pre-existing views are continually reinforced (Sunstein 2001, 2009).

These kinds of echo chambers are not only created by users themselves but are facilitated by platform affordances of social media sites. For example, Twitter operates on an algorithm, where users' activities and interactions are recorded. Twitter then uses this data to personalise content for each user, filtering messages according to prior activity. Due to users' tendency to engage with content they agree with, dissenting opinions are likely to be filtered out, resulting in users being shown more of what they already engage with, and ultimately reinforcing their pre-existing views (Grossetti et al 2019). The term 'filter bubbles' has been coined to refer to this kind of algorithm-produced, personalised content (Pariser 2011). According to Recuero et al (2020:570), these practices can provide users with a less varied 'media diet', creating the perception of 'false consensus' and increasing the chances of users being exposed to- and impacted by- disinformation.

Finally, Twitter makes particular topics and discussions searchable via the hashtag function. Hashtags in tweets can serve various purposes: they can mark topics, offer evaluations, and organise text (Zappavigna 2015:274). When a hashtag appears in a tweet (e.g., My thesis is about #disinformation), it is automatically aggregated into a searchable feed of all other tweets containing the same hashtag. Such a practice thus enables the creation of threads of discussion around certain subjects. Although hashtags appear to primarily function to aggregate content, Zappavigna (2011) argues that there has been a shift on social media, from solely content-based searches to searching to see what other people are saying/thinking about a topic. In this sense, hashtags can thus contribute to 'forming communities of shared value' (Zappavigna 2011:804). In other words, platform specific affordances like hashtags enable users to not only congregate around particular topics but create ambient affiliative networks with other users (as demonstrated in Sharma's (2013) work on Blacktags) and expand audiences beyond a user's immediate followers to much larger groups of potentially like-minded people.

On social media, the interaction between user behaviours and platform technologies demonstrates how information can quickly circulate and diffuse across groups of users, particularly where content has been targeted towards certain communities that lack exposure to alternative perspectives.

2.4.2 Disinformation research on social media

Given the increased attention to social media disinformation, academic research has also turned its focus to disinformation aspects. This section will discuss how disinformation has been researched and understood from a range of perspectives. After briefly mentioning how disinformation has been approached in popular sources and mainstream media, I will review disinformation research as it has been developing in relation to the way that disinformation accounts navigate the platform of Twitter, on the one hand, and the discourses that they produce and engage with, on the other. Then, I will highlight how this thesis fills a gap in current research.

Within mainstream media, efforts have been made to encourage users to scrutinise online sources more closely in order to reduce the impact of disinformation (Calabrese 2020). In this context, inauthenticity and disinformation tend to be approached in relation to specific markers. It is suggested that handles that contain long series of numbers, or profile pictures that are ‘anonymous’ (e.g., default profile pictures or photos of celebrities and cartoons) can be viewed as indicators of automated, inauthentic, and nefarious activity (Barojan 2018; Calabrese 2020; Dotto and Cubbon 2019). Users are encouraged to avoid engaging with such accounts, or to scrutinise accounts further before engaging by conducting a reverse image search or searching for more information about the information source. In other words, popular discussions tend to approach disinformation through a lens of automation, that is, disinformation as produced by automated bots that can be detected using particular indicators.

This popular framing of disinformation that juxtaposes disinformation accounts and associated practices with “ordinary” accounts is likely to have been informed by research from a computer science perspective. For example, Ferrara (2017) proposes that the most important indicators of social media bots relate to metadata information and usage statistics. Such indicators include the following: a) bots are more likely to use default settings and lack personalisation, b) bots tend to have an absence of geographical metadata, c) bots tend to show incessant activity and excessive numbers of tweets, d) bots tend to retweet content much more than producing their own material, and e) bots tend to have fewer followers in comparison to the number of users that they follow. Although such research is useful in providing an insight into how users may avoid engaging with political bots (as seen in relevant guidance targeted at the general public), Zannettou et al (2019:218) suggest that ‘automated content diffusion is only a part of the issue’ and highlight that humans play a key role in the dissemination of disinformation on Twitter (cf. Seargeant and Tagg 2016).

In their research, Zannettou et al (2019) investigate the behaviour of 1,000 Russian state-sponsored ‘trolls’ (i.e., sharers of disinformation) by comparing their activity to that of ordinary Twitter users. Zannettou et al (2019:219) investigate quantitatively a wide range of behaviours and find that disinformation accounts adopt ‘significantly different behaviours’ compared to ordinary Twitter users. These behaviours include among others: a) using Twitter Web Client rather than Twitter mobile, b) using hashtags more often than ordinary users, and c) ‘resetting’ their profiles by deleting their existing content and changing their details (e.g., screen name). Zannettou et al’s (2019) study thus provides an insight into how disinformation accounts that are not necessarily automated may navigate the platform of Twitter using behaviours that can be identified using technological tools (e.g., detecting the device used for accessing the platform, quantifying the use of hashtags, etc.).

Within research into the diffusion of disinformation across networks, there is also a focus on how citizens interact with disinformation tweets. Golovchenko et al (2018:979) highlight that research into the online spread of pro-Kremlin information predominantly focusses on state agents or state-controlled agents and overlooks the involvement that citizens have in the dissemination process. To investigate how citizens are mobilised by the kinds of disinformation shared by state-sponsored agents, Golovchenko et al (2018) use social network analysis to investigate how a particular disinformation narrative (Ukrainian responsibility for the MH17 crash) spread across Twitter. They analyse a network of retweets to establish which kinds of users had the greatest impact in disseminating disinformation and find that it is citizens, rather than state-controlled agents, who have the highest impact in generating popular disinformation content.

The use of network-mapping to study how narratives spread online is also prevalent in relevant research on health-related topics like COVID-19. For example, Jemielniak and Kremvovysh (2021) combine tweet coordination analysis with bot identification to analyse how fear-mongering vaccine tweets are spread and by whom. Similarly, Nogara et al (2022) use an algorithm to detect interaction networks and map how a small number of anti-vaxx activists used Twitter to disseminate 65% of COVID-19 related misinformation⁴ on social media. These studies provide an insight into how disinformation narratives are disseminated across networks and demonstrate how organised campaigns may reach, influence, and exploit ordinary citizens who end up propagating further such content.

⁴ I use misinformation here in line with Nogara et al’s (2022) adoption of the term.

This line of research appears to be motivated by a focus on automatic disinformation detection. Indeed, research that seeks to contribute to combating disinformation places emphasis on the development of technological tools that automatically detect disinformation. Some of these models use the findings mentioned above in order to automatically detect automated activity or ‘bots’ (Cai et al 2017; Lee et al 2011; Morstatter et al 2016), while others focus more closely on the content of messages. Within research that focuses more on the content of the messages, machine learning models have been adopted as means of automatically detecting inaccurate news stories, drawing on computational models to conduct large-scale studies of text sentiment, topic, and syntactic structures (Dubey et al 2022; Mohamed et al 2022; Patel et al 2022; Yafooz et al 2022). Such methods prove useful in exploring behaviours and patterns of disinformation sharing that are linked to particular technological aspects (e.g., networks, device) and they have influenced developments in artificial intelligence. However, they are limited in revealing more qualitative aspects of disinformation such as social and interpersonal processes that may be at play when users encounter and engage with disinformation accounts and social media content.

With respect to identity performance and the type of personas that disinformation agents may assume, Linvill and Warren’s (2020) research provides some insight into the kinds of identities adopted by IRA-controlled disinformation accounts. Using both quantitative and qualitative methods to understand the form and function of Twitter disinformation accounts, Linvill and Warren (2020) study both some of the behaviours mentioned above and the content that accounts used to populate their pages. For example, they identify dominant hashtags and themes in different kinds of accounts and suggest that IRA-controlled disinformation profiles can be split into 5 main categories: a) right troll, b) left troll, c) news feed, d) hashtag gamer, and e) fearmonger. Right trolls were characterised by spreading right-wing, populist messages, and commonly using hashtags such as #MAGA, #AmericaFirst, and #IslamKills, while left trolls posted socially liberal messages and regularly used hashtags like #BlackLivesMatter, #PoliceBrutality, and #BlackSkinIsNotACrime. Moreover, news feed accounts were those that presented themselves as news aggregators, used hashtags like #news, and created links with legitimate news sources. Hashtag gamers participated in online trends like #ThingsILearnedFromCartoons, and, finally, fearmongers engaged in disinformation narratives relating to fabricated crises like non-existent illness outbreaks, explosions, and war crimes. Overall, Linvill and Warren’s (2020) study provides a general overview of the types of accounts adopted by the IRA when spreading disinformation.

Studies of disinformation from discourse analysis perspectives can also contribute to a more nuanced understanding of social and interpersonal aspects of disinformation such as the performances as particular identities on the accounts. Quantitative discourse analytic studies, for example, draw on large datasets to establish common themes, key topics, and general tweeting practices used by different types of accounts. In the case of Xia et al's (2019) study, the focus is on a specific IRA account (known by the username Jenna Abrams). Drawing on mixed methods, their research examines how the IRA presented the persona of 'Jenna' across various platforms over an extended period of time. After collecting a random sample of tweets using Twitter's API, they used thematic analysis and quantitative methods to analyse dominant strategies in the construction of a particular persona. According to their findings, Jenna Abrams constructed an authentic identity via two main strategies: performing cultural competence and performing a likeable persona. Xia et al (2019) suggest that cultural competence was performed by tapping into 'talking points' where there was evidence of shallow knowledge of topics like Black Lives Matter, feminism, and immigration, and by using culturally relevant hashtags to spread messages to a wider community of American conservatives. Alongside this culturally charged performance, Jenna constructed a likeable persona by sharing content not related to politics and using affective words and punctuation. Finally, the study analyses temporal metadata patterns to gain an insight into how Jenna Abrams grew her following over time: for example, Jenna used interactive tweet formats like replies more frequently when growing her following, and less frequently once she had gained more followers. Through discourse analytic studies like the one conducted by Xia et al (2019), one can gain a more detailed insight into how 'right troll' accounts (i.e., one of the account types identified by Linvill and Warren 2020) design and construct performances of particular personas.

Beyond the focus on identity performance in disinformation accounts, discourse analytic studies of tweet content also aim to investigate disinformation sharing strategies. For example, in their study of disinformation during the 2018 Brazilian Presidential Election, Soares and Recuero (2021) use critical discourse analysis to establish strategies of manipulation used in political conversations on Twitter and find that biased framing and polarised ideological discourse are the most commonly employed strategies for manipulation. Similarly, Igwebuike and Chimuanya (2020) use a discourse legitimation approach and critical discourse analysis to analyse the strategies used to make disinformation items appear valid in the context of the 2019 Nigerian election. They find that the strategies of authorisation, moralisation, and rationalisation are employed to legitimate news items, together with linguistic features such as appeals to emotions, hateful comments, and coercive verbs.

Corpus methods have also proven useful in revealing topical trends and disinformation sharing strategies. For example, Monogarova et al (2021) use keyword frequency to track the topical changes of viral COVID-19 disinformation stories in Russia. They find that while neutral terms initially dominated disinformation items, there was a rise in specialised terms from relevant fields of biology, medicine, and physics in the second half of 2020, suggesting that disinformation items adopted different terms in line with growing public interest and knowledge of COVID-19 and associated scientific fields. Overall, this line of discourse analytic research on disinformation tends to focus on the textual content of disinformation posts in order to establish the broad themes and strategies used to spread disinformation narratives across networks.

Across both fields of computer science and discourse analysis, there is an overwhelming interest in how messages are amplified and spread across social media. As evident in the above discussion, attention has been primarily focused on either technologically determined behaviours (e.g., networks, algorithms) or discursive practices manifest in textual content of the shared messages. In fact, research on Twitter seldom examines how features of the platform architecture (i.e., usernames and profile pictures) interplay with user-generated tweet content (i.e., tweets, retweets) in the dissemination of disinformation. Although such research appears scarce, studies of hate speech and disinformation on ‘cloaked’ Facebook profiles, and of IRA framing of Black Lives Matter discourses, demonstrates how this dual perspective can provide meaningful insights into the sharing of disinformation.

Farkas et al (2017) investigate how several ‘cloaked’ (i.e., disguised) radical Islamist pages in Denmark created misleading content that exacerbated assumed differences between Muslims and ethnic Danes and led to hate speech towards Muslims. The accounts were active between March and September 2015 and shared hateful posts towards the Danish people and state. The posts were picked up by the mainstream media and resulted in thousands of comments containing Islamophobic and anti-immigration sentiment. According to Farkas et al (2017), the accounts successfully deceived users during their operation by tactically appropriating the platform affordances of Facebook. The fact that Facebook page administrators can stay anonymous and can remove other users’ comments enabled the cloaked Facebook pages to remove comments from anyone questioning the authorship of the pages. In terms of identity performance, Farkas et al (2017) highlight several recurrent features that were used to construct identities as radical Islamists. These features include: a) the use of ‘El’ or ‘Al’ in the middle of usernames of individuals (e.g., Fatima El-Mohammed); b) the use of ‘stolen pictures of real individuals’ in profile pictures, as well as the sharing of imagery that showed aggression by

Muslims towards Danish people and Denmark (e.g., images of the Danish flag being burnt); c) the inclusion of hyperlinks to existing Islamic organisations in the page's 'about me' sections; and d) the adoption of pre-existing, dominant discourses relating to the 'fear of the Islamisation of Europe' such as immigrant rape discourse and the victimisation of white European women.

Platform afforded identification options are also acknowledged in Arif et al's (2018) research into online Black Lives Matter discourses. From a computer-supported cooperative work perspective, Arif et al (2018) use interpretative network graphs and qualitative content analysis to investigate how right-leaning and left-leaning clusters of IRA accounts participated in online discourse about the Black Lives Matter movement and police-related shootings. They find that within both clusters, there are systematic patterns of profile presentation. In the left cluster, accounts largely take the form of a) personal profiles of 'proud African Americans', signalled by profile pictures of Black people and bios that mention the Black community, or b) grassroots political/media accounts committed to racial justice, signalled by the inclusion of 'news' in their bio and URLs that connect to external webpages. Similarly, in the right cluster, accounts took the form of c) personal profiles of 'proud white conservatives' signalled by profile pictures of white people and bios that indicated Southern location and/or support for the right to bear arms, or d) grassroots conservative news accounts, signalled by profile pictures showing the American flag and, again, inclusion of 'news' in their bios. These types of accounts demonstrate distinct ways of tweeting about Black Lives Matter and police violence: accounts from the left-leaning cluster share anti-police, pro-BLM messages while accounts in the right-leaning cluster share anti-BLM, pro-police messages. Arif et al's (2018) investigation of tweets also highlights how accounts in political clusters sometimes interacted with one another. Accounts from the same clusters retweeted and replied to one another as a means of signalling membership to a social clique. Arif et al (2018:15) suggest that by exploiting stereotypes and creating consistent 'brands' across accounts, the IRA accounts were able to micro-target their audiences and blend into the communities they aimed their messages at. By acknowledging identification options provided by the platform alongside tweet content, Arif et al's (2018:13) research highlights how IRA accounts presented Black and white Americans as 'binary groups that are internally homogenous with respect to politics'. Without contextual information provided by profile information (such as profile pictures showing people/bios that refer to race and politics), such insights may not have been possible.

Farkas et al (2017) and Arif et al (2018) draw on both platform-determined features like profile pictures and usernames, which have primarily attracted the interest of researchers concerned with automated detection of disinformation, and user-generated discourse evident in

post content, which has largely been examined from a discourse analysis perspective in disinformation research.

In this section, I have discussed how disinformation on social media has previously been researched. Despite the fact that both platform-specific features (such as metadata information, networks, algorithms) and textual content of social media posts have been the focus of existing research, these areas have been mainly approached from distinct perspectives. To further understand how and why the two areas would be fruitful to be combined, especially in examining the processes through which disinformation accounts construct authentic personas, In the following section, I will further discuss how these concepts are reflected in research that has been developed for approaching and studying social media discourse.

2.5 Social media discourse and platform design

Social media is an umbrella term that refers to ‘internet-based sites and services that promote social interaction between participants’ (Page et al 2014:5). In contrast to mass media, where communication tends to be one-to-many, on social media, content can be published by anyone and still reach potentially large audiences (Page et al 2014:5). Social media discourse thus refers to language that is communicated within the mediated context of social media, where language is produced and interpreted within a specific environment that shapes and is shaped by users’ behaviours and platform technologies.

Although social media discourse has received ample attention since the early days of the internet, discourse analysts have focused initially on the texts generated by users and approached the technology as part of the context (Georgakopoulou and Spilioti 2016; Herring 2007; Tannen and Trester 2013; Vásquez 2022). As a result, the discourse of platforms, as well as their design and how they shape the texts generated by users, has been relatively overlooked. Nevertheless, within any communication situation, different social actors, environments, and technologies allow for various actionable possibilities. In order to illustrate this, Eisenlauer (2014:73) discusses the scenario of a newspaper being given to a toddler; the newspaper most probably will not afford the act of reading but would be suitable to crumple up or rip apart. In the context of media communication, Hutchby (2006:21) also argues that ‘different genres of media talk have their own frameworks of participation and dynamics of address that operate within, and necessarily shape, the “message” that reaches the audience at home’.

Similarly in social media, discourse opportunities are facilitated and delimited by the technological environment which mediates the communication between profile owners and recipients (Eisenlauer 2014:73). In other words, the content generated and shared online is not

just authored by the users but is co-shaped by the design of the platform itself. The impact of platform design on social media discourse is particularly evident in metadata, defined in previous literature on Twitter as ‘characteristics of a Twitter profile that (are independent of tweet content)’ (Morgan-Lopez et al 2017:3). Metadata on Twitter profiles can be understood as any data not included within a tweet, such as the number of followers that an account has and the profile picture that it uses. These features are facilitated by Twitter: accounts always have a profile picture (even if it is the default image) and display follower numbers. According to Eisenlauer (2014:78), social media metadata can be split into two categories: automated text actions and creative text actions. Automated text actions are those that rely on a standardised, automatic action run by the systems of the platform, but are activated by human interaction. For example, pressing a ‘like’ or a ‘follow’ button is an example of an automated text action: a human presses the button, but the system automatically generates ‘@user likes X’ and adds the liked tweet to the list of ‘likes’ on a Twitter user’s profile. Similarly, if a user presses ‘retweet’, the system automatically reposts the content to the user’s own profile. In contrast, creative text actions are those that are more closely controlled by users. One such example is the writing of bio information: there is a bio space and a blank template where the user can write what they choose.

Although users have the power to populate creative text actions with their own content, the information used in usernames, handles, and profile pictures remains predetermined to some extent by the platform and its norms as there is a presupposition that these features should be used for identification. This is demonstrated in the prompts given to users when creating their profiles; users are encouraged to use their ‘favourite selfie’ and pick a username that will identify them to their friends (Twitter 2020). The expectation for metadata to be used as an identification resource is reinforced by proposals in recent years that social media platforms should provide users with the options to verify their profile pictures to ‘ensure it is a true likeness’, or to use ‘two factor authentication’ using a phone number or government-issued document like a passport (UK Government Department for Digital, Culture, Media and Sport 2022).

The significance of metadata discourse is illustrated in previous research according to which metadata features like usernames and profile pictures can influence the way that Twitter accounts are perceived and assessed by users (Andrews et al 2016; Jäkälä and Berki 2013; Morris et al 2012). For example, Morris et al (2012) study how profile pictures and naming strategies influence perceptions of tweet credibility in the context of breaking news on Twitter. They find that users with anime characters or the default image as profile pictures tend to be

deemed less credible than those who use photos of people. Moreover, the kinds of photos that people use also appear relevant; according to one participant, a particular tweet seemed untrustworthy because the man in the photo ‘looked like a stalker’ (Morris et al 2012:443). The idea that certain kinds of profile pictures can be understood as potential markers of nefarious activity is reinforced within the social media discourse of Twitter. In 2017, in a bid to reduce abuse and spam, Twitter started providing users with the option to mute notifications from users who have a default profile picture (Newton 2017). This update to the platform suggests a ‘symbiotic’ relationship between user behaviours and platform affordances (Jones 2009:116), where normative social discourses emerge; although the default Twitter profile picture may initially have been just another identification option, it now appears relatively established that such images will be assessed with suspicion.

Research on social media usernames also illustrates the impact that metadata units have on the content generated by users and, particularly, how it is received and understood by other users. For example, Morris et al (2012) reveal that the kinds of names users identify themselves with have an impact on assessments of credibility. In their study, they differentiate between internet, traditional, and topical names⁵ and find that internet names were viewed as having the least credibility when users were assessing tweets about breaking news. In contrast, when tweets were topically related to tweet content, these names contributed credibility to the information they shared (e.g., @celebnews posting news about a celebrity). Similarly, Andrews et al (2016) investigate the sharing of false rumours on Twitter and find that the tweets responsible for the highest volume of sharing and interaction are those that mimic traditional news distributors via their handles and usernames. It therefore appears that within the context of Twitter breaking news, naming strategies that are topically related to the information being shared and orient to existing practices of Twitter news accounts can be associated with increased trustworthiness, credibility, and influence.

Other research suggests that where ordinary usage is considered (i.e., where breaking news is not being assessed), names that use conventional, ‘traditional’ (cf. Morris et al 2012) practices can be drawn on as a means of authentication. Jäkälä and Berki (2013) differentiate between different kinds of online identities and suggest that naming practices play a role in these identity

⁵ ‘Internet name’ was used to refer to names that reflected a style commonly associated with the internet, such as those that include numbers, nonstandard grammar, and words that are not used as names offline. Examples given included *tenacious27* and *25th_Hour* (Morris et al 2012). ‘Traditional name’ was used to refer to names that used a conventional first name and surname structure such as *Alex_Brown*. Finally, ‘topical name’ was used to refer to names that had topical relevance to the subject of the tweet such as *AllPolitics*.

performances. They differentiate between 5 categories of online identity performance: eponymity, nonymity, anonymity, pseudonymity, and polynymity. Put briefly, these categories refer to the degree of personal information that a user shares about themselves. Anonymity and nonymity refer to situations where users do not identify themselves with any kind of name, while polynymity refers to the use of various names across different contexts. Eponymity is used to refer to the identification of oneself using first and/or surname, while pseudonymity is used to refer to the adoption of a pseudonym that is not someone's 'real, correct name' (Jäkälä and Berki 2013:8). According to Jäkälä and Berki (2013:7), the use of eponyms (names) can create a sense of familiarity between the 'real' and 'virtual' worlds, can be used to identify and authenticate online users, and can consequently be associated with increased trustworthiness. In contrast, pseudonyms can be viewed as a means of concealing a person's identity with 'fabricated, artificial and false characteristics' (Jäkälä and Berki 2013:8) and are thus often regarded suspiciously.

The ability to conceal one's 'real, correct' name (cf. Jäkälä and Berki 2013) with another on social media is often drawn on in discussions of online trolling and bullying within the mainstream media, where it has been suggested that 'vile online trolls' are emboldened on social media by the relative anonymity afforded to them by social media platforms (Lester 2021; Knowles 2022). Consequently, within every day, ordinary usage of Twitter, it appears that the use of pseudonyms or obvious obfuscators of personal identity can be viewed with suspicion, while the use of first names can authenticate identity performances by reducing perceptions of anonymity and thus orienting to accepted, established naming practices.

The above discussion suggests that within social media discourse, metadata units play a role in identification and, thus, identity performance, as well as in the audience's uptake of such performances. The interplay between the predetermined technological features of the site of Twitter and the way that users employ and experience the site shows how normative discourses and practices on social media often emerge and develop while users engage with and interact with the platform design and the acts it affords. As a result, this study will pay attention to both the technological discourses that shape identification in the context of Twitter (i.e., metadata units) and the discourse produced in the disinformation account posts.

2.6 Conclusion

The review of existing literature on authenticity has demonstrated a range of approaches to the concept which primarily view authenticity either as an essential quality or as a discursive performance. In line with the second perspective, I approach authenticity as a dynamic

performance constructed through discourse and define it as valid representation of “realness” which, in the context of this thesis, takes the form of “real personas” (e.g., “real conservatives”, “real liberals” etc.). While an essentialised perspective to authenticity would consider it futile to examine known disinformation accounts through this lens as there is no correspondence between the alleged account holder and an offline, embodied self where the ‘real’ can be located (cf. Lovelock 2016:206), van Leeuwen (2001:396) has already argued that we need to ‘ask not “how authentic is this”, but “who takes this as authentic and who does not” and “on the basis of which visible or audible cues are those judgements made”’. Given that these accounts were taken as authentic, at least by mainstream media that used some accounts as evidence of public opinion (see section 2.3.3), my research will focus primarily ‘on the [...] visible or audible cues’ on which those judgements were made, to use van Leeuwen’s (2001:396) words.

In my study of the process through which disinformation accounts authenticate their identity performances as “real” U.S. personas, I will pay attention to cues that may appear in both metadata units (platform design) and tweets (user-generated content). As discussed in section 2.4.2, this is an approach that has been rather rare in existing disinformation research that appears to examine either technological aspects of content propagation primarily through a computer science perspective or user-generated content from a discourse analytic perspective.

In order to examine how disinformation accounts operated by the IRA constructed authentic identity performances, I will specifically analyse:

1. How are handles and usernames employed for authenticating particular personas on Twitter disinformation accounts?
2. How are profile pictures used for authenticating particular personas on Twitter disinformation accounts?
3. How do disinformation accounts manage (co-)tellership to (dis)identify with different social groups when storying news events such as the Charlottesville Unite the Right Rally?

In chapters 4-6, I aim to answer these questions by analysing firstly handles and usernames, secondly, profile pictures, and finally, the management of (co-)tellership when narrating a news event. In the following chapter, I discuss further the methodological approach undertaken for the collection and analysis of the disinformation data. I provide an overview of the disinformation accounts and data collected for the purposes of this thesis and I discuss the methods implemented for the analysis of the data in each chapter.

Chapter 3 Researching Twitter Disinformation Accounts: Data and Methodology

3.1 Introduction

To explore how authenticity was performed on Twitter disinformation accounts, this thesis draws on a set of Twitter accounts that have been identified as being part of an organised disinformation campaign associated with the Internet Research Agency. The IRA were a Russian disinformation organisation that will be discussed in more detail in section 3.3.2. From this wider dataset, specific samples have been selected for further analysis to answer the three main sub-questions (see section 3.4) in chapters 4-6. This chapter presents an overview of the samples and methods used in the three analytical chapters. In addition, each analytical chapter will provide further detail about the specific analytical methods employed and their contribution to the study of authenticity in the context of disinformation communication. This chapter begins with a discussion of the overarching principles of digital discourse analysis that guide this study (section 3.2). Section 3.3 discusses and critically reflects on the data samples used in this study and the context within which the data is situated. In section 3.3.4, I explain the ethical considerations made when conducting this research. Finally, sections 3.4.1, 3.4.2, and 3.4.3 provide an overview of the samples and methods used in each analytical chapter.

3.2 Digital discourse analysis

My research draws on qualitative methods and descriptive statistics situated within the broader area of digital discourse analysis. As with earlier approaches to digital discourse, such as Herring's (2007) computer-mediated discourse analysis (CMDA) model, digital discourse analysis does not refer to one strict method but is an approach that brings together language-focussed paradigms that can be used to analyse online verbal interactions. Herring (2004b:339) suggests that in the broadest sense, 'any analysis of online behaviour that is grounded in empirical, textual observations is CMDA', with a CMD research agenda being empirical study of CM language and varieties of CMD. This research, as an investigation into how disinformation accounts used language to authenticate identity performances, consequently fits the broad definition introduced by Herring (2004b) along with more recent pluralistic approaches to digital discourse analysis that aim to aid better understanding of how 'our contemporary realities are shaped by and reflected in online texts and digital practices alike' (Vásquez 2022:2).

Digital discourse analysts acknowledge that in the online sphere, identity performances are created not solely through language, but through ‘the use of technologically specific semiotic resources’ (Vásquez 2022:6). The relevance of technological features, together with other contextual aspects, is evident in my study in my attention to platform-specific practices and choices such as metadata (in the form of profile handles, usernames, and pictures) and co-tellership actions (e.g., ‘retweet’ function). The need to pay attention to aspects related to the specific social media platform under study is also advocated by Jones et al (2015:1) who suggest that the differences between media platforms - or ‘sites of display’, as they call them - encourage the combination of tools of analogue analysis with ‘new concepts and new methodologies to address the unique combinations of affordances and constraints introduced by digital media’.

In this line of research, social media platforms as ‘sites of display’ are argued to be ‘inherently ideological’: their affordances and constraints enable and contribute to the construction of social realities which in turn influence what kinds of practices and actors are accepted, and which are not (Jones 2009:116). Jones (2009:116) consequently argues that platforms and the social practices associated with them are in a symbiotic relationship, where the platforms amplify and constrain social practices, while ‘social practices affect the kinds of functionalities sites of display develop’. In other words, platforms and the behaviours exhibited on them influence one another. This more nuanced, symbiotic, and inter-related connection between technological affordances or constraints, and discursive practice which elaborates on earlier CMDA claims and is developed by Jones (2009), is also assumed in this study.

Thurlow (2017:136) acknowledges the value of Herring’s contribution to the area of digital communication and elaborates on her work by suggesting three broad organising principles for any research into language and social media: discourse, multimodality, and ideology. Understanding discourse as ‘the social function of language, the interactional accomplishment of meaning, the significance of communicator intent, and the relevance of context’ (Thurlow 2017:137), this thesis also pays attention to how discourse (in the form of naming strategies and storytelling) is being used to create meaning within both the context of Twitter and broader contexts relating to the socio-political climate at the time the data was created. Furthermore, acknowledging multimodality, i.e., the range of other semiotic systems that contribute to meaning making by interacting with language, this study will focus and analyse profile pictures as a means of identity performance and meaning making. Finally, Thurlow (2017:138) sees ideology not only in ‘a) the ways that micro-level interactional and social practices constitute our social worlds and b) the ways our everyday communicative/representational practices are structured by larger systems of belief and hierarchies of knowledge’ but also in ‘social media’

due to their potential to control both ‘normative and counter-normative representations’. This thesis pays attention to ideology by acknowledging the norms that appear to govern social media use, both in terms of user behaviours and what is enabled by the platform itself, and by acknowledging broader dominant discourses related to social groups.

The relevance of social media platforms in these three principles is evident. As Thurlow and Mcozbek (2011:xxv) argue, technologies should be approached as ‘prosthetic extensions of people’s abilities and lives, rather like a hearing aid and the paper clip’. The need for an approach that acknowledges the power arising from the symbiotic relationship between technology and users is further highlighted by the shift and blur of traditional power dynamics associated with senders and receivers of information on social media (Antony and Thomas 2010:1283). For the first time ever, citizens can not only share and discuss but potentially create and structure news items. Such a shift in power dynamics is exemplified in Twitter discussions of the trial of a police officer who killed an unarmed Black teenager in Ferguson in 2014. According to Groshek and Tandoc (2016), non-journalists were the most influential sources of online news during the trial, largely because of the interactivity affordances of Twitter as a platform. It is in this communicative dynamic, where ordinary citizens both create and disseminate news, that Russian disinformation accounts operated and managed to disseminate harmful information while masquerading as American citizens.

Overall, this thesis draws on digital discourse analysis whereby methodological tools developed for non-digitally mediated communication (e.g., narrative analysis, social semiotics) are applied and adjusted to the study of social media discourse and communication. In the application and adjustment of such methods to the study of Twitter disinformation accounts, my study is also informed by Jones’s (2009) approach to social media as sites of display and, as a result, takes into account the platform affordances that facilitate and constrain opportunities for identity performance.

3.3 The data: Twitter disinformation accounts

3.3.1 Context

The growing use of social media as a platform for waging ‘information warfare’ and manipulating public engagement in politics has received increasing levels of attention in recent years, so much so that information warfare, online disinformation, and information distortion have been continually included in the World Economic Forum’s global risk reports since 2017. High profile instances of the use of disinformation campaigns to disrupt society and democratic

processes include the storming of the Capitol Building in the US following the spread of the inaccurate claim that the 2020 election result had been illegally ‘stolen’ from Donald Trump by the left (Nawaz 2022), and the vandalism and destruction of 5G towers in the UK following claims that the towers damage the immune system and cause COVID-19 (BBC News 2020a; Nogara et al 2022). Disinformation is therefore not a local phenomenon tied to one country or culture. There has been evidence of its existence and influence across both East and Western Europe, including Estonia, Latvia, Lithuania, Ukraine, Moldova, Belarus, Czech Republic, Slovakia, France, and Italy (Fletcher et al 2018; Helmus et al 2018; Smoleňová 2015). In 2016, the issue of disinformation on social media received increased public attention following claims of Russian interference in the 2016 US Presidential election. This thesis focuses specifically on a set of disinformation accounts that were identified as being part of a co-ordinated Russian disinformation campaign that targeted the aforementioned election and its aftermath.

3.3.2 The datasets

This thesis draws on a dataset of 3,836 accounts identified as distributors of disinformation, specifically those associated with the IRA, a ‘troll factory’ where bloggers were paid to circulate pro-Russian news viewpoints and attack perspectives perceived as Anti-Russian (Mejias and Vokuer 2017). The data is publicly available (Twitter 2022a) but was provided to me in Microsoft Excel format by a Research Institute within Cardiff University. In this section, I will explain the scope and size of the dataset(s) used in this thesis.

The origin of the dataset used in this thesis can be found in the initial list of 2,752 accounts that Twitter identified as being associated with Russian disinformation and submitted to the Senate Select Committee on Intelligence (2017) as part of an investigation into claims of Russian interference in the 2016 US election. Testimony to the Committee from Twitter’s general counsel, Sean J. Edgett, details how Edgett examined the content of accounts, using mentions of the American election combined with technical associations with Russia e.g., Russian email addresses and phone numbers, to put together the submitted list of accounts. Once Twitter identified the accounts, they suspended them and deleted their tweets from public view (US Senate Committee 2017). Twitter continued to carry out this process and in October 2018 released a more substantial dataset of 3,841 accounts affiliated with the IRA and 770 other accounts that potentially originated in Iran. This large dataset included more than 10 million tweets and more than 2 million images, GIFS, videos, and Periscope broadcasts (Gadde and Roth 2018).

Based on this data release, Cardiff University's Crime and Security Research Institute (CSRI) compiled their own comprehensive dataset of Russian disinformation data which was made available to me in January 2019. The CSRI dataset includes 3,836 Twitter accounts linked with Russian IRA activity. This data was presented in Microsoft Excel format and provided an initial insight into potential areas for further exploration (displayed in Screenshot 1, Appendix). The dataset is comprised of metadata for each of the accounts, specifically, the reported locations, bios, following and follower counts. With respect to handles and usernames, they are also included in the Excel sheet for accounts with over 5,000 followers, but they are 'hashed' for accounts with less than 5,000 followers (i.e., usernames and handles had been changed into encrypted codes). This dataset contains accounts that had been active since 2009, with the majority of accounts appearing to have been tweeting from 2014 until the moment they were suspended by Twitter.

These accounts initially tweeted about mundane topics 'consistent with the spoofed owner's persona' to gain a following, and then eventually switched narratives to begin disseminating political disinformation (Dawson and Innes 2019:249). On average, the accounts were most active during the 2016 US Presidential election and the following summer, in 2017. The account activity began to decline at the start of 2018 and none of these accounts are still active. Previous research into datasets linked to the IRA find that the accounts tended to post partisan content and could broadly be associated with right or left-leaning ideologies (Linville and Warren 2018). Their ultimate purpose was to push ideologies that supported Russian political interests, discredit 'anti-Kremlin candidate', Hilary Clinton, and undermine public faith in the U.S. democratic process (Polyakova and Boyer 2018:2).

Considering that the initial data of 3,836 accounts and the more than 3,000,000 tweets these accounts posted would be impossible to analyse from a more detailed discourse analytic perspective, I decided to compile a smaller dataset that includes the Twitter activity of five popular accounts in the sample. These accounts were chosen due to their popularity not only in terms of follower numbers but also within the mainstream media and popular culture. As a result, although the five accounts appear to vary in terms of follower numbers (see Table 2), they were all within the list of the top ten most-followed accounts in the sample. The reasons for choosing these five accounts will be explained in more detail below. For the compilation of this dataset, I was also supported by colleagues at the CSRI who provided me with a Microsoft Excel spreadsheet that contained all the tweets the five accounts posted from their date of activation to the date that they were suspended from Twitter (26,845 tweets) (see Screenshot 2, Appendix). In addition to the tweets, this sub-set of data includes (i) the date and time each

tweet was posted; (ii) how many retweets, likes, quotes, and replies the tweet received; (iii) details of any other accounts mentioned in the tweets; and (iv) information as to whether the tweet itself is a retweet, quote tweet, or a reply. The dataset also contains information as to whether the tweets include images, GIFS, videos, or URLs. The CSRI provided separate files that contained any attached additional media.

The accounts in this sub-set operated between 2015 and 2018. As shown in Table 1, the overall number of tweets for each account varies, with @TEN_GOP being responsible for most tweets, and @wokeluisa tweeting the least. With respect to the overall number of followers and general uptake of their tweets, Table 2 shows the number of followers, highest retweet count and the highest reply count that content from each account has received, evidencing the specific accounts’ reach and density of engagement by other Twitter users. Sections 3.4.1-3.4.3 provide more detailed explanations of how the aforementioned datasets were used in each analytical chapter.

Table 1: Summary of tweet numbers for the five key accounts

Account	Number of tweets
@TEN_GOP	10,794
@Pamela_Moore13	6,203
@wokeluisa	2,288
@SouthLoneStar	3,600
@USA_Gunslinger	3,960
Total	26,845

Table 2: Summary of engagement rates for the five key accounts

Account	Follower count	Highest retweet count	Highest reply count
@TEN_GOP	147,767	25,790	3,249
@Pamela_Moore13	72,121	11,972	2,039
@wokeluisa	57,295	51,472	2,173
@SouthLoneStar	53,999	8,243	1,195
@USA_Gunslinger	44,895	8,310	630

One of the reasons why these five accounts were selected was their influential status during their time of operation. As shown in Table 2, there have been moments where content posted by a specific account was retweeted by 51,472 other accounts or received more than 3,000

replies. Profile characteristics such as follower counts have previously been stated to be a measure of influence as they ‘define how many users receive a message’ (Ausserhofer and Maireder 2013:293), while replies and retweets demonstrate an active audience. Thus, the follower counts and interactivity statistics of the five users indicate how influential they were during their time of operation.

The selection of these accounts is also related to the fact that they received attention within the mainstream media and were interacted with by high-profile public figures. For example, a tweet by @SouthLoneStar appeared in British tabloids as evidence of an online backlash against a Muslim woman pictured walking along Westminster bridge following the Westminster terror attack in 2017 (Duell 2017). Similarly tweets from @wokeluisa featured in news stories published by the BBC, USA Today, Time, Wired, HuffPo, BET and others (O’Sullivan 2018). @TEN_GOP was retweeted and followed by a number of prominent American political figures, including Michael Flynn (former national security adviser), Donald Trump Jr., Kellyanne Conway (Trump’s campaign manager), and Brad Parscale (digital director and Trump’s 2020 campaign manager) (Kosoff 2017). The remaining two accounts, @Pamela_Moore13 and @USA_Gunslinger, also featured in news stories: @Pamela_Moore13 was referred to as a ‘popular conservative Twitter personality’ (Tornoe 2017) and was followed by Michael Flynn, while @USA_Gunslinger featured in an Australian news report, with the account’s 70,000 followers being discussed as evidence of the successful infiltration of disinformation accounts into mainstream society and media use (Seidel 2017).

In terms of political orientation, the specific accounts appear as largely right-leaning, with bios that declare stances that are pro-life, pro-Trump, against political correctness, anti-Islam, and in support of the Second Amendment right to bear arms. Only one of these accounts masquerades as being left leaning (@wokeluisa), posing as a young Black female political science major from New York who discusses racism, Black excellence, and support for Hillary Clinton amongst others. The dataset therefore contains Twitter accounts from both ends of a polarised political spectrum. Although the distribution of political affiliation may appear unbalanced, the sample does reflect wider patterns in the list of disinformation profiles, as according to Linvill and Warren (2018) there were more than twice as many ‘right troll’ accounts than ‘left troll’.

The data collection process for this thesis consequently involved three datasets: (i) a dataset publicly available and released from Twitter, (ii) a sub-set of this data created by the CSRI, (iii) a dataset relating to five popular accounts from the sub-set provided by the CSRI. I then

undertook further data collection and research for each chapter. These processes will be explained in more detail in sections 3.4.1-3.4.3.

3.3.3 Using pre-existing datasets: opportunities and challenges

To investigate authenticity in disinformation accounts, it was firstly necessary to find a dataset that contained texts created and shared by users known to be involved in disinformation distribution. Doing this in any other way than using logged data from previously existing datasets would have been problematic for two reasons: first, because Twitter disinformation accounts aim to go undetected in order to garner as much influence as possible, and so they would never reveal themselves as such; and second, because verifiable disinformation accounts (needed for a reliable and valid dataset), are quickly disabled.

It was consequently inappropriate, as well as not feasible, to create my own dataset from active Twitter accounts. The need to draw on and employ existing datasets of disinformation accounts is also evident in previous research. While there have been numerous methods employed to analyse disinformation data, the data collection process for a number of studies involved the use of the same large corpus of IRA-related accounts that laid the foundation for this study (Arif et al 2018; Freelon and Lokot 2020; Linvill and Warren 2018, 2020; Stewart et al 2018; Xia et al 2019; Zannettou et al 2019). Although it may be problematic to employ the same dataset, the size and scope of the corpus enables the investigation of both the same logged data and various smaller sub-sets of relatively underexplored data from multiple and complementary perspectives and fields.

While studies into active Twitter accounts as sources of disinformation have become increasingly prevalent in recent years, it is arguably difficult to attribute intention to tweets and accounts that are still active on the platform. In Golovchenko et al's (2018) study into disinformation surrounding the crash of MH17 over Ukraine, tweets were defined as disinformation if they attributed responsibility to Ukraine rather than Russia based on the Joint Investigation Team (JIT) report (created by police and judicial authorities from the Netherlands, Australia, Belgium, Malaysia, and Ukraine) that established Russian responsibility. While such a decision can be understood when considering the nature of the JIT report as an objective forensic account, the acknowledgement by Golovchenko et al (2018), that some items may have been shared by users without them knowing that the information was false, can be viewed as problematic in a study that aims to research disinformation, rather than misinformation.

Many studies have made a distinction between mis- and dis- information. Although citizens may take part in the spread of disinformation messages via retweeting and repeating false claims, such instances could be defined as misinformation, i.e., spreading information that is 'initially presumed to be true' (Lewandowsky et al 2017:355). In contrast, disinformation as a term refers to false information which is spread with the **intention** of causing disruption (Bennett and Livingston 2018). The difficulty of establishing intention online is further highlighted by recent research into active Twitter accounts engaged in the spread of inaccurate and potentially harmful narratives relating to COVID-19 vaccines. While these studies provide a valuable insight into how information can spread across social networks, the terms mis/disinformation appear to either be conflated (Scannell et al 2021), or misinformation favoured (Jemielniak and Krempovych 2021; Nogara et al 2022), arguably due to the aforementioned difficulty in attributing intention to tweets written by unknown users. In the case of my research on authenticity performances on disinformation accounts, it is important to work with data that has been established as intentionally misleading. For the reasons discussed above, it would be inappropriate to attempt to use active Twitter accounts.

Overall, the use of a pre-existing dataset has various strengths. On the one hand, the fact that the specific dataset proved fruitful for other studies suggests it is an appropriate and relevant dataset to use when investigating issues related to disinformation. On the other hand, the use of such a dataset speeds up the research process as it reduces the time the researcher spends in obtaining and compiling a dataset. While the data still require a level of cleaning (e.g., removing data I would not use), which will be discussed further in sections 3.4.1-3.4.3, the data was easily accessible and will remain so for others who wish to research the data further.

Having said this, it is very rare that a pre-existing dataset will be used exactly as it is. It may prove insufficient and further research may be needed to enrich it. In this case, the dataset provided by the CSRI did not contain profile pictures, nor did it contain information about the change of handles or usernames. A closer investigation of the data also revealed that some pre-existing categorisations relating to tweet format were unreliable; not all tweets that were quote tweets were labelled as such, and where tweets were replies, the original tweets being responded to were not included. Consequently, for each research question, I closely investigated the data and extended the datasets where necessary, obtaining extra information needed for my analysis and checking any pre-existing categorisations that were made within the existing dataset. It therefore appears that while using pre-existing datasets can save time for researchers, it is necessary to closely (re)investigate the data in order to ensure an accurate dataset that fulfils

the needs of one's particular research aims. Without close inspection, the time-saving benefits of using pre-existing data can be at the expense of accuracy and reliability.

3.3.4 The ethics of researching Twitter disinformation accounts

My research on Twitter disinformation accounts has been conducted in line with the ethical guidelines of Cardiff University's School of English, Communication and Philosophy (ENCAP) Ethics Procedure (2021/22). The data used in this study was all at one time self-published to Twitter, a public platform that does not require a password to view. The notion - that once posted, material can be accessed by others on Twitter and through third party services-, is explicitly acknowledged within Twitter's (2022b) Privacy Policy. Most of this data remains publicly available now on various websites (Twitter; Defending Democracy Together (see page <https://russiatweets.com/>); FiveThirtyEight (see page <https://github.com/fivethirtyeight/russian-troll-tweets>) meaning that I did not have to enter into any private spaces to gain access to data that users did not know was being posted to a public platform. The emphasis placed on the distinction between public and private spaces is summarised by Spilioti and Tagg (2022:96): 'the general principle is that the more public the site and the more open the access to it, the less urgent is the need to protect participants' privacy'. Of course, this principle is considered alongside the specific context of data, and where participants are vulnerable or the data is sensitive in nature, participants' privacy should be protected in order to reduce any risk of potential harm.

While the majority of the data is publicly available and directly linked to the IRA or public figures, some data produced by private individuals has been obtained. In line with the guidelines of Cardiff University's ENCAP Ethics Procedure (2021/22), data relating to private individuals has been handled sensitively, without the reproduction or sharing of information that could relate to the identification of any individual. To protect the data and the identity of participants, the data has been stored in password-protected files.

This data has not been gathered from a group of participants who have given their consent for the information to be analysed. Such a scenario is not uncommon within social media research, where data is often considered public due to its self-publicised nature into a public setting (Patterson 2018) and where there are significant impracticalities -and in some cases potential for harm- associated with contacting users directly to gain their consent (Bishop and Gray 2018:175; Solberg 2010:318). According to recent ethical recommendations within linguistics, all research involving human subjects requires informed consent from participants (Spilioti and Tagg 2022:95). As a result, influential studies such as Labov's (1986) department

store work may be considered unethical by modern standards due to his observation and analysis of participant's speech without their knowledge (Meyerhoff 2016). While in the offline sphere, this may appear relatively straightforward, Spilioti and Tagg (2022:95) acknowledge that in digital research, the understanding of what a "human subject" is, and thus, how they should be treated, can become complicated. Within the ongoing discussion of digital ethical practice, some argue that the practice of collecting data from participants without their consent should be considered unethical, just as it would be offline. Williams et al (2017) argue that although users may have posted something into the public domain, the material still belongs to them, and consequently they should have control over what happens to it.

Such a broad generalisation of ethical guidelines to the heterogeneous arena of online communication has been argued by some to be problematic. According to Herring (2011:154), 'rules that seek to generalise indiscriminately across all varieties of CMC do not "fit" the nature of the phenomenon', suggesting that the application of ethical guidelines should pay attention to the specific context of each case under study (cf. BAAL 2021). As summarised in the AoIR Ethical Guidelines (2019:6), 'the best we can do is to "develop guidelines, not recipes"'. A common strategy to protect user privacy and confidentiality is to anonymise datasets and minimise any harm that may arise if third parties trace back material to its authors. In the instances where researchers are more interested in linguistic patterns across groups of users, rather than the identity of those who posted (Herring 2011), concealing any identifying information is recommended as an appropriate practice for protecting participants from potential harm or risk, while at the same time ensuring thorough and robust linguistic investigation.

In the case of disinformation research, tweets come from accounts that belong to fictional personas who do not exist as "real" persons in the physical world, i.e., they were invented. In this way, the authors already anonymised themselves; even if the profile names and handles are published, the data cannot be tracked back to any specific individuals in the physical world. While in other situations it is broadly recommended that usernames and online personas are anonymised/pseudonymised (AoIR 2019:41), in this specific dataset, the nature of the data as operated by an organisation (the IRA) reduces the risk of harm associated with the publishing and use of online usernames. Moreover, the accounts are no longer active, so it would be impossible to use identifying information to find the accounts and communicate with the authors behind the pages. While the use of data that has been removed from Twitter arguably raises questions in relation to the right to be forgotten, the nature of these accounts as parts of a co-ordinated disinformation effort appear to require particular consideration. Firstly, the

accounts were removed by Twitter rather than the authors. Secondly, the public interest in the data as an attempt to disrupt democracy arguably outweighs the right to be forgotten in this instance. Such considerations encouraged my decision to proceed with my data collection and analysis without seeking informed consent.

Finally, there is precedent for researching, using, and publishing this data without seeking informed consent, and despite it no longer being active online. An exhaustive list of disinformation accounts, content, and statistical information can be found online on websites published by FiveThirtyEight and Defending Democracy Together. The information is therefore already within the public sphere, can be considered publicly available and thus, available to use in research (Patterson 2018).

While the dataset is largely comprised of tweets created and posted by the IRA, the tendency for users to interact with one another on Twitter means that some of the data includes content produced by other users; the disinformation accounts may have retweeted content produced by ordinary users, for example. While the data produced by these users was publicly posted to Twitter and could therefore be argued to be within the public domain, it is important to acknowledge that since the creation of the dataset, these users may have deleted the content that they shared, deactivated their Twitter accounts, or made their profiles private. It has been suggested that within the ever-changing nature of accessibility online, there is a need to ‘re-ethicise’ our approach to social media data when it has been removed from the public domain (Spilioti and Tagg 2016:166). Debates around the ‘right to be forgotten’ suggest that users’ withdrawal of their data from the public sphere could also be interpreted as a withdrawal of consent to be studied (Tromble 2017:2). Moreover, it appears ethically precarious to publicise data connected to ordinary people in relation to disinformation campaigns used as tools for political disruption. Especially now that disinformation has become a popular topic of public concern, users are entitled to protection from any potential harm that may come from their association with a known disinformation campaign.

Taking into consideration the above issues, this thesis avoids reproducing examples of tweets that have been produced by ordinary people e.g., in the form of retweets. Where such examples are included (e.g., in chapter 6), user handles are omitted, and any examples are carefully selected so that potential identification of specific individuals is not possible. Further details of how data produced by ordinary people has been managed will be provided in the discussion of each chapter’s data collection and coding processes (see section 3.4.1.1, 3.4.2.1, 3.4.3.1). Moreover, the study makes use of data produced by public figures that is no longer available on Twitter. In the case of Donald Trump, for example, Twitter suspended his account and thus,

removed his tweets from the platform. While the tweets are no longer in the public domain where they were originally published, the data still remains within the public record in other forms (such as mainstream media news reports), and they are associated with a public figure (Trump) rather than a private individual. More generally, scholars have argued that the benefits of disinformation research outweigh any minimal risks caused by the management of such data. The efforts by disinformation accounts to disrupt democracy are such a clear demonstration of unethical behaviour that it is within the public interest to tackle this issue as comprehensively as possible (Froehlich 2017).

3.4 The methods: Investigating authenticity on Twitter disinformation accounts

As mentioned in Chapter 2, my study of the discursive performance of authenticity in popular disinformation accounts involves the detailed examination of the following three inter-related areas:

1. How are handles and usernames employed for authenticating particular personas on Twitter disinformation accounts?
2. How are profile pictures used for authenticating particular personas on Twitter disinformation accounts?
3. How did disinformation accounts manage (co-)tellership to (dis)identify with different social groups when storying news events such as the Charlottesville Unite the Right Rally?

These questions are addressed in the following three analytic chapters, where specific subsets of the overall data sample are analysed depending on the aim of each chapter. In the following sections, I provide a brief overview of the samples and methods used in each chapter.

3.4.1 Collecting and analysing handles and usernames

3.4.1.1 The sample

The sample used for the analysis of handles and usernames includes first and foremost data from the five main accounts: @TEN_GOP, @Pamela_Moore13, @SouthLoneStar, @wokeluisa, and @USA_Gunslinger (see section 3.3.2). In my research, I do not only focus on the data employed at a specific point in time, but for each account I searched all the potential handles and usernames that have been used while the account was active. Drawing on previous literature (John 2012:171) that examined texts produced on social media pages at different times, I have used the Wayback Machine (see page <http://web.archive.org/>)- software that

enables a user to access a snapshot of a webpage at different points of time, even if the page is no longer accessible. The decision to take snapshots of the accounts at different time periods is informed by previous research that states that accounts can change their profile information in line with agendas they aim to push at any given time, such as narrative switching or follower fishing (Dawson and Innes 2019:246).

In addition to the handles and usernames associated with the five key accounts, and in an attempt to situate these findings in the context of disinformation accounts, I have also collected the handles and usernames employed by accounts from the wider data sample. As a result, the sample used in this chapter also includes the handles and usernames of all accounts that had more than 5000 followers⁶ (167 profiles). The 167 accounts were filtered down by excluding profiles that used languages other than English. The filtering process resulted in 86 accounts, including the five accounts in question. While for the five accounts, I searched all potential handles and usernames during their period of activation, only one handle and username was recorded for each account in the wider data sample.

The 86 accounts were further categorised based on the number of followers and were grouped into three large categories: (i) accounts with 100,000 followers or more; (ii) accounts with 10,000 – 99,999 followers; and (iii) accounts with 5000 – 9,999 (see Table 3). This categorisation process was informed by practices established within the social media industry, where social media influencers, for example, can be split up into nano-influencers, micro-influencers and popular or celebrity influencers. According to this categorisation, nano-influencers have between 1,000 and 5,000 followers, micro-influencers have more than 10,000 but less than 100,000 followers, and popular or celebrity influencers have anywhere from 100,000 followers to millions (Lieber 2018; Yang et al 2019). These categories set up clear boundaries at 5,000+, 10,000+ and 100,000+ followers which have also been used in this thesis. The table below depicts the number of disinformation accounts (including the five main accounts) that belong to each category.

Table 3: Distribution of disinformation accounts according to number of followers

Number of followers	Disinformation accounts (n)
>100,000	1

⁶ Within the data sample provided by the CSRI, accounts with less than 5,000 followers were anonymised and thus information pertaining to the usernames and handles of these accounts are not available. Although this data may be valuable, it appears more appropriate to focus this investigation upon those with more than 5,000 followers in order to allow reasonable comparison with more popular accounts.

10,000 - 99,999	62
5,000 – 9,999	23
Total	86

Throughout this study, I will differentiate between the five accounts focussed on throughout the thesis, and the wider sample of other disinformation accounts (81). Within the wider sample, I will primarily distinguish between ‘more popular accounts’ i.e., accounts with tens of thousands of followers (n: 58) and ‘less popular accounts’ i.e., those with thousands of followers (n:23). For the analysis of naming strategies in chapter 4, the handles and usernames of the 81 accounts in the wider sample were collected and stored on a Microsoft Excel spreadsheet (see Screenshot 3, Appendix).

3.4.1.2 The method

The analysis of handles and usernames draws on qualitative methods while also taking a quantitative approach to coding. The qualitative part of the analysis focused on the five main accounts and was used in order to reveal the categories that would be used for the analysis of the wider sample of 81 remaining disinformation accounts using descriptive statistics. For both handles and usernames, the analysis focused on two areas: (i) the type of naming strategy employed (e.g., eponym or pseudonym) and (ii) the identity aspect evoked in the data (e.g., gender, location, political affiliation). Further details about the coding, the methods that informed this coding, and the specific tools employed are provided in chapter 4.

Drawing on the categories that emerged from the qualitative analysis and enriching the labels where appropriate, I analysed the identity aspects and the naming strategies used by all 81 accounts in the wider sample using descriptive statistics. Descriptive statistics have been used for this part of the analysis to indicate the percentages of accounts that used different naming strategies (eponyms/pseudonyms) and evoked particular identity aspects. This quantitative part also allowed comparisons between more popular and less popular accounts in terms of their use of handles and usernames.

3.4.2 Collecting and analysing profile pictures

3.4.2.1 The sample

Similar to the study of handles and usernames, the sample of profile pictures included: (i) the profile pictures used by the five main accounts, which formed the basis for the qualitative part of the chapter and (ii) the profile pictures employed by the wider sample of 81 accounts (see Table 3), which were analysed using descriptive statistics. Given that the dataset provided by

the CSRI did not include profile pictures, I searched and collected this sample first through the means of the ‘Wayback Machine’ which provides access to screen grabs of Twitter profiles as they appeared during their activation, thus showing their profile pictures. Where profile pictures could not be found on the Wayback Machine because, for example, the Twitter profile pages have never been recorded, I used a pre-existing repository of data linked to suspicious accounts which included some of the accounts in the CSRI dataset (see page <http://suspicioustweets.archive.org/?userid=all>). Similar to chapter 4, I collected all profile pictures employed by the five main accounts until their suspension, while I only collected one profile picture for each account that belongs to the wider sample of 81 accounts.

For the purposes of studying authentication through profile pictures, I have also investigated the provenance of such pictures, i.e., the extent to which they have been used elsewhere. To explore this, a reverse image search has been carried out on the profile pictures, using TinEye (see page <https://tineye.com>). TinEye invites users to upload an image to their search engine, which then identifies duplicate images and shows other web locations where the image has appeared. The site has been described as a platform that ‘facilitates transparency’ (Zelizer 2017:252) and is described as a preventative measure for avoiding dating scams and inaccurate online information reports and detecting fraud and copyright infringements (Cowan and Haydock 2021; Jarry 2021; TinEye 2022). In my research, the search engine has been useful in aiding the attribution of context to profile pictures. For example, if a profile picture contains a symbol or logo that has 200 other results, the results and the websites where the image is used can inform what the symbol is usually associated with. If an image is of a celebrity, for example, this can give important context to explorations of authentication processes.

The profile pictures for all of the 86 accounts were collected and saved in an images file. The handles and usernames for the wider sample of 81 accounts were stored in a Microsoft Excel spreadsheet and the images file was used for cross-reference (see Screenshot 4 and Screenshot 5 in Appendix).

3.4.2.2 The method

For the analysis of profile pictures, I draw on previous models of visual analysis, particularly Kress and van Leeuwen’s (2006) visual modality model, Ledin and Machin’s (2018) adaptation of Kress and van Leeuwen’s (2006) social semiotic model, and El Refaie’s (2010) work on visual authenticity. Based on these approaches, the close-up qualitative analysis of the profile pictures used by the five accounts focused on the following areas: (i) the objects being represented in the images; (ii) how the objects are depicted; (iii) what sociocultural connotations

the images evoke; and (iv) how the identity aspects signalled in the images interact with those signalled in handles/usernames.

The analysis of areas (i)-(iii) revealed the range of identity aspects evoked by the profile pictures and led to the generation of categories (e.g., person, location, political affiliation) which were then applied to the wider sample. To avoid the difficulties that can come with storing images on Excel, I created a new spreadsheet where the handle and username for each account were stored, and their corresponding profile pictures were saved in a separate images file. I then attributed the labels to the spreadsheet. For the study of the extent to which there is consistency in the identities projected by handles/usernames and profile pictures, I labelled whether the identity aspects being represented in the profile pictures were the same as those presented in handles and/or usernames, or whether they contributed additional identity aspects.

3.4.3 Collecting and analysing tweets from a narrative perspective

3.4.3.1 The sample

For the purposes of chapter 6 that shifts the focus from metadata content to tweet content posted by the disinformation accounts, the sample includes all the tweets that appeared on the five main accounts' feed during the two-week period that followed the Charlottesville Unite the Right rally (i.e., 12th-24th August 2017). The accounts include: @TEN_GOP, @Pamela_Moore13, @wokeluisa, and @USA_Gunslinger. In the case of @SouthLoneStar, it should be noted that the account was suspended by the date of Charlottesville (12th August 2017). For that reason, I used the back-up account @SouthLoneStar2. Within disinformation campaigns, it appears common for accounts to have back-ups where followers are directed in cases where accounts are suspended (cf. @ELEVEN_GOP that states in the bio 'This is our back-up account in case anything happens to @TEN_GOP'). @SouthLoneStar2 is the back-up account of @SouthLoneStar, as evident in the tweet posted on 13th August: 'hi y'all! @southlonestar is back'. @SouthLoneStar2 also uses the same profile picture and naming strategies as @SouthLoneStar and has previously been identified as a back-up based on synchronicity analysis by Dawson and Innes (2019:253).

The reason why the sample includes all tweets posted during a specific timeframe, rather than just posts containing the word 'Charlottesville' is related to the type of analysis undertaken. Social media narrative analysis acknowledges that narratives on Twitter can be fragmented and distributed across multiple tweets (Page 2012:107). The focus on the overall framing of the narrative of Charlottesville involves attention not only to tweets specific to the events of

Charlottesville but also to other tweets that are posted close to the time of the event and, thus, may remain relevant or contribute to the telling of the story.

The two-week period has been chosen because the event itself remained in the news and public debate for a prolonged period of time, as evident in news reports from September 2017 that continue to document efforts by senators to influence Trump to amend his comments following Charlottesville (Levy 2017).

The specific timeframe was also appropriate due to the activity on the accounts under study. The end of the two-week period coincides with the time when four of the five accounts stopped tweeting: @SouthLoneStar2 stopped tweeting on 22nd August 2017; @USA_Gunslinger on 23rd August 2017; and @TEN_GOP and @Pamela_Moore13 on 24th August 2017. These accounts were all suspended at similar times following their discovery as disinformation items by social media companies (Isaac 2017). While @wokeluisa was the only account that continued tweeting until 21st March 2018, only the tweets until 24th August 2017 (the latest date of activation for the other accounts) has been included in the analysis for reasons of consistency.

The sample of tweets collected includes 685 posts, of which 272 were posted by @TEN_GOP, 184 by @Pamela_Moore13, 45 by @SouthLoneStar2, 53 by @wokeluisa, and 132 by @USA_Gunslinger (see Table 4). For all accounts but @SouthLoneStar2, the tweets from the relevant timeframe were manually exported from the CSRI dataset to a separate Microsoft Excel spreadsheet (Screenshot 6 in Appendix). Data for @SouthLoneStar2 was gathered manually using publicly available datasets (see section 3.3.3).

Table 4: Summary of sample used in analysis of the storying of Charlottesville

Account	Number of tweets
@TEN_GOP	272
@Pamela_Moore13	184
@SouthLoneStar2	44
@wokeluisa	53
@USA_GunSlinger	132
Total	685

3.4.3.2 The method

The analysis in chapter 6 is guided by Georgakopoulou’s (2015) small stories heuristic which is further explained within the chapter. Based on Georgakopoulou’s (2015) model where the analysis of small stories focuses on three separate but interrelated levels (tellers, sites, and ways of telling), my investigation addresses the following areas:

1. Who the 5 main disinformation accounts engaged with when storying Charlottesville.
2. How the accounts used platform affordances to engage with other users when telling the story of Charlottesville.
3. How the disinformation accounts managed (co-)tellership to (dis)identify with different social groups when telling the story of Charlottesville.

To gain an insight into these areas, I carried out three stages of analysis targeted at the levels established by Georgakopoulou (2015).

3.4.3.3 Tellers

To address the first area, I mainly analysed the type of tellership and the type of co-tellers involved in the storying of Charlottesville. In terms of tellership type, tweets were coded as ‘single tellership’, i.e., the account holder is the sole teller and represents the only voice in the tweet, or as ‘multiple tellership’, i.e., there are multiple tellers and voices embedded in the tweet.

In cases where there was multiple tellership, I coded the type of co-teller(s) in terms of the following areas: (i) whether they had a verified⁷ status or not; (ii) the professional domain that they were associated with; and (iii) their political affiliation. The focus on these areas was chosen with the aim of providing a more detailed understanding of how the accounts managed (co-)tellership to index community membership, and how such interactions may have contributed to authentication processes of the accounts’ identity performances.

In order to retrieve such background information about the co-tellers found in the data, I used publicly available information such as self-categorisations (e.g., in Twitter bios), previous research, and shared knowledge (e.g., in the case of news organisations). In cases where co-tellers were not verified, I used resources like the Wayback Machine, particularly for accounts that are no longer active on Twitter. As previously mentioned, due to the sensitivity of such information and the right to be forgotten, such content is included in the qualitative part of the chapter on a selective basis, where potential harm to participants has been minimised. In line with Cardiff University’s Standard Operating Procedure for Social Media Data, information

⁷ At the time of analysis, and during the period that the accounts were active, a ‘verified status’ (used to ‘let people know that an account of public interest is authentic’) was attributed to accounts deemed ‘authentic, notable, and active’ (Twitter 2021a). As defined by Twitter, the notability of an individual’s account was assessed on the basis of high numbers of followers, high numbers of mentions, increased search activity on Google trends, a stable Wikipedia article about the individual, ties with official organisations, and expertise in matters of public interest.

safely stored in existing password-protected spreadsheets will be kept for no longer than five years after the completion of the project.

3.4.3.4 Sites

To address the second area, I mainly analysed and coded the tweets in terms of the tweet format, i.e., original posts, replies, retweets, and quote tweets. As further explained in chapter 6, different tweet formats afford different possibilities for users in terms of narrative roles, power dynamics, and reciprocity. In a quote tweet for example, the person writing the tweet introduces the initial teller's material and becomes a potential problematiser. As a result, power dynamics related to the interaction shift, and the user quoting the tweet gains control of the conversational turn. By quoting a tweet, users engage in one-to-many communication, sharing the message with a wider audience than if they replied to the tweet, and the other user, directly.

In some cases, this information was easily discernible from the data provided by the CSRI; tweets that were retweets often included the letters 'RT' at the start of the message, for example. The data provided by the CSRI also contained some basic information about replies and quote tweets, but this information, when manually checked against archive tools, was not always accurate. For example, some tweets that were quote tweets were not recorded as such. As a result, I looked at external records of every tweet using archive tools such as the Wayback Machine and the 'suspicious tweets' archive of IRA activity (see page <http://suspicioustweets.archivelab.org/?userid=all>). By looking at how the tweets were presented at the time of activity via screenshots, I could see whether they involved any other users and what format they took. As a result, I was able to manually label all tweets in terms of whether they were original posts, retweets, replies, or quote tweets.

3.4.3.5 Ways of telling

To address the third area, I zoomed in further on instances of multiple tellership and analysed how users positioned themselves in relation to potential co-tellers. Guided by Georgakopoulou's (2015) small stories heuristic, I used narrative analysis to investigate how events, characters, and co-tellers were described and assessed. For the analysis of ways of telling, I drew on the concepts of narrative and non-narrative stancetaking. These concepts refer to whether (dis)affiliations are projected using conventionalised storytelling devices, or whether users detach themselves from the story being told to provide more generalised social commentary (Georgakopoulou 2014, 2017). I also analysed whether evaluations were character- or event-focussed, as these are the two main forms that Georgakopoulou (2014:526) suggests that narrative stancetaking can take. She uses character-focussed narrative

stancetaking to refer to evaluations that assess or revoice characters' actions or speech, while event-focussed narrative stancetaking is more closely related to 'what exactly happened' (Georgakopoulou 2014:526). Finally, I suggest how these evaluations relate to wider social group membership and identity performance.

3.5 Conclusion

This study introduces a methodological framework that fills a gap in disinformation research by approaching identity performance from a discourse analytic perspective that acknowledges both metadata and tweet content. By investigating identity performance in terms of both the visual (profile pictures) and the verbal (naming strategies and tweets), this study also offers a multimodal perspective on performances of identity and authenticity online. Despite an emphasis on multimodality within the broad field of social media research (Georgakopoulou and Spilioti 2016:4), relatively little research into disinformation approaches identity performance from a multimodal perspective. To my knowledge, Farkas et al (2017) provide the only detailed insight into the role that multimodal identification features (i.e., naming strategies **and** profile pictures) can play within identity performance in a disinformation context. While it is acknowledged that images can have a stronger impact on audiences than textual cues alone (Powell et al 2015), Hameleers et al (2020:281) argue that 'we know too little about the role of visuals in disinformation'.

Moreover, this thesis demonstrates how internet archives can be used to contextualise data beyond the initial, textual dataset described in section 3.3.2. As discussed in Chapter 2, research on disinformation tends to make use of large datasets, where tweet content is analysed as a means of establishing dominant themes and tweeting practices of disinformation accounts. In many of these cases, research focusses almost exclusively on the textual content of tweets. As a result, the content is viewed as somewhat detached from the context within which it was created as the tweets only constitute part of a Twitter profile. According to Georgakopoulou and Spilioti (2016:3), there has been a 'definitive move towards contextualised approaches that view language and communication in digital environments as both locally situated and as socio-culturally and historically shaped', with such approaches being advocated as both 'a desideratum and prerequisite' within the second wave of digital discourse studies. In this thesis, I use internet archives to gain further insights into the historical context of these accounts. For example, I access the accounts' self-descriptions, profile pictures, names, and interactions with other users, in order to understand identity performances as locally situated within the context of the individual Twitter accounts.

More generally, my approach towards both metadata (i.e., discourse acts shaped by the platform) and data (i.e., discourse acts produced through user-generated content) attempts to acknowledge the symbiotic relationship between platforms and users, as favoured in recent approaches to digital discourse analysis (Vásquez 2022:6). As acknowledged in Chapter 2, despite extensive research into these areas as distinct spaces, relatively little disinformation research has acknowledged how the interaction between platform facilitated opportunities and user-generated content can shed light on identity performance practices. Jones (2020:711,725) argues that an orientation towards post-human theory, recently adopted in applied linguistics, can aid our understanding of how we ‘play the information game’ online, that is, how we ‘negotiate access to various “territories of the self”’. Pennycook (2018:457) defines posthumanism as an umbrella term that ultimately refers to the need to ‘rethink the relations between humans and that deemed non-human’. In other words, posthumanism encourages consideration of how humans engage with technological resources to extend their thinking, rather than distinguishing between ‘internal cognitive processes and external computational ones’ (Mitchell 2003:38). By acknowledging how users utilise platform affordances to construct meaning and perform identity, this thesis orients to the aforementioned post humanist understanding advocated by Jones (2020), according to which technologies and humans are interconnected, rather than distinct, entities.

The need to conduct further data collection to contextualise data is intimately linked with the use of secondary and pre-existing, rather than primary, data in this thesis. As acknowledged in section 3.3.3, although using a secondary dataset, i.e., a dataset that has been collected by other researchers for the purposes of other studies, can speed up the data collection process, the researcher has less control over the dataset and thus often needs to conduct further data collection to fulfil research aims and enrich the data provided. Within the disinformation space, specifically, there are unique challenges relating to the data collection process. While secondary datasets compensate for the challenge that comes with disinformation data often being deleted from the space it was originally posted, there are still challenges when attempting to contextualise interactions. It can be difficult, for example, to understand the meaning of replies when the original tweet has been deleted or the account suspended. As a result, it is not uncommon for potentially valuable data to be removed from studies, because certain tweets appear out-of-context, and thus, are difficult to understand (cf. Xia et al 2019:1651). This thesis attempts to remedy this challenge by conducting further research using internet archives. Having said this, the ever-evolving nature of disinformation data makes this task rather time-

consuming and, at times, futile, as it can be difficult to establish when data is simply no longer accessible.

In line with the tendency for digital discourse analytic studies to be ‘typically situated within qualitative, interpretative paradigms’ (Tovares 2022:25), my application of Georgakopoulou’s (2015) small stories framework enables a detailed qualitative analysis of specific tweets situated in their digital contexts. Despite the advantages of qualitative, interpretive approaches, namely the rich and detailed investigation of specific cases (Hinrichs 2016:30), ‘soft’ qualitative approaches can be criticised for being ‘too subjective’ and not producing generalisable results (Jaworska 2018:330-381). As pointed out by Hinrichs (2016:30), analysts’ subjective perspectives are ‘indispensable to the interpretation’ of qualitative data, and research findings can thus rarely be verified in follow-up studies. In an attempt to redress some of the challenges that can be associated with entirely qualitative approaches, I also employ a quantitative approach to coding, using descriptive statistics to establish dominant patterns within larger samples and enhance replicability, where relevant (see chapters 4 and 5).

Overall, the methodology used in this thesis contributes to the field of digital discourse analysis by demonstrating how online data can be analysed in ways that acknowledge and actively explore how the discourses created by users online both shape and are shaped by the technological discourses of the sites being used. The findings from the various chapters will be interpreted using a dynamic approach to authenticity as a performance process. Each chapter focusses on a separate area of a Twitter profile that contributes to identity performance: naming strategies, profile pictures, and tweet content. The thesis begins with an analysis of the metadata units initially required to have a Twitter page; all users must have a handle and username (Chapter 4). This chapter focusses on how handles and usernames lay the first steps in identity performance by giving users choices about what identity aspects to present about themselves linguistically. I use the aforementioned approach to identity performance, i.e., that strategies of identification are inherently linked with the discourses of the sites of Twitter, to analyse how different naming strategies may contribute to authentication processes.

Chapter 4 Authenticating Identity Performance on Twitter

Disinformation Accounts: Handles and Usernames

4.1 Introduction

This chapter explores how handles and usernames contribute to identity performances as particular personas on Twitter disinformation accounts. Based on Eisenlauer's (2014:78) definition of metadata and his distinction between automated and creative text actions (see section 2.5), this chapter will investigate the use of handles and usernames as creative text actions, that is, they are controlled by humans rather than being generated automatically by the platform. As discussed in section 2.5, previous research suggests that the way that users name themselves online can influence how they are perceived by other users (Morris et al 2012), particularly in relation to authenticity (Jäkälä and Berki 2013). Consequently, this chapter aims to understand what kinds of naming strategies are used by disinformation accounts and how these strategies may contribute to authenticating their performances as specific types of personas (e.g., US citizens).

Firstly, I will outline and exemplify what handles and usernames are and how they are used on the site of Twitter. Then, I will qualitatively analyse how the five main accounts use handles. Based on the categories that arise from the qualitative analysis, I will use descriptive statistics to analyse the type of handles used in the wider sample. This will provide a more comprehensive account of how handles are used across more and comparatively less popular disinformation accounts. I will then qualitatively analyse how the five main accounts use usernames, and again situate the findings in an analysis of the usernames used in the wider sample.

4.2 Defining Twitter profile handles and usernames



Figure 1: Layout of a Twitter profile



Figure 2: A tweet as it appears in a home timeline

Figure 1 and Figure 2 display handles and usernames as they appear on Twitter profiles (Figure 1) and in home timelines (sometimes referred to as news feeds) (Figure 2). This chapter focusses

on handles and usernames because they represent the creative text actions whose identification traces always appear alongside a tweet within a home timeline (as shown in Figure 2). They have also been described as ‘the key piece of your identity on Twitter’ (O’Reilly and Milstein 2012:21).

To illustrate and explain metadata, I will use the account of Facebook Newsroom. Figure 1 depicts how the Twitter profile belonging to Facebook Newsroom appears when the page is visited. In contrast, Figure 2 shows how tweets appear within a user’s home timeline. A home timeline displays tweets posted by accounts that a user follows. If a user was following @fbnewsroom for example, every time @fbnewsroom posted a tweet, it would appear in a list amongst tweets posted by all other accounts a user follows. As shown in Figure 2, a profile’s username and handle are included in the stream alongside a tweet, while other types of metadata (e.g., bio, cover photo) are only visible when visiting a Twitter account page.

The terms “handle” and “username” are often used interchangeably within literature about social media. The way that the terms are distinguished, if at all, differs between and across researchers (see for example Olivier 2014 versus Cavazos-Rehg et al 2014).⁸ In my study, I will differentiate between the two and use them to refer to different text actions. In general, both terms can be used to refer to the names or identifiers social media users must choose when creating a profile. Without these, users cannot participate in the network of Twitter. On Twitter, the platform prompts those who sign up on the network to create two types of names: 1) a unique identifier used to log in to an account and 2) a display name. The unique identifier used to log in to an account always follows the @ symbol and appears in the profile’s unique Twitter URL (see @fbnewsroom in Figure 1- url: <https://twitter.com/fbnewsroom>). In line with previous use in social media research (Cavazos-Rehg et al 2014; De Cock and Pedraza 2018), I will refer to this type of name as a handle. In contrast, I will use the term “username” to refer to the name that appears in bold on an account’s profile page and in the Twitter home timeline (see Figure 1 and Figure 2). These two types of identifiers also differ in terms of their character limit and uniqueness. Handles have a 15-character limit and need to be unique, whereas usernames can extend to 50 characters and do not have to be unique.

⁸ Olivier (2014) employs the term ‘username’ to refer to the name following the @ symbol. In contrast, Cavazos-Rehg et al (2014) call this a handle.

4.3 Identification and authentication through Twitter handles

4.3.1 Eponyms and pseudonyms

As mentioned in section 3.4.1, to establish how identity is conveyed through the handles, the focus of analysis is on what is being represented in each case. More specifically, the analysis focuses on the types of information represented in the handles. I code the meaning of the words used in Twitter handles in terms of their potential for identification with specific categories. The first stage of the analysis establishes whether handles use first names (and are thus ‘identifiable’) or other identification aspects (and are more ‘anonymous’ (Peddinti et al 2014:83)). While the categories of identifiable and anonymous can be useful, Jäkälä and Berki’s (2013) work arguably provides a more detailed insight into how identification can take place online, drawing attention, specifically, to the role of eponyms and pseudonyms.

As acknowledged in section 2.5, Jäkälä and Berki (2013) differentiate between online users that present their ‘real, correct’ names (eponymous identities) or ‘false names or nicknames’ (pseudonymous identities). This work appears to formulate categories of identity by drawing on the distinction between eponyms and pseudonyms as naming strategies. Broadly, eponyms can be understood as ‘proper names’ that are formed after (and can therefore identify) a person (Deseriis 2013:41). In contrast, pseudonyms can be viewed as ‘improper names’ that are constructed to obfuscate identity and are thus less likely to identify a person or people (Deseriis 2013:41). Although Jäkälä and Berki (2013) make reference to naming, their work primarily focusses on types of identity rather than types of names. For example, according to their work, my adoption of the name Kevin Smith would most likely be viewed as pseudonymous as my ‘real, correct’ name is Aurora Goodwin. Having said this, it is important to acknowledge that within the context of Twitter in everyday use, there are limited cues available to users when they make evaluations of identity performances (Morris et al 2012:444). Consequently, users are most likely not able to establish whether online identity performances are eponymous or pseudonymous to the standard discussed in Jäkälä and Berki’s (2013) work. Users can, however, establish whether a name **could** conceivably be ‘real and correct’ (cf. Jäkälä and Berki 2013:8) or ‘proper’ (cf. Deseriis 2013). Drawing on the previous example, while Kevin Smith is not my ‘real, correct’ name, it could more conceivably be used on a person’s birth certificate (or other legal documentation) than an ‘internet name’ like tenacious27 (cf. Morris et al 2012:45). Within the analysis of both the five accounts and the wider sample, all handles are therefore labelled as eponyms or pseudonyms based on whether they use ‘proper’ first names or other identification strategies.

The second stage of the analysis identifies what is being coded within the categories of eponym and pseudonym. In the case of eponyms, all handles use ‘proper’ names. These kinds of names can also index further information about a person’s identity, such as gender and race.

4.3.2 First names indexing gender and race

As indicated above, first names have the potential to reveal information about gender. Liu and Ruths (2013) research gender inference online and find that first names and profile pictures are the strongest indicators of gender. For that reason, I have coded first names based on the gender that they are typically associated with. To source this information, I have searched each name online e.g., ‘Jenna first name’. Typically, search results return information from name meaning sites where gender is usually referenced e.g., ‘the name Jenna is primarily a female name’ (see page <https://babynames.com/name/jenna>). Where names are most often described as ‘female names’ (like Jenna), I have further categorised the name as ‘female’. Where names are most often described as ‘male names’ (like John), I have further categorised these names as ‘male’. Where names are described as ‘unisex’ and thus not typically associated with a binary gender category (like Casey), names have been coded as gender neutral. In line with previous research (Peddinti et al 2014:85), gender has not been coded in cases where there is a lack of clarity such as where I could not access clear gender statistics. This happened in one case within the wider data sample (Yahaya).

In addition to potentially indicating gender identification, previous research suggests that names can also give off information about race (Bertrand and Mullainathan 2004; Fryer and Levitt 2004; Gaddis 2017:471; Sweeney 2013). According to Sweeney (2013:7), Black-identifying and white-identifying names are those that occur with sufficiently higher frequency in one race than another. Previous research suggests that up until the early 1970s, there was little distinct difference between names given to Black and white children (Fryer and Levitt 2004:767). However, following the rise of the Black Power movement in the 1970s, there was a shift in how Black people perceived their identities and, as a result, distinctly Black names began to emerge (Fryer and Levitt 2004:770). While names are of course imperfect indicators of race, i.e., not all people with a particular name will be of the same race, research suggests that first names do tend to be used to make racial assumptions. Surveys, for example, reveal that respondents tend to identify ‘Black-sounding’ and ‘white-sounding’ names from lists (Bertrand and Mullainathan 2004; Gaddis 2017), while research into search engines suggests a structural racial bias for Black-sounding and white-sounding names (Sweeney 2013). Research

into job applications also suggests that names can be used as the basis of racial discrimination (Bertrand and Mullainathan 2004).

When investigating associations between race and names, previous research has made use of birth records and surveys to establish whether names are predominantly given to children who are Black or white, and to establish whether the general public view certain names as being associated with racial identity (Bertrand and Mullainathan 2004; Gaddis 2017; Sweeney 2013). These studies have resulted in lists of names that can be understood as distinctively “Black” or “white” (Bertrand and Mullainathan 2004; Gaddis 2017; Sweeney 2013). Across these lists, certain patterns appear distinct within categorisations relating to race and gender: white female names appear to repeatedly contain a higher concentration of names ending with /i:/ e.g., Carrie, Emily, Amy; Black female names are more likely to end in ‘isha’ e.g., Tanisha, Denisha, Lakisha; white male names tend to have a smaller number of syllables, e.g., Luke, Jake, Jay; Black male names are the only ones that start with the ‘De/Da’ morpheme e.g., DeAndre, DeShawn, DaQuan (Bertrand and Mullainathan 2003; Gaddis 2017; Sweeney 2013). These examples suggest that certain types of names can be typically associated with particular races and genders.

In addition to the use of lists created from birth frequency records, Sweeney (2013) proposes the use of Google Image search as a tool to detect racial distinction. She shows, for example, that search results for the names Latanya and Latisha show predominantly Black faces, while the names Jill and Kristen result in images showing predominantly white faces.

Based on the aforementioned research that suggests names can index race, I have used previously established racial categories relating to names (e.g., the lists published by Bertrand and Mullainathan 2004; Gaddis 2017; Sweeney 2013) and Google Images (cf. Sweeney 2013) to establish whether names have a racial distinction. If a name has been categorised as ‘Black’ or ‘white’ in previous naming research, or if Google Images displays images of predominantly Black/white faces, I have coded the name as indexing race, and further categorised whether the racial distinction is Black or white. Where it was not possible to establish racial distinction, e.g., because the names were not included in previous research, or because Google Image results did not return images of a) people or b) more than a 50% predominance of Black or white faces, race has not been coded.

It is important to acknowledge that there are numerous racial identities outside of the binary categories of Black and white. Having said this, previous research into IRA accounts suggests that left-leaning Black users were targeted separately from other demographics during the 2016

election disinformation campaign (Freelon et al 2022:572). Moreover, Freelon et al (2022:563) argue that IRA accounts can be categorised as Black and non-Black, with a racial asymmetry between Black- and white-presenting accounts existing alongside political asymmetries between left and right. Consequently, this research will orient to the categorisations of Black and white established by Freelon et al (2022) and used within previous research into names (Bertrand and Mullainathan 2004; Gaddis 2017; Sweeney 2013). An example of the coding strategy is shown below:

Handle: @AmandaVGreen

Username: Amy Green

Name type: Eponym

Identity aspects: Name, gender (female), race (white).

4.3.3 Pseudonyms indexing various identity aspects

Although pseudonyms have been described as names that present ‘false’ characteristics intended to ‘obfuscate’ a person’s identity (Deseriis 2013:41; Jäkälä and Berki 2013:8), pseudonymous naming strategies (i.e., those that do not use a first name) can still reveal information that identifies an individual in relation to social (i.e., ‘one’s awareness of being a member of a social category or group’), rather than ‘personal’ (i.e., one’s individual characteristics such as name/psychological characteristics) identity aspects (Amaral and Monteiro 2002:576;579). Where handles do not use first names, I have coded each handle based on the kind of ‘other’ identification they use. In cases where the identification meaning of the handles is ambiguous, I have drawn on usernames that assist in meaning disambiguation. Taking into account content encoded in both handles and usernames is supported by previous studies of Twitter handles. In his study of South African nicknames in Twitter handles⁹, Olivier (2014) uses the relationship displayed between username and handle to create categorisations for handles such as ‘individual name and surname swop’ for ‘Aubrey Matshiqi @MatshiqiAubrey’, and ‘shortened username’ for ‘GCIS Media Liaison @GCIS_Media’. Where handles present more than one identity aspect, all identity aspects are labelled. An example of the coding strategy is shown below:

⁹ Olivier (2014) employs the term ‘username’ to refer to what I call a ‘handle’; in contrast, what I call ‘username’ is referred to as ‘real names’ or ‘proper names’ in Olivier’s (2014) study. The fact that usernames do not always take the form of names in my sample justifies the adoption of the more generic term ‘username’ in my sample.

Handle: @rightnpr

Username: Right and Proud

Name type: Pseudonym

Identity aspect: Political affiliation

4.3.4 Identification through Twitter handles: A qualitative analysis of the five main accounts

(i)@TEN_GOP

@TEN_GOP had 147,767 followers and represents the account with the most followers in the data set. As indicated in Table 3, it is the only account in the '>100,000 follower' category. Created in November 2015 and tweeting from December 2015, the account never changed its handle which consists of two parts separated by a punctuation mark (the underscore). The first part, TEN, is rather ambiguous as it could refer to a number, which in turn could represent a variety of identity aspects, including age, birthday, lucky number, etc. However, attention to the username allows the disambiguation of its meaning and the interpretation of it as a shortening for Tennessee. The second part of the handle, 'GOP', can be interpreted as an abbreviation of 'Grand Old Party' that has been used to refer to the Republican party in American politics since the 1800s (Engs et al 2002:ix). In this instance, the handle @TEN_GOP identifies the account in terms of geographical location and political affiliation.

More specifically, Tennessee is a part of the Southern 'Bible belt' and has been a Republican stronghold since 1996. The state is characterised by its traditionalist reputation, with 'strong societal support for traditional values and institutions' and 'acceptance of traditional inequalities' (Lyons et al 2001:17). Historically conservative, the state still upholds the death penalty and is ranked the '13th best state for gun owners' (Wood 2019). The coupling of both stereotypically conservative location with explicit declaration of political affiliation consequently reinforces the identification of the account as right-leaning. The correspondence between reported location, political orientation, and historical socio-political context suggests that the account fits stereotypical expectations associated with the persona of a Southern Republican. Based on previous research, according to which an adoption of normative practices can lead to constructions of authentic collective identity (Coupland 2003; Cutler 2003; Herring 2004a; Maragh 2018), it could be argued that the adherence to stereotypical expectations in

@TEN_GOP's handle may contribute to the valid representation of a right-wing persona from the South.

It is important to acknowledge that location and political affiliation foreground collective identity aspects; political parties are defined by the fact that they are groups of people sharing ideas and beliefs, while location can provide the shared foundation for local, national, and international identities. Through their handle then, @TEN_GOP uses features that can be considered less personal to contribute to their identity performance; the handle represents information that could not be used to characterise the owner of the account in anything other than collective terms.

(ii) @Pamela_Moore13

@Pamela_Moore13 had 72,121 followers. The account was created in November 2015 and started tweeting from February 2016. The handle did not change during the account's activation. This account uses a first name and surname structure, with the two elements being divided by an underscore. This interpretation is supported by the username, which will be discussed in section 4.4.1. The addition of a number (13) at the end of a name is a practice used by social media users in order to create a unique handle, as previously discussed in relation to the function of the handle as a unique URL. Considering that both Pamela and Moore are popular names (Social Security Administration 2020; US Census Bureau 2010), both number and underscore are likely to have been used as means of differentiation from other individuals bearing the same first and last name. While, as previously mentioned, numbers can be used to represent personal information, without further contextual cues, one 'cannot interpret the aforementioned use of numbers any further (Olivier 2014: 66). The name Pamela does not appear in previous research as a name with a racial distinction. When searched on Google Images, the name Pamela predominantly shows images of white women, suggesting there is a racial distinction associated with the name (Sweeney 2013).

In terms of other identity cues given off through the handle, the specific form suggests that the account owner is most likely a) an individual person, b) called Pamela Moore, and c) a woman. As will be discussed in chapter 5, the profile picture used by @Pamela_Moore13 further corroborates this interpretation and supports Liu and Ruths (2013) claim that the two work together in indicating gender. Compared to @TEN_GOP and its focus on location and political affiliation, @Pamela_Moore13's use of a 'proper' first name identifies the account in more personal, rather than collective, terms (Deseriis 2013:41; Olivier 2014:53).

(iii) @wokeluisa

@Wokeluisa had 57,295 followers. The account was set up in March 2017 and started tweeting from that time. The handle stayed the same throughout the activation period. Like @Pamela_Moore13, the correspondence between @wokeluisa's handle and username (Luisa Haynes) suggests that the handle includes a first name. Like the case above, the first name can be taken as an indication of gender, which is further supported by the account's profile picture (see section 5.4). The handle can be split into two parts despite the lack of punctuation separating them: 'woke' + 'luisa'. The first part of the handle 'woke' is a slang term, with origins in African American Vernacular English. With strong associations with the Black Lives Matter movement, the term was 'popularised by the cultural force known as "Black Twitter"' (Yao 2018:439). The term is used to refer to the concept of being culturally enlightened, i.e., aware of injustices and inequalities stemming from institutional discrimination (Minamore 2017). The term is often used when referring to the political left and those with liberal viewpoints. Despite being used as a positive evaluative descriptor within left-leaning communities, the term has become a slur amongst the political right, used to 'mock the hypersensitivity of the left' (Ramaswamy 2019). The name Luisa does not appear to have been identified as having a racial distinction in previous research. When searched on Google Images, the name does not return results that can be used to interpret racial dominance within the categories of Black and white¹⁰.

The selection of the particular slang term reveals an orientation towards indexing a personal attribute or characteristic; the handle signals a presentation of the account holder as being, or presenting themselves as, 'woke'. While based on the term's definition, one could suggest that the term demonstrates political affiliation, the usage of the term as a qualifier of the individual name, 'Luisa', indicates that it is mainly being used as character description. In comparison to the explicit political affiliation demonstrated in @TEN_GOP, the use of 'woke' as presentation of political affiliation appears somewhat ambiguous in that there is no explicit mention of political party, and the handle could be being used ironically as is widely done by the political right (Ramaswamy 2019). As with @Pamela_Moore13, @wokeluisa's use of a 'proper' name (Luisa) draws on identity aspects that can be considered primarily personal rather than collective. Having said this, the lack of a surname in @wokeluisa's handle presents information that can be considered somewhat less individually identifying than @Pamela_Moore13.

(iv) @SouthLoneStar

¹⁰ The search predominantly returns images of a character from animated film, Encanto.

@SouthLoneStar had 53,999 followers. The account was created and began tweeting from November 2015. The pseudonymous handle can be split into three units, indicated by the use of capital letters to separate words: South + Lone + Star. ‘Lone Star’ is a collocation often associated with the US state of Texas. Stemming from its former status as an independent republic, Texas earned the nickname “the Lone Star state”. The reference to ‘South’ in the first part of the handle arguably supports the interpretation of ‘Lone Star’ as a reference to geographical location, more specifically, Texas. The interpretation of the handle as indexing geographical location is further disambiguated by the username and profile picture, which will both be discussed in later sections.

In terms of account identification, @SouthLoneStar and @TEN_GOP employed similar strategies in their handles by orienting to geographical location. Neither handle presents information pertaining to an individual that could be traced to a person offline using traditional identification strategies such as ‘proper’ names.

(v) @USA_GunSlinger

@USA_GunSlinger had 44,895 followers. The account was created in November 2015 and started tweeting from December 2015. The handle is split into two units by the underscore as a separator. The first unit, USA, indicates national location: United States of America. As for the second unit, the term ‘gunslinger’ is popularly used to refer to the old Western practices of gunfighting and expert fire (Cambridge Dictionary 2020). Consequently, the term in the handle represents a declaration of topic of interest or hobby. The handle therefore draws on similar identification strategies as @SouthLoneStar and @TEN_GOP; there is an absence of traditional naming strategies in the form of first and/or last names, and identification is instead centred on collective identity aspects.

Based on the analysis of the handles in the five accounts, I have found the following labels relevant to the type of identification achieved through the handles: ‘geographical location’, ‘political affiliation’, ‘proper name’, ‘gender’, ‘race’, ‘personal attribute’, and ‘hobbies/interests’. The results of the initial analysis can be presented as follows:

Table 5: Identity aspects signalled in five account handles

Handle	Proper Name	Geographical Location	Political affiliation	Personal attribute	Hobbies/Interests	Gender	Race
@TEN_GOP		✓	✓				
@wokeluisa	✓			✓		✓	
@Pamela_Moore13	✓					✓	✓

@SouthLoneStar		✓					
@USA_GunSlinger		✓			✓		

These labels demonstrate that online naming strategies can be composed of different types of identity aspects. Based on the degree of personal information revealed, @Pamela_Moore13 and @wokeluisa can be considered more personal, while @USA_GunSlinger, @SouthLoneStar and @TEN_GOP are seemingly less personal, drawing on collective identity aspects related to social categorisation (cf. Amaral and Monteiro 2002). The following analysis will draw on and enrich the labels presented above.

4.3.5 Identification through handles: Using descriptive statistics to analyse the wider sample

Following the categories/labels that emerged by the qualitative analysis of the handles of the five main accounts (section 4.3.4), I categorised all accounts in both data samples (more popular and less popular accounts) based on the type of identification conveyed in their handles. Closer investigation of the wider sample presented the need to introduce new labels to the categorisation: news, animals, organisation, and religion. News was used to code accounts that contained words traditionally associated with news organisations such as ‘Today’, ‘Post’ or ‘Daily’. This category has previously been established in studies of disinformation accounts; Linvill and Warren (2018) used ‘news feed’ to refer to handles that ‘overwhelmingly presented themselves as U.S. local news aggregators and had descriptive names such as @OnlineMemphis and @TodayPittsburgh.’ Animal was used in cases where the handles referred to animals e.g., ‘@screamymonkey’, organisation was used where accounts presented themselves as being organised groups e.g., @SpaceDept, and finally, religion was used when handles used language that invoked associations with religion e.g., @Jihadist2ndWife. Definitions and examples of each coding category are displayed in Table 25 (see Appendix).

Before moving to the types of identification displayed by the remaining (81) accounts, the data was initially analysed in terms of naming strategy. As demonstrated in Table 6, the accounts in the 10,000-99,999 group (popular), versus the accounts in the 5,000-9,999 group (less popular) use different types of naming strategies to construct identity. Handles in the more popular sample use predominantly pseudonymous naming strategies, while in the less popular sample, pseudonymous and eponymous naming strategies are used more equally. This suggests that the more popular accounts use handles that revealed less personally identifying identity aspects than those in the less popular sample. Across both samples, the majority of eponymous handles use first names alongside surnames (seven in both samples) e.g., @DickyIrwin, while

the others (three in both samples) use first names only e.g., @JemiSHaaaZzz, or first names combined with other identity aspects e.g., @AndyHashtagger. There are no cases where a surname is presented without a first name (e.g., @MissGoodwin).

Table 6: Naming strategies used in handles across both samples

Naming strategy	Handles of more popular accounts (n)	Handles of less popular accounts (n)	Total
Eponyms	10	10	20
Pseudonyms	48	13	61
Total	58	23	81

The distribution of identity aspects presented within the handles is shown in Table 7 and Table 8 below. The percentages in these tables demonstrate what proportion of each sample size display a particular identity aspect. Due to numerous accounts displaying more than one identity aspect (as demonstrated in section 4.3.4), the percentages do not add up to 100%.

Table 7: Identity aspects presented in handles for more popular accounts

Identity aspect in handle	Number of instances (n: 101)	Proportion of sample representing each identity aspect (%)
Geographical location	32	55%
News	30	52%
Proper Name	10	17%
Gender	11	19%
Hobbies/interests	11	19%
Race	9	16%
Political affiliation	4	7%
Personal attribute	3	5%
Animals	2	3%
Organisation	0	0%
Religion	0	0%

As evident in Table 7, geographical location is the identity aspect most commonly presented in the handles of more popular accounts. The table also shows that a similar number of handles identify the relevant accounts as news organisations (see ‘news’). A closer investigation reveals that this is because location and news often co-occur within the same handles: out of the 32 accounts in the more popular sample indexing location in their handles, 27 simultaneously use news-related language e.g., @KansasDailyNews.

The prominence of location as an aspect of identity performed in handles appears to account for the lower number of eponymous naming strategies used by more popular accounts; of all the accounts presenting location, none of them use location alongside a name. In contrast, all eponymous identifications use ‘proper name’ and index gender, while six of the names show a racial distinction: four with a white racial distinction, and two with a Black racial distinction. Race, specifically Black identity, is also drawn on as a collective identity aspect in three accounts that use pseudonyms: @BlackToLive, @Blk_Voice and @blackmattersus. There are no cases where white racial identity is constructed collectively. Consequently, it appears that accounts that present themselves pseudonymously and eponymously tend to do so by drawing on different kinds of identity aspects: eponymous presentations use proper names that index gender (and sometimes race), while pseudonyms tend to be centred on collective identity aspects like location.

Table 8: Identity aspects presented in handles for less popular accounts

Identity aspect in handle	Number of instances (n: 38)	Proportion of sample representing each identity aspect (%)
Geographical location	4	17%
News	4	17%
Proper Name	10	43%
Gender	11	48%
Hobbies/interests	3	13%
Race	7	30%
Political affiliation	5	22%
Personal attribute	0	0%
Animals	0	0%
Organisation	1	4%
Religion	1	4%

Table 8 shows that in the less popular sample, accounts draw on location and news less often than in the more popular sample. Having said this, where location and news are presented, they do tend to co-occur: three of the four accounts indexing location also use news-related language. It should be noted that this type of handle (location + news) is not documented among the five main accounts analysed (section 4.3.4). In fact, only one of the top ten most followed accounts in the entire data set uses news-related language in the handle. The prevalence of location-identified, news-branded accounts within the wider sample of popular accounts thus demonstrates the importance of looking beyond accounts with the most followers. Although

such accounts may not have the highest number of followers, location-identified, news-branded accounts appear to play a key role within the kinds of identity performances used to push disinformation by the IRA.

In contrast to the more popular accounts, in the less popular sample, handles that present names are the most common. Eponymous identifications index gender in all cases (five female, four male, one gender neutral) and race in seven cases (all of which have a white racial distinction). In contrast, pseudonymous handles draw on a wider range of collective identity aspects more equally (location, political affiliation, hobbies/interests) than in the more popular sample. In the less popular sample, race is not indexed through any pseudonyms.

4.4 Identification and authentication through Twitter usernames

Similar to the analysis of handles, I have also coded Twitter usernames in terms of their potential for identification with specific categories. The decision to follow a similar coding system is supported by previous research, according to which, handles and usernames can be investigated in parallel in order to understand how the information presented in each unit may be related (Olivier 2014). Applying the same categorisation to usernames will lead to a more comprehensive understanding of the use of naming strategies in disinformation accounts, highlighting differences and/or similarities in how naming is used as an identification strategy. Similar to section 4.3, I will first analyse the usernames utilised by the five main accounts. Then, I will use the codes generated through the initial qualitative analysis to carry out an analysis of the wider sample of more and comparatively less popular disinformation accounts.

4.4.1 Identification through usernames: A qualitative analysis of the five main accounts

(i)@TEN_GOP: Tennessee GOP

The username for this account changed from ‘Tennessee GOP’ to ‘Tennessee’ in July 2017. The reference to the American state in the username supports the interpretation of ‘ten’ in the account handle as an abbreviation of Tennessee and thus, as revealing locational information. The correspondence between the handle and username also reflects the tendency for handles to take the form of shortened usernames (Olivier 2014:58). The deletion of ‘GOP’ from the username in 2017 changes the username from displaying location and political affiliation to just location.

Similar to the handle @TEN_GOP, the account username presents identity aspects that draw on social group categorisation rather than more personal identity aspects (like first name, for example). The idea that naming strategies can present different degrees of identification is echoed in Olivier's (2014) study, where naming strategies that I describe as 'more personal' are categorised as 'individuals' and are distinguished from collective identity performances that in Olivier's (2014) study took the form of 'types of business and organisation username'. In contrast to this study, Olivier (2014) recorded no instances where non-business or non-organisation names drew on purely collective identity aspects (such as Tennessee @TEN_GOP).

(ii)@Pamela_Moore13: Pamela Moore

The username for this account, 'Pamela Moore', remained the same throughout the period of activation. The structure of the username, with a space between the two elements, follows a traditional first name + surname structure used in offline contexts. The correspondence between the handle and username supports the previous discussion of @Pamela_Moore13 as using an eponymous naming strategy and arguably follows a common model of online naming practice, as suggested by Olivier (2014). Olivier (2014:65) suggests that individuals often use handles that closely correspond to their usernames, (which he labelled their 'real names') with the addition of numbers and special characters for purposes of establishing uniqueness. This account therefore follows previously established naming practices associated with accounts that construct identity performances as individuals. The username does not present any additional or different information to the coding established in the handle: name, gender (female), race (white).

(iii)@wokeluisa: Luisa Haynes

The username for this account, 'Luisa Haynes', remained the same throughout the period of activation. The username follows a more traditional eponymous structure than the handle. While the handle contains information pertaining to a personal attribute and first name, the username instead follows the structure of a first name followed by a surname. The different kinds of identity aspects drawn on in the two metadata units in this case (handle: personal attribute + first name, username: first name + surname) appear to work in conjunction with one another to present different aspects of identification for the account holder. The different types of identification may serve different purposes in authenticating and projecting particular identities; full names can increase trustworthiness and authentication due to their link with the 'real world' (Jäkälä and Berki 2013:6), while the use of personal attributes in handles can give

information pertaining to how the user views themselves. This kind of identity presentation does not appear to be uncommon. According to Olivier (2014:62), some Twitter users present ‘nicknames’ that reveal personal traits in their handles while presenting their ‘real names’ in their usernames. In many of these cases, Olivier (2014:62) suggests that there is a degree of correspondence between the nicknames used in handles and the ‘real names’ used in usernames. Consequently, it seems that @wokeluisa follows a naming practice that is not rare on Twitter. The account presents different forms of identification in the handle and username that correspond with one another to ultimately reinforce particular identification strategies (e.g., the name ‘Luisa’).

(iv) @SouthLoneStar: South Lone Star

The username for this account was South Lone Star until it changed to Texas Lone Star at some point between January and March 2017. As discussed in section 4.3.4, Texas has the nickname the ‘lone star state’. The replacement of ‘South’ with ‘Texas’ in the username in 2017 arguably disambiguates the handle for any users who may not be familiar with the historical context of the state; the use of ‘Texas’ clarifies that the user identifies as Texan and not just Southern. Similar to @wokeluisa, @SouthLoneStar appears to present slightly different information in the username and handle (from early 2017 onwards). Although both metadata units (handle and username) draw on location, the username is more explicit and specific than the handle. In line with previous research, an emphasis upon location can be viewed as a performance of collective identity (Amaral and Monteiro 2002). @SouthLoneStar therefore uses a pseudonymous approach to naming, presenting less personal information than what would be signalled by the use of first and/or surnames.

(v) @USA_GunSlinger: American GunSlinger

The username for this account was American GunSlinger until it was changed to GunSlinger Girl at some point between January and June 2017. With this change, the username presents additional information to the handle: the account becomes gendered. ‘GunSlinger’ remains the same in both the handle and username indicating that it is likely the most salient aspect in terms of their identity performance. The combination of both @USA_GunSlinger and GunSlinger Girl consequently identifies the account in terms of location, hobby, and gender. The use of hobby/interest-based usernames can be linked back to the history of internet chatrooms and forums as places for offline communities to manifest and discuss certain activities (Williams 2006).

The change of @USA_GunSlinger’s username to GunSlinger Girl makes it the third account in the data set to perform a female identity. Unlike @wokeluisa and @Pamela_Moore13 where the gender is inferred from the inclusion of first names, @USA_GunSlinger explicitly states gender through the inclusion of the word ‘Girl’. Rather than a personal naming strategy (e.g., eponym) that evokes a link between the online and the offline, physical world (cf. Jäkälä and Berki 2013), the identification is achieved through reference to a social categorisation (Amaral and Monteiro 2002:579).

More specifically, the use of ‘girl’, rather than other gendered terms (e.g., woman), invokes connotations with youth. Considering that the term ‘girls’ may widely be understood as referring to ‘female children’ (Stokoe 1998:222), research suggests that it is common for the term ‘girl’ to be used to in a way that infantilises women, invoking connotations with youth and childhood (Bolinger 1980:116) and, to some extent, ‘fun’. @USA_GunSlinger therefore presents more information about the individual via their username by emphasising collective identity membership (female gender) and mobilising language that can be associated with specific gender stereotypes.

Table 9 provides an overview of the key categories that emerged from the qualitative analysis of the usernames of the five main accounts. These categories will serve as the basis for the coding and use of descriptive statistics when analysing the usernames employed in the wider sample.

Table 9: Identity aspects signalled in five account usernames

Handle	Proper Name	Geographical Location	Political affiliation	Hobbies/Interests	Gender	Race
Tennessee GOP		✓	✓			
Tennessee		✓				
Luisa Haynes	✓				✓	
Pamela Moore	✓				✓	✓
South Lone Star		✓				
Texas Lone Star		✓				
American GunSlinger		✓		✓		
GunSlinger Girl				✓	✓	

4.4.2 Identification through usernames: Using descriptive statistics to analyse the wider sample

Based on the categories identified in the qualitative analysis, I coded the usernames of all remaining accounts. In this section, Table 12 and Table 13 include the results of the analysis of usernames, together with the results of the analysis of handles (see section 4.3.5), in order to make identifiable any patterns in the naming strategies employed across both metadata units.

Before moving to the types of identification achieved through the usernames, Table 10 and Table 11 present the extent to which the usernames of the accounts in the wider sample take the form of eponyms or pseudonyms. Table 10 displays the distribution of naming strategies used in the more popular accounts, while Table 11 displays the frequency of the naming strategies used in the comparatively less popular accounts.

Table 10: Naming strategies used in more popular accounts

Naming strategy	Handle (n)	Handle (%)	Username (n)	Username (%)
Eponyms	10	17%	12	21%
Pseudonyms	48	83%	46	79%
Total	58	100%	58	100%

Table 11: Naming strategies used in less popular accounts

Naming strategy	Handle (n)	Handle (%)	Username (n)	Username (%)
Eponyms	10	43%	9	39%
Pseudonyms	13	57%	14	61%
Total	23	100%	23	100%

The results of the username analysis reflect similar patterns to those established in the analysis of handles. In the sample of more popular accounts, pseudonyms are the most common naming strategy used in both usernames and handles. On the other hand, in the sample of less popular accounts, a more even spread across eponyms and pseudonyms is documented in both usernames and handles. Across both samples, there is a slight increase in the number of eponymous naming strategies that use both first and surname in their usernames (11 accounts in the more popular sample and nine accounts in the less popular sample). In each sample, there is one username that uses first name without surname, and no usernames that draw on surname only. The analysis of the identity aspects invoked in usernames also reveals that the identification patterns across the two metadata units are rather similar.

As previously mentioned, the sample sizes for the two groups are different: there are more accounts in the more popular group than the less popular group. Table 12 and Table 13 demonstrate the proportion of handles and usernames representing each identity aspect.

Table 12: Aspects of identity presented in more popular accounts

Identity aspect	Handles		Usernames	
	Number of instances (n:101)	Proportion of sample representing each identity aspect (%)	Number of instances (n:101)	Proportion of sample representing each identity aspect (%)
Geographical location	32	55%	30	52%
News	30	52%	29	50%
Proper Name	10	17%	12	21%
Gender	11	19%	11	19%
Hobbies/interests	11	19%	10	17%
Race	9	16%	10	19%
Political affiliation	4	7%	3	5%
Personal attribute	3	5%	3	5%
Animals	2	3%	2	3%
Organisation	0	0%	0	0%
Religion	0	0%	0	0%

Table 13: Aspects of identity presented in less popular accounts

Identity aspect	Handles		Usernames	
	Number of instances (n:38)	Proportion of sample representing each identity aspect (%)	Number of instances (n:38)	Proportion of sample representing each identity aspect (%)
Geographical location	4	17%	4	17%
News	4	17%	3	13%
Proper Name	10	43%	9	39%
Gender	11	48%	12	52%
Hobbies/interests	3	13%	1	4%
Race	8	35%	5	22%
Political affiliation	5	22%	6	26%

Personal attribute	0	0%	0	0%
Animals	0	0%	0	0%
Organisation	1	4%	2	9%
Religion	1	4%	1	4%

As evident in Table 12, the more popular accounts primarily reveal an identification with geographical location and news in their selection of usernames and handles. On the other hand, as shown in Table 13, the less popular accounts primarily mobilise proper names and gender aspects in both usernames and handles. What is notable across both samples is the consistency of identity aspects presented across both metadata units. For example, the handles and usernames tend to use the same naming strategies and present the same identity aspects e.g., handle: @KansasDailyNews, username: KansasCityDailyNews. This pattern corroborates Olivier’s (2014) suggestion that handles and usernames often correspond either via shortenings or exact replications. Across both samples, there are no instances where conflicting information is presented in handles and usernames (such as an American state in the handle but a British city in the username).

Similarly, in most cases across both samples, the same personal name, gender orientation, and racial distinction are presented in both handles and usernames. In some cases, race is indexed in one unit but not the other because an eponym and pseudonym are used alongside one another e.g., @TheFoundingSon John Davis. Within the less popular sample, there are no handles nor usernames that show a Black racial distinction; all names that have a racial distinction can be associated with white people.

Within the more popular sample, there is a more even distribution of racial identities: five usernames index white identity via eponyms, while five usernames index Black identity (three eponyms, two pseudonyms). Where Black identity is signalled via pseudonym, these presentations draw on race as a shared identity aspect e.g., @blackmattersus Black Matters. The presentation of Black identity as a shared identity aspect is not uncommon on Twitter, a platform that has been used as a means of contesting erasure in the Black community, organising protests against state violence, and playing a key role in the evolution of the Black Lives Matter movement (Hill 2018). In comparison, whiteness is ‘a marker of dominance against which other groups gain identity, and as such has no identity beyond its ability to dominate’ (Gallaher 2021:249). As a result, performances of whiteness as a shared identity aspect are indicative of extremist alt-right ideology rather than large-scale anti-racist activism (Gallaher 2021). Such accounts consequently do not play the same kind of role on Twitter as accounts that foreground Blackness.

Overall, based on the kind of identifying information presented in both handles and usernames, both more and less popular accounts can be broadly split into two main categories: accounts associated with individuals (personal accounts) and those associated with news (news accounts). While personal accounts that use eponyms appear to behave in the same way across both more and less popular accounts by predominantly using first name + surname structure, there is some variation noted for news accounts. In the more popular sample, news accounts draw on location and news-related terms in both handle and username in very similar ways (cf. @KansasDailyNews). In the less popular sample, however, there is more variation in the way that news accounts are identified and presented. News accounts in the less popular sample do not always present the same information across handles and usernames e.g., @MissouriNewsUS has the username ‘People for Justice!’. In another case, a news account is identified as related to a location **and** a particular political affiliation e.g., @redlanews Red Louisiana News. The presentation of location-based news accounts as affiliated with political positions does not appear in the more popular sample. A more detailed exploration of the differences between the more and less popular accounts and implications in relation to authenticity will be included in section 4.5.

When compared against the five accounts focussed on throughout the thesis, there are some similarities between the ways that the main accounts and other more popular accounts use naming strategies. Although the five accounts do not use news branding, they use pseudonyms more than eponyms in both handles and usernames, and location is drawn on as a social identifier, i.e., in relation to social group categorisation.

Considering that gender stereotypes can be exploited to fulfil interactional goals online (Herring and Stoerger 2014; Hussain and Griffiths 2008; Jenson et al 2015; Yee et al 2011), I investigate gender identification and distribution in this data set by drawing on information gathered from both handles and usernames. Although first names remain the main practice for conveying gender in the data, there is also one instance in the more popular sample where gender is presented via pseudonym (@TheFoundingSon) and three instances in the less popular sample (Marlboro Man, Resistance Girl, and Jihadist Wife). Gender identification is therefore primarily achieved in accordance with more personally identifying information (first name). Pseudonymous naming strategies can use gendered terms to reveal more information about an individual, but such handles and usernames do not demonstrate the same level of personal identification as eponymous strategies.

In terms of gender distribution, there is some variation between the more and less popular disinformation accounts. Identifications with the female gender are more frequent in accounts

with more followers (73%), compared to those with less (50%). The sample of the five main accounts appears to reflect the pattern found in the more popular accounts: all usernames and handles that encode gender (@Pamela_Moore13 Pamela Moore, @wokeluisa Luisa Haynes, and @USA_GunSlinger GunSlinger Girl) identify the account holder as female.

The prevalence of female identity performances amongst the more popular sample could be explained via the social capital that women are reported to have on social media: women are more likely to have wider social networks and higher numbers of ‘friends’ and are more likely to be ‘influencers’ (Djafavora and Rushworth 2017:5; Magdy et al 2017). Within this dataset, some pre-existing patterns related to gender and engagement thus appear to be reflected in the more popular accounts; accounts with more followers are more likely to be performing female identities. It has been suggested that disinformation narratives do not need to be internalised to the point of complete trust to have an unconscious impact on decision making (Bastick 2021; Canan and Warren 2018). Consequently, despite associations between masculinity and increased credibility (Armstrong and McAdams 2009), the relatively fewer performances of men within the more popular sample may suggest that female identities are favoured within disinformation presentations due to their association with increased attention, engagement, and thus, exposure.

4.5 Discussion

This chapter aimed to explore what kinds of naming strategies are used by disinformation accounts and how these strategies may contribute to authenticating their identity performances as American personas. The analysis of handles and usernames suggests that this sample of Twitter disinformation accounts not only constructed identity performances as American citizens, but also as news accounts. This section will discuss how the accounts employed a range of naming strategies to authenticate their performances as different types of Twitter accounts.

Previous research suggests that depending on the context of an online interaction, different kinds of naming strategies can be used to contribute to authentication (see section 2.5). In a disinformation and false rumour spreading context, research into metadata and naming strategies has focussed primarily on identity performance practices of news accounts. Studies have focussed on the sharing of ‘fake news’, with source perceptions being situated within the context of Twitter as a news source (Andrews et al 2016; Morris et al 2012). The analysis in section 4.4.2, however, reveals that news accounts are only one type of account used to disseminate disinformation. The analysis shows that accounts use both eponymous and

pseudonymous naming strategies to spread disinformation via two main types of accounts: news accounts and personal accounts. These types are created by the adoption of identification practices associated with different communicative contexts: the use of Twitter by news organisations and the use of Twitter by ordinary citizens. The findings consequently suggest that rather than one kind of naming strategy (e.g., pseudonyms or eponyms) being used to authenticate identity performances on disinformation accounts, authentication is constructed via different naming and identification strategies depending on the type of account being used to propagate disinformation.

While there are two types of accounts in the disinformation sample serving different purposes and orienting to different naming strategies, consistency of identification information across both handles and usernames appears to be important regardless of the type of account. Section 4.4.2 shows that both types of accounts in both samples tend to present the same identity aspects in their handles and usernames. Such consistency is also noted in studies that explore naming practices beyond the disinformation context. In Olivier's (2014) study of South African Twitter handles and usernames, the majority presented corresponding information in both naming units such as in the form of shortened usernames, first name and surname swaps, or the addition of numbers to pre-existing name such as GCIS Media Liaison @GCIS_Media, Aubrey Matshiqi @MatshiqiAubrey, Graeme Smith @GraemeSmith49. As acknowledged in chapter 2, previous research suggests that performances of authenticity are situated within particular communities of practice, where normative practices of interaction and identification are co-constructed from below to index particular identities (Jenson et al 2015; Marwick 2013; Sharma 2013; Whitehead 2015). In the case of these disinformation accounts, the orientation to existing naming practices on Twitter may contribute to constructing valid representations of Twitter users. By naming themselves in ways that appear readily established on Twitter, i.e., presenting the same identity aspects across handles and usernames, the accounts appear to demonstrate a knowledge of (and thus, potentially a belonging to) the communicative environment of Twitter.

4.5.1 News accounts

As mentioned previously, the findings of this study corroborate previous research that suggests some disinformation accounts employ news-related language to construct identity performances as news accounts (Andrews et al 2016; Linvill and Warren 2018). This section will discuss how news accounts use naming strategies to authenticate their identity performances, particularly in relation to pre-existing naming practices on Twitter.

An orientation to pre-existing norms is evident in the types of handles and usernames employed in news accounts within the more popular account sample. As previously mentioned, this largely takes the form of a combination of location and news branding terms. A repetition of existing news naming practices can also be seen via the construction of, to some extent, shared branding. As previously acknowledged, it is not uncommon for news organisations to have location-specific pages, such as ‘CBS Chicago’ and ‘CBS New York’. These accounts are clearly connected to one another via the recurrent use of ‘CBS’ and the adoption of an identical naming structure (news organisation + location).

While news branded accounts in the more popular sample did not use established news organisation names like ‘CBS’, the accounts could be placed into smaller groups based on news branding terms and naming structures. Table 14 provides an overview of the key news branding terms used in the data: Today; Online; Daily; Top News. Notably, in each case, the accounts not only use the same key news branding terms, but adopt identical naming structures, potentially indexing a shared brand and adopting existing shared naming practices associated with established news organisations.

Table 14: Shared news naming strategies in the more popular sample

News-branding term	Handle	Username
‘Today’	@TodayNyCity @TodayPittsburgh	New York City Today Pittsburgh Today
‘Online’	@OnlineCleveland @OnlineMemphis	Cleveland Online Memphis Online
‘Daily’	@DailySanDiego @DailySanFran	San Diego Daily San Francisco Daily
‘Top News’	@ElPasoTopNews @SanAntoTopNews	El Paso Top News San Antonio Top News

Unlike the more popular accounts, the less popular accounts in the data set seldom display this level of uniformity between handles and usernames. In fact, only one account among the less popular ones mobilises the ‘Online’ term in both metadata units (handle: @BaltimoreOnline, username: Baltimore Online). Even in this case, we notice that the account uses a 0 to represent the letter ‘o’ in the handle. This common feature of ‘netspeak’ which has associations with informality and unprofessionalism amongst the general public (Squires 2010:473) appears marked in news contexts where there are expectations for Standard English. Such a use of language therefore distances the account from the genre it aims to be identified with, as well as from the ‘more popular’ accounts that use standard language. Inconsistency is

also evident in the less popular account, @MissouriNewsUS, that has the username ‘People for Justice!’. The username uses neither ‘topical’ branding, nor does it mirror norms of pre-existing news organisations, practices that have previously been established as credible and influential within a news sharing context (Andrews et al 2016; Morris et al 2012). The differing degree of orientation and adherence to pre-existing news-sharing naming practices may explain why certain accounts have been more popular than others; those with more followers may indicate that they were more frequently taken as valid representations of news-sharing Twitter accounts. In this context, it seems that authentication processes appear to draw on the extent to which the naming practices of the disinformation accounts mirror what users stereotypically associate with verified and established online accounts of news networks. By mirroring established norms and practices, accounts can construct representations of Twitter news accounts that can be taken as valid by users based on their expectations about what news accounts tend to look like.

As previously mentioned, Jäkälä and Berki (2013:7) suggest that the use of eponyms can increase perceptions of trustworthiness and authenticity due to the link they establish between the virtual and ‘real’/physical world. Based on the prevalence of news-branded accounts that closely adopt -and orient to- pre-existing norms associated with news sharing on Twitter, I suggest that trustworthiness can also be created by mirroring practices and norms established in the so-called virtual world. Rather than authentication being mainly achieved through a link between the online and offline sphere, authenticity can also be performed by creating a sense of familiarity with what users ordinarily experience in different contexts and environments on an online platform.

4.5.2 Personal accounts

Previous research has explored how the content posted by individual Twitter disinformation accounts can contribute to their performances of authentic American personas (Xia et al 2019). This section will discuss how the use of naming strategies can also contribute to processes of authentication within accounts that present themselves as individuals.

As previously acknowledged, research suggests that eponymous naming strategies can be perceived as trustworthy and authenticating because of the ‘notion of familiarity’ that they create between the offline and online spheres. This chapter has shown that across both more and less popular accounts, it is not uncommon for accounts to draw on first names and surnames when identifying themselves via handles and usernames. Along with the suggested link that such practices may evoke with a person’s offline identity, this practice also appears to be already established amongst Twitter users. In Olivier’s (2014) study of Twitter handles and

usernames, he finds that across various categories (such as names with special characters, names with the addition of numbers etc.), first names and surnames are consistently drawn on as identifiers. Similarly, in their research into online anonymity, Peddinti et al (2017:85) find that of a random sample of 100,000 Twitter accounts, 68% were ‘identifiable’ in that they used both first names and surnames. This suggests that the disinformation accounts in this sample that use first names and surnames adopt naming practices that can be considered somewhat typical of Twitter accounts linked to individuals.

Not only does it appear that certain kinds of names are used more frequently than others on Twitter, but as mentioned in section 2.5, research also suggests that certain kinds of names can be associated with different value judgements. In their study of users’ perceptions of different identification practices on Twitter, for example, Morris et al (2012) find that ‘internet names’ are regarded as less credible than ‘traditional’ or ‘topical’ names. Moreover, associations between anonymity and nefarious behaviour have led to calls for Twitter to change their naming policy and make users utilise their ‘real names’, which can be verified using official documents (Twitter 2021b). As a result, it appears that the use of first names and surnames on Twitter may be regarded with less suspicion than other, less identifiable kinds of names. The use of eponymous naming strategies by disinformation accounts therefore appears to demonstrate an adoption of established naming practices that could arguably create a ‘notion of familiarity’ (Jäkälä and Berki 2013:7), not only with the offline, so-called “real” world, but with what users typically see on Twitter. Despite the role that eponymous naming strategies play within these disinformation accounts, it is important to acknowledge that accounts posing as individuals did not exclusively use first names and surnames. Some of the most popular Internet Research Agency accounts (like @USA_GunSlinger, @SouthLoneStar2, and @TEN_GOP), draw on other identity aspects, with location appearing to be particularly relevant.

While geographical location does not identify an individual to the extent that a proper name does, the use of this identification strategy can also arguably evoke familiarity with aspects of the offline world. It could be argued that geographical location can be perceived as providing a tangible link with the places a person inhabits in the physical world, potentially increasing links between a user’s online (the so-called “virtual”) and offline (the so-called “real”) presence (cf. Jäkälä and Berki 2013) and going some way to authenticating their online identity performance. Beyond the idea that location can index a sense of familiarity with the offline world, it is also necessary to acknowledge that location appears to play a specific role within the communicative context of Twitter.

On Twitter, research suggests that users ‘usually specify their geographical location in their profiles’ (Takhteyev et al 2012:73). On Twitter profiles, there is limited space for information about an individual; when setting up or editing their profile, users see five pre-determined boxes for information: name, bio, location, website, and birth date (Twitter 2022c). If shared, this kind of information all appears in the same area of a user’s Twitter profile e.g., underneath the profile picture, handle, and username (see Figure 1). Although users do not need to populate the last four boxes with information, in their study of the geography of Twitter networks, Takhteyev et al (2012:75) found that of a random sample of 481,248 tweets, 75% shared information about their locations. This took the form of either a location that a user had posted to their profile themselves, or in a small number of cases, a location that had been automatically generated by Twitter. The pre-determined space for locational information within Twitter profiles suggests that location may play a different role on Twitter than other, less frequently invoked identity aspects like political affiliation, religion, etc. (see Table 10 and Table 11) that do not have designated spaces within Twitter’s pre-determined options for profile information.

Not only does research suggest that users share information about their location in their bios, but Olivier’s (2014) research into naming strategies finds that both individuals and businesses sometimes make reference to geographical location, specifically country, in their handles and usernames (such as Brendan Boyle @BrendanSAfrica). It therefore appears that the invocation of location in naming strategies is not specific to this sample of disinformation accounts but is also a practice observed within Twitter more widely.

Beyond naming strategies and bio information, users can also share information about their location via the geolocation of individual tweets. When tweets are geolocated, information about location appears alongside the tweet itself. Although Jurgens et al (2015:188) argue that less than 1% of tweets are geolocated, geolocated tweets have proven a valuable source of data for studies into sociolinguistics and language variation, geographically preferred topics, and disinformation identification to name but a few (see Deligiannis et al 2018; Ilbury 2020; Pavalanathan and Eisenstein 2015). It therefore appears that the symbiotic relationship between the pre-determined options for identification on Twitter, and the ways that users share information about themselves, has led to location having an established role in communication and identification on Twitter.

Considering the relevance of collective identities in public participation in media contexts (see section 2.3) and the prevalence of pseudonyms and collective identity aspects (such as geographical location) in my data, I argue that Jäkälä and Berki’s (2013) suggestion that naming strategies can be used to authenticate identity performances and invoke trustworthiness can be

extended beyond eponyms and encompass pseudonyms as well. Rather than pseudonyms being approached as strategies that hide personal identity features, they can instead be used to perform identity by drawing on collective identity features. Consequently, pseudonymous naming strategies can contribute to authenticating processes by drawing on collective identity features in ways that can be taken as valid within the communicative context at hand.

4.6 Conclusion

Based on the discussion of both news-branded and personal accounts, I argue that authenticating processes are not associated with a particular type of naming strategy. Instead, they are afforded by the degree of information shared that allows identification with norms and practices associated with what we experience in different contexts and environments, both online and offline. Eponyms and pseudonyms, thus, can be placed in a continuum of authenticity as depicted in Figure 3; both naming strategies can authenticate identity performances depending on the extent to which they mirror normative practices associated with the specific context within which the communication is situated.

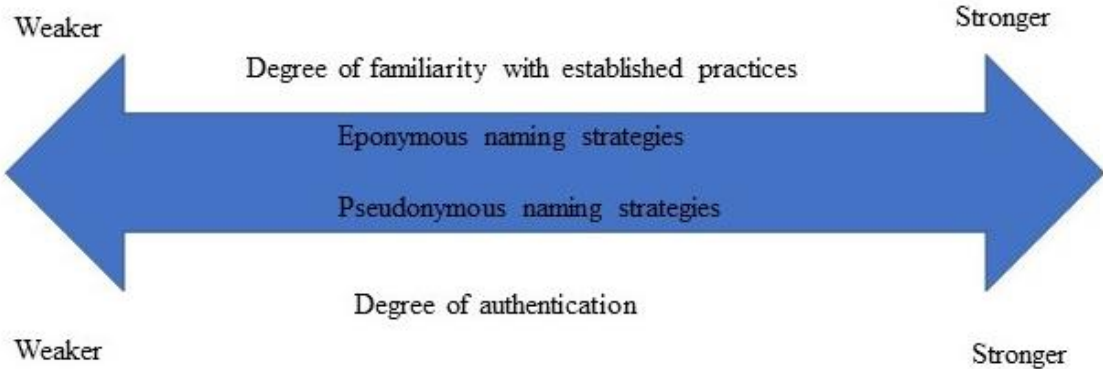


Figure 3: Continuum of naming strategies in relation to authenticity

As discussed in section 2.2.2, mainstream media discussions of bots, trolls, and disinformation suggest that certain metadata features, such as handles containing long series of numbers, can indicate ‘inauthentic activity’ online (Dotto and Cubbon 2019). As a result, users are warned against trusting such sources. However, this analysis reveals that popular disinformation accounts overwhelmingly orient to normative expectations associated with particular contexts. I argue that these disinformation accounts constructed valid representations of individual people and news organisations by appropriating existing naming practices related to these types of accounts. Previous research suggests that the concept of normativity emerges

and is assessed from below and can be located in ‘different layers of language use’ (Kytölä and Westinen 2015:18). This chapter has examined the layer of naming strategies, but research suggests that ‘other semiotic resources such as pictures’ can be understood as another layer of language use where normativity is negotiated and co-constructed on social media (Kytölä and Westinen 2015:18).

As acknowledged previously (see section 2.2.2), images are consistently associated with identification practices, both in the form of official authorised documents like passport photos, and online, in the less strictly regulated form of profile pictures. As depicted in Figure 2, profile pictures appear alongside naming strategies and tweet content in a user’s home timeline. These features have consequently been described as those available to users ‘at-a-glance’ when assessing account credibility (Morris et al 2012:444). As a result, profile pictures are often discussed alongside naming strategies in relation to ‘inauthentic’ activity online. In these discussions, certain markers, such as the use of stock images or celebrities as profile pictures, are highlighted as being potentially unreliable, and thus more likely to be associated with disinformation (Calabrese 2020; Dotto and Cubbon 2019). To understand authentication processes on Twitter, it is therefore necessary to extend the study of identification practices to profile pictures.

Chapter 5 Authenticating Identity Performance on Twitter

Disinformation Accounts: Profile Pictures

5.1 Introduction

On Twitter, all users have a profile picture associated with their account. As displayed in Figure 1 and Figure 2, these images appear in a circle on a user's profile and are shown alongside all of their tweets. All users are encouraged by Twitter to upload a visual identifier, but they also have the option to abstain from using an image of their choice. In these cases, a 'default' profile picture is provided by Twitter, as displayed in Figure 4.

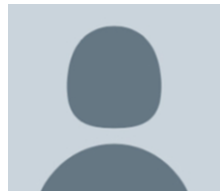


Figure 4: A default Twitter profile picture

At the time of writing, Twitter users have the freedom to choose any kind of profile picture, and there are currently no means to verify that an image holds a physical likeness to a user. Users can employ different kinds of images as profile pictures and image choice can influence the way that other users assess accounts and the information that they share in terms of their credibility (Morris et al 2012). Figure 5, Figure 6 and Figure 7 are used as profile pictures in my dataset and they exemplify some of the different possibilities for visual representation on Twitter: anime or cartoon characters, logos, and photographs of people (Morris et al 2012:445).

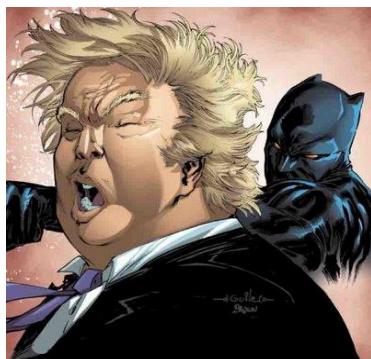


Figure 5: Cartoon as profile picture



Figure 6: Logo as profile picture



Figure 7: Photograph of person as profile picture

This chapter aims to investigate if and how profile pictures on Twitter disinformation accounts contribute to the authentication of particular personas. Firstly, I will discuss previous approaches to visual analysis and explain how they can be used to analyse and understand performances of authenticity in the context of Twitter profile pictures. The analysis will draw on Kress and van Leeuwen's work on visual modality (2006), El Refaie (2010)'s approach to visual authenticity, and Ledin and Machin's (2018) adaptation of Kress and van Leeuwen's (2006) social semiotics model. Then, I will qualitatively analyse how the five main accounts use profile pictures. Similarly to the handles and usernames, the profile pictures will be analysed in terms of what identity aspects they present. I then use the categories generated from the qualitative analysis to analyse a wider sample of more and comparatively less popular disinformation accounts, thus situating the findings from the qualitative analysis within the context of broader disinformation practices.

5.2 Visual analysis and authenticity

According to Ledin and Machin (2018:47), Barthes' (1977) work provides basic tools that can be used to explore the kinds of questions we can ask of a photograph. Drawing on Barthes (1977), they suggest that one useful way of analysing images is to distinguish between 'denotation' and 'connotation'. Denotation is a term used to refer to what is visibly present in an image, what is 'literally there', while connotation refers to the meanings that are associated with the images in question. In the case of an advertisement, for example, a female figure wearing an evening gown may be denoted, while connotations of classical elegance and beauty may be associated with the woman due to her posture and clothes (Penn 2000:238). To grasp a sign's connotation, 'additional cultural or conventional knowledge is required' (Gaskell and Bauer 2000:351). This means that connotations can change based on social norms, with an image having different meanings for people from different social groups or cultural backgrounds. Acknowledging what is denoted in an image (describing what is there rather than how we view it), as well as its connotations, is part of a systematic analysis of images (Ledin and Machin 2018:47).

For the analysis of profile pictures and their role in authentication processes, attention will be given to both what the pictures depict, and the potential sociocultural meanings that the pictures evoke. Visual means, such as images, play a significant role in the communicative environment of social media, where, as mentioned in chapter 2, the uploading of the user's profile picture is included in the process of signing up and creating an account. In other words, these images are key identification means for social media users and may contribute to how

authentic such identifications are perceived to be by the online audience. As a result, the focus in this chapter is on a) what identity aspects are being (re)presented visually and b) how these visual (re)presentations may contribute to authentication processes of the identity in question.

The meaning potential of visual means such as colour, for example, have been noted by Kress and van Leeuwen (2006). Kress and van Leeuwen (2006) draw on Hallidayan linguistics and suggest that visual modes can be analysed in similar ways to verbal language. With respect to meanings related to authenticity, they suggest that Halliday's concept of verbal modality (the expression of judgement of probability and obligation) can manifest in images in terms of whether, and to what degree, a visual presentation is taken as true, real, and/or valid (El Refaie 2010:163; Kress and van Leeuwen 2006:204). They refer to this idea as 'visual modality' and suggest that images can be investigated through three levels of meaning: ideational (representation of the world around and inside us), interpersonal (enactment of social interactions as social relations) and textual (presentation of a coherent 'world of the text') (Kress and van Leeuwen 2006:15). In other words, they suggest that images can be analysed in terms of a) what is represented in the image, b) the social relationship and interaction between the image and the viewer and c) how the image is organised to create meaning. According to Kress and van Leeuwen's (2006) visual modality model, degrees of validity and the representation of truth can be interpreted from the resources that make up an image, such as colour. They refer to these factors as 'validity markers', a term which is used to include: contextualisation, depth, colour differentiation, colour saturation, colour modulation, illumination, brightness, and representation.

5.3 Identification in profile pictures: visual authenticity on Twitter

To understand how the profile pictures in my sample may have contributed to authentication processes, it is first necessary to consider how images can make meaning. Firstly, I will present Kress and van Leeuwen's (2006) model and, particularly, the ways in which dimensions of colour contribute to meaning and modality. Then, I will discuss Kress and van Leeuwen's (2006) theory in relation to other visual analysis approaches, and especially, their potential application to social media research. Finally, I will explain how the profile pictures in my study will be analysed.

5.3.1 Kress and van Leeuwen: visual modality

According to Kress and van Leeuwen (2002:356), meaning making is achieved through colour and, particularly, through its dimensions of saturation and brightness. In terms of saturation,

colours are distinguished between those with high saturation, e.g., the styles of bright primary colours used in cartoon rainbows, and low saturation such as more muted, dilute colours like black and white or sepia effects. According to Kress and van Leeuwen (2002:356), highly saturated colours are often associated with positivity, exuberance, and adventurousness, while low saturation evokes ‘subtle and tender, but also cold and repressed, or brooding and moody’ meanings. Colours that appear in their most ‘extreme’ form, such as in their highest or lowest saturation, have usually required editing. As a result, Kress and van Leeuwen (2006:160-161) suggest that colours that sit around two-thirds up the continuum of saturation reflect natural life most accurately; they are what the naked eye tends to see in the natural world, or in practice, what we tend to see in standard 35mm photography. As Kress and van Leeuwen (2006) define modality in terms of how truthfully something appears to be represented, colour saturation that is associated with reality can thus be understood as having the highest modality, i.e., it is very likely to be interpreted as a valid representation of what is true or real. However, it is not the case that the same relationship between saturation and truth will always be observed; colours of median saturation will not **always** evoke ideas of truthfulness. Instead, the significance and meaning of colour (and other markers) can differ based on what is considered valid within different contexts.

Kress and van Leeuwen (2006) use the term ‘coding orientation’ to refer to the idea that different principles of reality will dominate how a text and/or image is received by specific social groups in specific contexts. Specifically, they distinguish four coding orientations: technological, sensory, abstract, and naturalistic. The attribution of a particular orientation to a type of discourse is determined by the aim of the text or image in question within its context of use: technological (‘the effectiveness of a visual representation as a blueprint’), sensory (‘pleasure’), and abstract (how much an image ‘reduces an individual to the general and how much it essentialises concrete qualities’) (Kress and van Leeuwen 2006:265).

Naturalism refers to the notion that an image is interpreted as a valid representation of what is true and real based on how closely the image matches what one would see with their own eyes. As intimated above, within the naturalistic coding orientation, images can be understood as having the highest modality when they reflect what we expect to see in a standard photograph (around two-thirds up on all modality markers) (Kress and van Leeuwen 2006:160-163). According to Kress and van Leeuwen (2006), naturalism is the coding orientation in which most things in the western world can be analysed, due to its ‘common sense’ characterisation.

5.3.2 Beyond visual modality

Despite the invaluable contribution of this model of visual analysis to the field, the idea that naturalism can be captured using stylistic markers appears problematic. As pointed out by El Refaie (2010), not all images have the same relationship with naturalism as suggested by Kress and van Leeuwen (2006), i.e., being most valid when they are two-thirds up the scale of each validity marker. El Refaie (2010:162) uses the genre of autobiographical comics to demonstrate how an authentic presentation may have more to do with having a ‘transparent relationship to its object of representation’, a defining feature of the concept of visual authenticity, rather than the types of markers set out by Kress and van Leeuwen (2006). The concept of ‘authenticity’, specifically when applied to visual genres (see Knieper and Müller 2003), suggests that the truthfulness of an image is more closely related to how faithfully an image creator presents the truth ‘as they see it’, or experienced it, rather than by the image style that they use to do so. In the case of autobiographical comics, an artist could use line drawings of stick people to depict their interpretation of an experience that they had. El Refaie (2010) suggests that this kind of spontaneous, ‘unrealistic’ drawing style has come to signify greater authenticity in the context of comics than the kind of hyper-realist styles that may more closely correspond with the idea of naturalism. Visual authenticity, thus, as defined by El Refaie (2010) is closer to the approach to authenticity that I adopt in this thesis: authenticity is not a static feature related to specific markers, but a dynamic process that can be performed and negotiated, with interpretations of authenticity being influenced by what is taken as valid within the context of the interaction.

Building on Kress and van Leeuwen (2006), Ledin and Machin (2018) provide a model for analysing images that can be applied to specific domains, such as photographs, documents, and packaging. They draw on existing approaches to visual analysis, such as Barthes (1977) work on denotation and connotation, and Kress and van Leeuwen’s (2006) work into the meaning potential of colour, to present systematic toolkits for breaking down and analysing photographs. Objects in photographs, they suggest, can be analysed in terms of how the object’s colours and any related symbols may evoke sociocultural meanings. On the other hand, people can be analysed in terms of descriptions of the participants in the photograph (How many are there? What do they look like?) and how they are positioned. When discussing what people look like, Ledin and Machin (2018:55) suggest that people can be represented in ways that visually categorise them on the basis of culture and/or biology. They suggest that cultural categorisation can be signalled via ‘kinds of dress, hairstyles, body adornment and so on’, while biological categorisation is indicated through ‘emphasis on stereotyped physical characteristics’, which can be viewed as both positive and negative, depending on the sociohistorical context within

which they are situated. Consequently, objects and humans can both be analysed using Ledin and Machin's (2018) model.

5.3.3 The visual analysis of Twitter profile pictures

Profile pictures use 'signs and symbols to create meaning around one's presence' (Lough et al 2018:1279) and can be understood as one of the tools involved in the projection of a 'specific likeness to the public' within social media and internet use (Lough et al 2018:1270; Schau and Gilly 2003). The suggestion that features like profile pictures can be used to project 'likeness' demonstrates how and why profile pictures may contribute to authentication processes; a cartoon may project a different kind of 'likeness' than say, a photograph, particularly when considered against the 'common sense' coding orientation of naturalism.

Having said this, it has been acknowledged that the changing standards of photography and image production within the digital age may lead to a shift in what is considered 'natural' and normal (Ravelli and van Leeuwen 2018:284). Not only has the development in photography technology led to the standard average 35mm photograph no longer being a valid de facto naturalistic standard (cf. Kress and van Leeuwen 2006), but all users can edit, retouch, and filter the images that they produce and share on social media (Ravelli and Ven Leeuwen 2018:280). It appears that this practice is not only possible, but broadly accepted- and to some extent expected- online; Veum and Undrum (2017:93) analysed Instagram photos and found that 78% of selfies analysed had been 'obviously edited before being published', while other research suggests that users change dimensions of colour or even the shape and appearance of some facial features to create filtered images of perfection (Åberg et al 2020; Tiggermann et al 2020:176). Despite these changes to normalised expectations relating to image production, the preference for photographs of people over cartoons as identifiers on social media (Morris et al 2012:445) does demonstrate the continuing relevance of 'common sense' orientation of naturalism in social media contexts, a type of naturalism that is perhaps more closely related to what the image depicts (e.g., photograph of a person), rather than how it is presented (e.g., modification of saturation).

Drawing on the aforementioned discussions of various approaches to visual analysis, the denotation and connotations (Barthes 1977) of each profile picture used by the five main accounts will be analysed by attending to the following areas:

a) What is represented from the world around us? For example, is the profile picture of a human or nonhuman?

- b) How are the profile pictures and what they depict presented? To investigate this, the following dimensions from both Kress and van Leeuwen's (2006) and Ledin and Machin's (2018) models will be drawn on: colour, setting, participants, positioning, and categorisation. The analysis will consequently aim to answer questions such as the following: What colours are used in this profile picture? Is the profile picture situated against a background? If the profile picture features people, how are they positioned? Where is their gaze directed?
- c) What sociocultural connotations are evoked by what is presented in the profile picture? For example, how does the choice to present a participant wearing a certain item of clothing evoke particular social and cultural meanings? If the image is not of a human, what sociocultural meanings are attributed to this object?
- d) What identity aspects are mobilised by the visual representations and how do these identity aspects interact with the type of identification signalled by the other metadata units, i.e., handles and usernames?

The findings will be interpreted by using an approach to authenticity that follows El Refaie (2010), where authenticity is regarded as a performance, depending on the context within which the interaction is situated and on the specific audiences being addressed. Categorisations established in the qualitative analysis will then be used to analyse the identity aspects presented in the wider samples of more and less popular disinformation accounts.

5.4 Identification through profile pictures: A qualitative analysis of the five main accounts

5.4.1 @TEN_GOP

@TEN_GOP had two profile pictures while the account was active: Figure 8 is the initial picture used and Figure 9 illustrates the picture used from June 2017 until the suspension of the account.



Figure 8: @TEN_GOP profile picture (August 2016- May 2017)



Figure 9: @TEN_GOP profile picture from June 2017 onwards

A reverse image search (see section 3.4.2.1 for information about this technique) for Figure 8 resulted in 160 matches. Most of the results were black and white images of the state seal of Tennessee, revealing that the profile picture uses a popularly distributed pre-existing symbol. Figure 9 also brought up very similar results. Exact matches for both figures do not appear in the reverse image search, with the exception of instances related to the disinformation account, such as web pages that archived Internet Research Agency data.

What is represented from the world around us?

The first profile picture used from August 2016-May 2017 is an image of two ‘objects’, rather than humans: the seal of the state of Tennessee (Figure 10) and the confederate flag (Figure 11). The seal features a plow, a sheaf of wheat, a cotton stalk, and a boat, and it includes the roman numerals ‘XVI’ which represent the number 16. The other text on the seal reads ‘AGRICULTURE’ and ‘COMMERCE’ and around the outside it reads ‘THE GREAT SEAL OF THE STATE OF TENNESSEE 1796’. The second profile picture contains the same seal, but it is overlaid on the American national flag (Figure 12), rather than the confederate flag.



Figure 10: State seal of Tennessee



Figure 11: Confederate flag



Figure 12: American flag

How are the key objects presented in the profile picture?

The seal is presented in black and white and superimposed firstly over the confederate flag, and then over the American flag. The entire seal is visible which makes it easily recognisable and reflects the way the seal is presented when used on official documents. The seal has a translucent quality to it, meaning the respective background that it is superimposed upon can still be seen. Such stylisation suggests that multiple identity aspects are presented at the same time, due to the clear visibility of more than one symbol.

In the first profile picture, the image of the state seal is superimposed over the symbol of the confederate flag, made recognisable by the red background, blue X, and white stars. Despite the state seal covering parts of the flag, there is enough visibility of the aforementioned symbols to make the flag identifiable. The appearance of a flag is further exacerbated by the appearance of a ripple-like effect. The use of light and shade to create ripples works to create the appearance of something moving, or more specifically, billowing, in the way that a flag may appear when blowing at the top of a flagpole.

The second picture (used from June 2017) shows the Tennessee state seal superimposed over an American flag which appears in black and white, unlike the coloured confederate flag in the earlier version. The flag appears to have been zoomed into, as it is not visible in its entirety. Despite this, as well as the black and white effect, the flag is still recognisable due to the visibility of white stars on a darker background and stripes underneath that also alternate between light and dark.

What are the sociocultural connotations of the images depicted in the profile picture?

In both profile pictures, the central image is the seal of the state of Tennessee. The selection of this symbol signals an orientation towards identity aspects that foreground local identity linked with the geographical area of the American South, as opposed to the broader national identity of the US. Seals of states have been historically used to identify and/or authenticate legal documents. As is the case historically with symbols of national identity (such as flags), state seals also serve as a way of symbolising pride in one's state, evident in the nickname of the Tennessee seal, 'the Great seal of Tennessee' from as early as its creation in 1801 (White and Sevier 1901). The elements of the seal itself serve to underline the local identity of Tennessee as a farming state and the prominent role of the three large rivers (Tennessee River, Cumberland River, and Mississippi River) in trading. The Roman numerals signify that Tennessee was the 16th state to be formed (in chronological order) and there is a reference to 1796 when the constitution of the state took place. The state seal remains in use in its official capacity as an

authentication stamp of legal documents. As a result, the use of the state seal in a profile picture appears to represent local identity features linking to state pride.

The confederate flag, on the other hand, has long been a controversial symbol that traces back to the American Civil War, where confederate states fought to uphold the right to own slaves. The flag features a red background, with a blue cross and 13 white stars. Each star represents a confederate state in the American South: South Carolina, Mississippi, Florida, Alabama, Georgia, Louisiana, Texas, Virginia, Arkansas, Tennessee, North Carolina, Missouri and Kentucky. The display of the confederate flag alongside the state seal of Tennessee consequently reinforces the presentation of identity aspects related with the American South. While some argue that the confederate flag symbolises pride in ‘non-racist Southern spirit’ and celebrates a ‘legacy of ancestors whose valor became legendary in military history’ (Jones 2015), the flag is also widely recognised as a symbol of intolerance and racism. Politically, Republican legislators and supporters of the Republican party are more likely to support the display of the Confederate flag than Democrats (Huffmon et al 2016: 727,737), indicating a potential link between the confederate flag and conservative politics. In Figure 8, the characterisation of Tennessee as a confederate state is dually represented via the display of the State’s seal and the confederate flag. More importantly, the state is not only represented in geographical terms, but also in political ones.

The removal of the confederate flag from the picture and its replacement with the US flag in June 2017 minimises the overt connotations with slavery and racism associated with confederate imagery. Instead, the profile picture appears to tie the state and the nation together, indicating an affiliation with a broader collective group (Americans), rather than those who share confederate ideologies.

The shift from highly saturated primary colours (red and blue) in the first profile picture, to black and white in the second profile picture, also appears to signify a change in mood. As mentioned in section 5.3.1, colour is a semiotic resource that is ‘multifunctional in its uses in the culturally located making of signs’ (Kress and van Leeuwen 2002:356). The use of colour to act as a signifier of a particular mood or stance on social media has previously been employed in response to political decisions and landmark events¹¹. While here, the change to black and white does not appear immediately associated with viral trends, it could be argued that the change from a colourful confederacy to a black and white national flag could be reflective of

¹¹ For example, following the decriminalisation of abortion in Northern Ireland in October 2019, anti-abortion campaigners changed their social media profile pictures to black backgrounds as a reference to the landmark ruling being the ‘darkest day’ for Northern Ireland (Belfast Telegraph 2019).

different attitudes that the profile owner has towards the different social groupings represented by each flag (confederacy versus USA more generally): the change to black and white evokes more subdued feelings of enthusiasm and warmth towards the national symbol. At the same time, though, the removal of the confederate flag broadens the group with which the user potentially affiliates with. Consequently, the change in profile picture could function to demonstrate a shift in mood in relation to attitudes towards the confederacy and the USA more generally, while simultaneously broadening the potential for affiliation with wider groups.

What identity aspects are mobilised by the profile pictures and how do these aspects interact with the identity aspects signalled by the handles/usernames?

Both profile pictures used by @TEN_GOP present identity aspects linked to location: the specific state of Tennessee and the nation of the USA. More specifically, the inclusion of the confederate flag evokes identity aspects linked to both location (the American South) and affiliation with right-leaning politics. The historical link between the American South, the confederacy, and conservative politics contributes to the construction of a valid representation of a conservative persona by presenting identity aspects that draw on pre-existing stereotypical sociocultural associations. Furthermore, the use of the state seal, a symbol associated with the authentication of legal documents, on a social media profile, arguably carries connotations of authenticity due to its functional purpose in offline contexts. Despite changes in other parts of the profile picture, the continued use of the seal suggests that the local area of Tennessee is the most salient identity feature for the page owner. This interpretation is further reinforced by the reference to the state in both the handle and the username. The removal of the confederate flag from the profile picture coincides with the removal of ‘GOP’ from the username. This change results in location becoming the only identity aspect presented by the profile picture and username at the same time. The presentation of similar identity aspects in the naming units and profile pictures (i.e., location and political affiliation) may contribute to authentication processes by constructing an identity performance that appears stable and coherent across different metadata units; the suggestion that (in)consistency can contribute to constructions of (in)authenticity online has previously been discussed in studies into online identity performance (see section 2.3.1).

5.4.2 @Pamela_Moore13

@Pamela_Moore13 used the same profile picture throughout the account’s activation (see Figure 13). A reverse image search does not produce any results, indicating that the image was not used in other online contexts. It is therefore not possible to know whether the profile picture

is from a photography exhibition or other artwork, or whether it was shot particularly for the purposes of the account.



Figure 13: @Pamela_Moore13 profile picture throughout activation

What is represented from the world around us?

The profile picture used by @Pamela_Moore13 is a black and white image that features a topless woman who appears to be white (although it is difficult to be sure due to the black and white saturation). Her bare chest appears to be covered by tape, she is partly wrapped in a USA national flag, and she is wearing a niqab-style hair and face covering.

How is the participant presented?

The profile picture is presented in black and white. The use of black and white is commonly associated with ‘mellow’ and ‘reserved’ moods (Ledin and Machin 2018:50), as well as coolness, distance, and ideas of repression (Kress and van Leeuwen 2006:233). In the context of social media in particular, black and white is also used to signal historicity, echoing the aesthetic of old black-and-white photographs signalling retro-ness (Beil 2015:140). The saturation of the image suggests editing; standard photography no longer presents images in black and white, but rather, as Kress and van Leeuwen (2006) suggest, colours are usually around two-thirds up the scale of each validity marker.

The participant is posed and shot in a way that appears to have been taken by someone other than the person depicted; in other words, it is not a selfie. Zappavigna (2016:278) refers to these types of images as being shot from a ‘with photographer perspective’, that is, they appear to have been photographed by someone other than the subject in the photo. In contrast, she uses ‘as photographer subjectification’ to refer to selfies, where the photographer and the subject of the image appear to be the same person. In the profile picture above, the ‘with photographer subjectification’ is made evident from the distance between the subject and the camera, the direct gaze towards the viewer, and the posed nature of the participant, with both of the subject’s

hands being placed in a way that creates the effect of being partially wrapped in the American flag. The juxtaposition of nudity with a clothing item that can be culturally categorised as associated with Islam and modesty exacerbates the distance between the ideology associated with the garment and the connotations of nudity evoked by the woman.

What are the sociocultural connotations of the person depicted in the profile picture?

The way in which the woman is presented in the photo challenges general notions of modesty and more specific norms associated with Twitter profile pictures. Apart from certain sub-communities of accounts (e.g., those committed to pornographic content), it is unusual to see Twitter profile pictures featuring topless women.

The challenge to modesty is exacerbated by the combination of nudity with a niqab-style head covering. This combination is very marked in terms of societal expectations, as it brings together nudity and head-face wear that is typically associated with modesty (Piela 2013:785). The covering shown in the profile picture resembles a niqab, an Islamic garment that covers the face. Often, niqabs are accompanied by ‘full-length jilabs (long robes), and hijabs (headscarves)’ and are described by their wearers as demonstrating a commitment to faith and showing their identities as practising Muslim women (Zempi 2016:1740-1743). Within this profile picture, a demonstration of allegiance to the American flag appears to highlight and reinforce a rejection of Islam. This is achieved through juxtaposing exposed skin and nudity with a religious garment associated with modesty. The exposed woman being wrapped in an American flag consequently indexes a rejection of Islam in the name of American nationalism. The relationship between American nationalism and Islam draws on pre-existing anti-Islamic discourses in the US, that frame Islam as being ‘anti-Christian, anti-government, and anti-liberalism’ (Sririam 2016:50), and as a result anti-American.

What identity aspects are mobilised by the profile picture and how do these aspects interact with the identity aspects signalled by the handles/usernames?

Based on the discussion above, the profile picture presents aspects of identity relating to location and political affiliation through the use of objects like the flag and the niqab. At the same time, the depiction of the topless woman indicates identity aspects related to gender (female) and race (probably white). The appearance of a person in the picture reinforces the construction of a female persona which was also revealed in the analysis of the account’s handle and username. While the naming features were not associated with political affiliations (cf. @TEN_GOP), the composition of the image reveals ideological stances that can be summarised as provocative to the Islamic religion and nationalistic towards the US. Consequently, the image

is found to contribute additional identity features to the account, in contrast to @TEN_GOP where the picture appears to complement and directly map to the identity features evoked in the naming units. The use of black and white in the image suggests editing as the subject is not presented in a way that one's naked eye would see the woman, nor in a way that we would typically see in modern standard photography (cf. Kress and van Leeuwen 2006). The distance that the photograph is shot from, together with the black and white filter and the facial covering, makes the facial characteristics of the subject less visible, limiting the potential for the profile picture to be used to identify a person in the offline sphere.

While the absence of identifiable facial features may violate established practices relating to social media profile pictures (i.e., that profile pictures should present a 'true likeness' of the individual (UK Government Department for Digital, Culture, Media and Sport 2022)), the account's consistent presentation of female identity aspects across metadata units may add validity to the representation of @Pamela_Moore13 as a woman.

5.4.3 @wokeluisa

@wokeluisa used the same profile picture throughout the account's activation (Figure 14). A reverse image search shows that the image has previously appeared on Pinterest and Tumblr in threads of users who are 'bombshells'¹². There are also results linking the photo to hair advertisements. None of these websites are still active. The results in the reverse image search show the person's body (one shoulder and chest) unlike the image displayed in Figure 14. The profile picture therefore appears to be a cropped version of the image that was used elsewhere.



Figure 14: @wokeluisa Profile picture throughout activation

What is represented from the world around us?

¹² The threads on Tumblr and Pinterest were linked with profiles that were not called 'Luisa' but contained many other photos of the same person i.e., it appears this image was taken from an existing social media user.

The profile picture is a selfie of a young, Black woman, wearing makeup and pouting. The style of the photo, as a close-up selfie, is in line with expectations associated with photos posted by young women on social media (Veum and Undrum 2017:95).

How is the participant presented?

The profile picture is taken from an ‘as photographer’ perspective (Zappavigna 2016:278) and embodies the kind of image that is strongly associated with the use of front-facing mobile phone cameras (Shah and Tewari 2016). The photo is a headshot taken from a slightly sideways angle with an upward gaze into the camera. These types of shots have been described as the ‘best’ way to take a selfie in mainstream media (Chávez and Cardellino 2017), suggesting that @wokeluisa’s profile picture shows an orientation to established selfie-taking practices. The image is slightly pixelized, potentially since the image is a cropped version of the original photo. The young woman is not smiling but rather posing (pouting) in a way commonly associated with young women on social media (Dobson 2015:38). The background is barely visible and does not appear to situate the subject in a particularly recognisable environment. The close-up nature of the photo does not afford any props or effects in the background that would signal extra meaning, making the woman’s face the most salient means for performing identity. This practice appears typical of selfies taken by young women, where close-up selfies often remove contextual cues and plain white walls are favoured as image backgrounds (Dobson 2015:38; Veum and Undrum 2017:95). As previously mentioned, the image has been edited via cropping. The bright lighting and saturation also suggest some level of enhancement or editing typical of the filtered images of ‘perfection’ that tend to be presented online (Åberg et al 2020:9); the image appears to have reduced sharpness, demonstrated by the woman’s extremely smooth skin and the clear contrast between light and shade. Despite potential edits, the image appears typical of what one may expect to see within increasingly normalised ‘more than real’ digital photographs (Ravelli and van Leeuwen 2018:284).

What are the sociocultural connotations of the person depicted in the profile picture?

Given that the profile picture has the hallmarks of typical photo sharing practices amongst young women, connotations of female identity are reinforced by the styling of the photo. More specifically, the pouting pose and use of makeup can be understood as reinforcing stereotyped, sexualised constructions of women dominant in contexts like advertising (Rosewarne 2007:82-84). Döring et al (2016:961) suggest that selfies posted on Instagram tend to reproduce beauty ideals and specifically, traditional gender stereotypes, to a larger extent than magazine adverts, describing certain visual features (such as the pouting pose in women) as ‘social-media-specific

gender expressions'. In fact, the results of the Reverse Image search support the claim that the profile picture has connotations of female beauty. On Pinterest and Tumblr, the image appeared in threads that were compiled to celebrate attractive women or 'bombshells', as evident in the threads' titles. The use of the image in an advertisement for hair pieces also suggests that the photo depicts an attractive woman with desirable hair. These connotations of female beauty contribute to the overall identification of @wokeluisa as a young, attractive woman.

What identity aspects are mobilised by the profile picture and how do these aspects interact with the identity aspects signalled by the handle/username?

The profile picture reinforces female identity aspects signalled in the account's handle and username. The picture further identifies the account in terms of race and contributes to the performance of a Black racial identity. Assuming a Black identity, the account draws on shared knowledge and beliefs about social media use and demographics. As previously acknowledged, "Black Twitter" is a term used to describe the strong Black presence on Twitter¹³ and challenges Black erasure from digital media studies, where users of colour have historically been cast as 'victims' with limited resources, rather than empowered social media users who set agendas of discussion, influence popular culture, and contribute to global debate (Florini 2014:225). As argued by Nakamura (2008), performances of Blackness online 'constitute an important mode of resistance to marginalization and erasure' and are viewed within Black Twitter as authenticating belonging to the online community. The profile picture evokes discourses that not only resist erasure of Black voices and faces, but also celebrate Black beauty.

The profile picture works together with other elements of profile metadata to create a more coherent identity performance. Although the name does not make explicit links to age or race, the term 'woke' is widely used to refer not only to those with liberal political views, but also to young people, specifically millennials (Caldera 2018:3). As a result, the display of the woman as being youthful builds on expectations created by the information revealed in the handle. Moreover, the origins of the term 'woke' within African American Vernacular English, combined with the popularisation of the term by Black Twitter and the Black Lives Matter movement, also creates an association between the term and Black culture (Yao 2018:439). Consequently, while the term is now used widely, with some arguing that it has been appropriated by white people (Hess 2016), there still exists an association between the term and

¹³ '26 percent of Black Americans use Twitter or another status update service, compared with 19 percent of whites (Pew 2009). According to a 2010 Edison Research and Arbitron study, although Black Americans make up only 12 to 13 percent of the U.S. population, they comprised 24 percent of the seventeen million Twitter users in the United States (Saint 2010).' (Florini 2014:225)

Black culture, including Black people. The use of an image of a Black woman therefore contributes to the authentication of a particular identity by performing identity aspects that match the cultural associations of the term used in the handle, thus contributing to the construction of a valid representation of a Black woman.

5.4.4 @SouthLoneStar

@SouthLoneStar used the same profile picture throughout the account's activation (see Figure 15). A reverse image search does not produce any results, indicating that the image does not appear in any other online contexts.



Figure 15: @SouthLoneStar Profile picture throughout activation

What is represented from the world around us?

The profile picture is a coloured photo featuring a person. The image is close-up and shows a smiling, young, white male with a beard, wearing a Stetson, and what can be seen of a checked shirt.

How is the participant presented?

The photo is presented with a limited amount of visible background, the details of which are blurred. The inclusion of the person's neck and hat suggest some distance between the subject and the lens. These factors contribute to the impression of the image as having been shot from a 'with photographer' perspective (Zappavigna 2016:278). The subject is looking directly at the camera and is smiling. Both the facial expression and gaze create a social connection between the viewer and the subject (Kress and van Leeuwen 2006:118). The photo appears crisp and there is no evidence of any identifiable filters.

The blurred background and use of culturally meaningful clothing such as the Stetson evoke the aesthetic norms associated with stock imagery. Stock images are very profitable for image banks, where millions of images are stored and sold to designers for use in advertising, public relations, and the media. Stock images are images that 'do not record anything but evoke an

idea or a feeling and can be used to add interest to a page' (Machin 2004:317). In other words, they are not photos that are taken of real-time events or record moments of history, but they are posed images by models that are used to symbolise certain activities, concepts, or moods. According to Machin (2004:320-323), stock images often blur backgrounds to decontextualise their settings, making them suitable for use in diverse scenarios. They tend to represent 'types' of people; such 'types' are demonstrated by distinct markers like clothing. For example, a woman wearing a hard hat is used to symbolise 'women in construction' (Machin 2004:322). In this case, @SouthLoneStar fits the defining principles of stock imagery based on the style and shot of the photograph; the high resolution of the image evokes connotations with professionalism, rather than the kinds of images shot and edited by social media users on their phones. The blurred background and the clearly identifiable cultural categorisation markers also suggest that it probably concerns an example of a stock image.

What are the sociocultural connotations of the person depicted in the profile picture?

The profile picture used by @SouthLoneStar has many connotations with the Texan culture of cowboys. Paying attention to cultural categorisation markers (Ledin and Machin 2018), the appearance of the Stetson and flannel shirt contributes to the representation of the person as a cowboy. This interpretation is reinforced by image bank searches of the term "cowboy" that resulted in several images similar to @SouthLoneStar's profile picture (i.e., predominantly white men in Stetsons and checked shirts). Cowboy culture has long been associated with the South of the US, and more specifically, Texas. The Stetson, in particular, has been described as a 'badge of the stereotypical Texan' (Texas State Historical Association 2019).

As well as the association with geographical location, cowboys have long been 'a model and myth of frontier masculinity' (Gibson 2016:733), with some arguing that cowboys have become the 'symbol of the American man' (Ford 2018:41). Although in the 'old west' an estimated one in four cowboys were African American (Billson 2020), popular culture overwhelmingly presents cowboys as being 'gun-toting, boot-wearing white men' (Manzoor 2013). As a result, the profile picture used by @SouthLoneStar capitalises on racial discourses that draw on stereotypical associations between 'traditional' Texan cowboys and white masculinity.

What identity aspects are mobilised by the profile picture and how do these aspects interact with the identity aspects signalled by the handles/usernames?

The above discussion demonstrates that @SouthLoneStar's profile picture further identifies the account in terms of gender (male) and race (white), beyond location (Texas) that is achieved through the handle and username. The depiction of @SouthLoneStar as a cowboy not only

evokes gender as part of the overall identity performance in the account, but a performance of gender that represents traditional ideas of American macho-ness embedded in local identity. Coherence in identity performance is achieved through the presentation of stereotypical Texan features in the profile photo as well as the references to Texas and the American South in the naming strategies, thus signalling identity aspects relating to location, local identity, and culture. This coherent identity portrayal adheres to societal expectations of what a Texan should be – or should look like, thus contributing to the construction of a valid representation of a conservative Texan man.

5.4.5 @USA_GunSlinger

@USA_GunSlinger had two profile pictures while the account was active: Figure 16 is the initial picture used and Figure 17 shows the picture used from June 2017 until the account was suspended.

A reverse image search for Figure 16 produces five results, three of which are directly linked to repositories for Internet Research Agency data. The other two are linked to websites that no longer exist: (i) www.pinsdaddy.com (which on closer investigation appears to be related to Pinterest) and (ii) www.thetattoo hut.com, where it appeared under a link entitled ‘tattoo-not-a-muslim’. In the case of the two inactive websites, the image did not include the basket.

From June 2017, the profile picture changed. A search of the new profile picture (Figure 17) resulted in four results, all of which were records of the picture’s function as @USA_GunSlinger’s profile picture recorded by data repositories for disinformation accounts.



Figure 16: @USA_GunSlinger profile picture (July 2016-June 2017)



Figure 17: @USA_GunSlinger profile picture from June 2017 onwards

What do the images represent from the world around us?

The first profile picture features a cartoon of the statue of liberty holding a gun, with bullets wrapped around its body and its hand positioned on its hip. The image is set against a backdrop of the American flag and the statue is positioned so that it is emerging from a basket. The second profile picture features a young white woman wearing a t-shirt featuring a similar image to the first profile picture: the statue of liberty holding a gun.

How are the visual elements and/or participants presented?

The first profile picture contains a coloured, cartoon depiction of various elements layered on top of one another: a basket is laid over the top of the lower part of the statue of liberty which is set against the backdrop of an American flag. In terms of colours, the statue is represented in a shade of green similar to the colour of the statue in “real” life and the American flag is displayed in highly saturated, flat, unmodulated shades of white, red, and blue. The design of the statue in Figure 16 gives the statue gaze, facial features, and expression that differ from the original figure as displayed in New York. The gaze of the cartoon, which is directed at the viewer, rather than gazing into the distance or at the gun itself, connects the anthropomorphic statue with the viewer. According to Kress and van Leeuwen (2006:118), the participant’s direct gaze (and the gesture, if present) ‘demands something from the viewer; it demands that the viewer enters into some kind of imaginary relation with him or her’. In instances where viewers are stared at by subjects with ‘cold disdain’, viewers are asked to relate to the subjects ‘perhaps as an inferior relates to a superior’ (Kress and van Leeuwen 2006:118). The statue is not illustrated as smiling, but simply with the lips apart, in what could arguably be described as a snarl. Finally, the positioning of the gun in the image also contributes to setting up a particular relationship between the viewer and subject. The act of aiming a gun into the air is often associated with celebratory gun fire, common during military conflicts to display victory over an enemy (Parada et al 2009:150). As a result, the non-smiling, cold stare of the statue, together with the military connotations of the aim and positioning of the gun, presents the statue in a position that symbolises power over a perceived enemy.

The second profile picture is also provided in colour; rather than being a cartoon, it is a photograph of a person. The photo appears to have been taken from a ‘with photographer’ perspective, as evident in the distance between the subject and the viewer, but also the subject’s stance. The subject appears to have their arms out, perhaps as though their hands would be on their hips, or simply splayed outwards. Either way, the positioning of the arms suggests that the photo was not taken by the subject themselves. The background of the image is blurred but brickwork and a window are still visible. Similar to @SouthLoneStar, the crisp, high resolution shot of the person is set against a generic, blurred background that decontextualises the image

and foregrounds the subject as the centre of the image. The subject of the photograph is smiling, tilting her head slightly to the right, and gazing directly into the camera. As discussed in relation to @SouthLoneStar, such poses create intimacy between the subject and the image viewer (Kress and van Leeuwen 2006:118), indicating that it may concern a posed photo. The display of the statue of liberty holding a gun on the T-shirt reveals that the profile picture itself contains two images: one of a smiling woman and the other of the emblem on the T-shirt. Displaying the emblem via clothing demonstrates a level of alignment between the T-shirt wearer and the image being displayed and, at the same time, an alignment between the profile owner, the T-shirt wearer, and the image displayed on the clothing. The use of this garment arguably also acts as explicit cultural categorisation; rather than drawing on conventional connotations of clothing elements (e.g., Stetson hat in the cowboy case), the T-shirt wearer, and consequently the profile owner, position themselves as supporters of the right to bear arms.

The two profile pictures therefore present similar messages relating to the right to bear arms but are presented in different ways: the somewhat combative stance and gaze of the first image is counteracted by the smiling pose of the subject in the second.

What are the sociocultural connotations of the images depicted in the profile pictures?

To understand what connotations are evoked by the profile picture, sociocultural associations with the statue as an artefact must first be acknowledged. The iconic image of the statue of liberty is formed by the figure of Libertas, the Roman goddess of liberty who holds a torch above her head in one hand and a Roman tablet with the date of American independence in the other. Next to the statue's feet lie broken shackles and chains, intended to commemorate America's abolition of slavery (National Park Service 2015). It is noted that the symbolism of anti-slavery associated with the shackles and chains in the original statue is not visible in the cartoon used by @USA_GunSlinger, where the statue's feet are not in view.

According to Berenson (2012:3), the statue of liberty 'doesn't even overtly refer to the United States'. Although the date of American independence is inscribed upon the statue, the date is barely visible and, as a result, the date and the statue hold different meanings to different groups of people around the world. Instead, the statue's 'abstractness, artistic banality, and colossal size' was intended to 'express a general, universal theme' (Berenson 2012:3). In this instance, however, the coupling of the statue of liberty with the American national flag, together with the addition of weaponry and the removal of anti-slavery connotations, represents a far more specific interpretation of liberty, relevant to a specific group of people: gun-carrying, white Americans. To some, freedom and liberty may be defined specifically by the absence of

weapons, but references to guns as symbols of freedom are found throughout discourses of pro-gun campaigners and the political right (Melzer 2009:233). The anthropomorphisation of the statue of liberty in this case, created by the positioning of hand on the hip (often associated with ‘sassiness’), direct gaze, and holding a gun in the air, can be used to position the statue as a symbol of power and freedom, with notions of freedom being directly associated with the American right to bear arms.

The connotations of the image of the gun-holding statue of liberty are carried over into the second profile picture.

For many people, guns are associated with violence, fear, and historically, masculinity (Braman and Kahlan 2006; Carlson 2015; Stroud 2016). Arguably, a smiling young woman sporting a T-shirt with pro-gun imagery subverts traditional expectations relating to who owns guns and supports the right to bear arms. Depictions of women in mainstream media have historically reinforced existing gender stereotypes of women as emotional, maternal caregivers who are less likely to be violent (Pennock-Speck 2016), and who are often defined by their appearance, ‘sex appeal’, and family relationships, rather than their achievements or professional credentials (Attenborough 2011:665). In the context of such stereotypical associations, women have been used to normalise and ‘de-demonize’ controversial political movements like the far right (Vieten 2016:626). It could therefore be argued that the pro-gun imagery in Figure 17 appears ‘softened’ by the presence of the young woman, her smiling pose, and associated gender stereotypes that deem women as ‘weak, passive and responsive’ in comparison to their ‘active, aggressive and spontaneous’ male counterparts (Creedon 1994:29). The portrayal of a woman showing pro-gun leanings may consequently add impact and increase visibility to the account due to the juxtaposition of what is stereotypically expected of women (softness, gentleness) with the aggression and masculinity associated with guns.

What identity aspects are mobilised by the profile pictures and how do these aspects interact with the identity aspects signalled by the handles/usernames?

Through the portrayal of the statue of liberty (a famous American landmark), the American flag, and the bearing of weapons, the first profile picture signals identity aspects relating to location (the USA) and topic of interest (guns). The profile picture builds on pre-existing pro-gun discourses and contributes to the construction of the profile as belonging to someone who believes that access to guns is both inherently American and libertarian. While it may be counter-intuitive to identify guns as a hobby or topic of interest, their diachronic place in American culture has normalised gun ownership in society: ‘gun culture is part of the broader

American culture’, evidenced by its characterisation as ‘a nation of gun clubs, training classes, shooting events, network meet-ups and gun collectors and shooters associations’ (Yamane 2017:7). The display of location and hobby in the profile picture also maps into the identity aspects presented in the username and handle: American GunSlinger @USA_GunSlinger.

The change to the second profile picture coincided with the username change from American GunSlinger to GunSlinger Girl. Consequently, gender became relevant both to the username and the profile picture at the same time, together with the ‘topic of interest’ which remained constant. The use of the second profile picture indexes identity aspects of the T-shirt wearer (gender and race) while aligning them with the pro-gun message evident through the symbol on the clothing. The use of ‘girl’ in the username also has connotations of female youth and the picture’s depiction of a smiling, young woman supports this interpretation. The use of the term ‘girl’ together with the image of a smiling woman reflects the practice of female users to ‘earn value in the virtual marketplace by appearing playful, sociable, and friendly’ (Herring and Stoerger 2014:576). Race appears to be the only identity aspect that is conveyed solely through visual means and the second profile picture.

Similar to the other accounts, the consistent representation of the same identity aspects across the naming units and the profile picture contributes to the authentication of the persona by creating a coherent identity performance as an American who supports the right to bear arms. The change of the profile picture results in the loss of any associations with a specific location as the American flag is no longer used and the statue itself has acquired a more global meaning (Berenson 2012). Nevertheless, the move from the first profile picture to the second arguably contributes to authentication processes by orienting to more established profile picture practices on Twitter, that is, for images to present a physical likeness with a person, thus making them more identifiable, less ‘anonymous’ and thus, less ‘suspicious’ (Barojan 2018; Calabrese 2020; Dotto and Cubbon 2019). While cartoons can be used to authentically depict life experiences in the context of, for example, autobiographical comics (El Refaie 2010), cartoons are evaluated as being less credible than photographs of humans within the context of social media profile pictures (Morris et al 2012:445). The use of a photo rather than a cartoon therefore appears to demonstrate an adoption of pre-existing dominant practices associated with identification on Twitter, thus contributing to the construction of a valid representation of a “real” pro-gun, American Twitter user.

Based on the analysis above, I have found the following labels that represent the identity aspects conveyed through profile pictures: (i) location, (ii) political affiliation, (iii) person, (iv) race,

(v) gender, (vi) hobbies/interests. The results of the labelling for the five accounts can be presented as follows:

Table 15: Identity aspects signalled in five accounts profile pictures

Account	Location	Political Affiliation	Person	Race	Gender	Hobbies/interests
@TEN_GOP August 2016-May 2017	✓	✓				
@TEN_GOP June 2017 onwards	✓					
@Pamela_Moore13		✓	✓	✓	✓	
@wokeluisa			✓	✓	✓	
@SouthLoneStar	✓		✓	✓	✓	
@USA_GunSlinger July 2016-June 2017	✓					✓
@USA_GunSlinger June 2017 onwards			✓	✓	✓	✓

Overall, the study of identification in profile pictures, together with the findings from the analysis of handles and usernames, has shown that identity aspects are presented in a coherent manner across the different metadata units. Not only do profile pictures and naming units tend to present the same identity aspects, but they present identity aspects in ways that reinforce normative expectations based on social stereotypes, ultimately contributing to the construction of valid representations of distinct “types” of American personas. Such consistent identity performances may contribute to the authentication of the identities performed and increase the credibility of the specific disinformation accounts on Twitter. The link between consistency and authenticity has been discussed in previous literature (section 2.3.1), especially in cases where consistency is achieved through adherence to normative expectations or as consistency within a person’s own identity performance over time or across platforms (Coupland 2001; Kytölä and Westinen 2015; Marwick 2013). To further investigate how disinformation accounts use profile pictures to perform identity, the analysis will be extended to a wider sample of other popular accounts (those with 10,000-100,000 followers) and comparatively less popular accounts (5,000-9,999 followers).

5.5 Identification through profile pictures: Using descriptive statistics to analyse the wider sample

Similar to the analysis of handles and usernames, I will use descriptive statistics to analyse the profile pictures used in the wider data sample in terms of the identity aspects presented within

them. This will situate the findings of the qualitative analysis within a wider data sample and enable findings to be contextualised in relation to wider practices associated with the use of profile pictures on Twitter disinformation accounts. The investigation of both popular and comparatively less popular accounts will make visible any differences between the ways that profile pictures are used by accounts with differing levels of success within the attention economy of Twitter, and thus enable a more comprehensive understanding of how profile pictures may contribute to authentication processes.

Based on the visual analysis of the five main accounts, I coded all profile pictures in both data samples (more and less popular) in terms of the identity aspects indexed or represented through these visual means. Where more than one identity aspect is present (e.g., topic of interest and location), both have been coded. The analysis draws on the labels that emerge from the five key accounts, as well as others that were noted in the wider sample. As a result, beyond the categories introduced in section 5.4, I have also included the following codes: animal, religion, organisation, news, and inanimate object. In line with the method employed in the analysis of naming strategies, ‘news’ was used to label profile pictures that contained words associated with news organisations e.g. ‘Today’, ‘Post’, ‘Daily’, ‘Online’ etc.

It is necessary to acknowledge that unlike the five main accounts, the wider sample also includes some profile pictures that are word-based. For example, Figure 18 represents a word-based profile picture, whereas Figure 19 shows an image-based profile picture. Despite their difference, both pictures have been coded as indexing location. Figure 18 signals location by displaying ‘Miami’ via text, while Figure 19 signals location (Syria) visually through the image of the national flag. Figure 18 has therefore been categorised as signalling both location and news, while Figure 19 has been coded as signalling location.



Figure 18: Word-based profile picture indexing location



Figure 19: Image-based profile picture indexing location

Finally, the analysis focuses on the extent to which the identity aspects conveyed in profile pictures are consistent with the identity aspects signalled through the usernames and handles. The following two sections will therefore explain how a) identity aspects are presented in profile pictures and b) how these contribute to identity performances across metadata units.

5.5.1 Profile pictures and identity in more and less popular accounts

The five accounts focused on throughout the thesis have been removed from the wider sample to gain a sample that can help to contextualise the findings of the qualitative analysis. The percentages in Table 16 and Table 17 demonstrate what proportion of the disinformation accounts in each sample display a particular identity aspect. Due to numerous accounts displaying more than one identity aspect (cf. section 5.4), the percentages do not add up to 100%. The distribution of identity aspects presented in profile pictures in the more popular sample will be presented first, followed by those in the less popular sample.

Table 16 shows that location is the most frequently presented identity aspect in profile pictures in the sample of more popular accounts. Similar to patterns of identity presentation through naming strategies, location co-occurs with news-related language in a number of accounts; of the 36 profile pictures indexing location, 27 also include identity aspects associated with news branding.

Table 16: Identity aspects presented in profile pictures in more popular dataset

Identity aspect in profile picture	Number of instances (n)	Proportion of sample representing each identity aspect (%)
Location	36	62%
News	27	47%
Person/people	14	24%
Race	15	26%
Gender	14	24%
Topic of interest	5	8%
Political affiliation	3	5%
Animal	2	3%
Inanimate object	1	2%
Religion	0	0%
Organisation	0	0%

With respect to the news-branded accounts, based on shared visual features, profile pictures indicate similar sub-groups to the ones identified through the type of handles and usernames;

today (Figure 20 – five accounts) online (Figure 21 – nine accounts), daily1 (Figure 22 – four accounts), daily2 (Figure 23 – three accounts) and top news (Figure 24 – four accounts). The accounts in each group all display the same profile picture design. For example, in the ‘today’ group, all profile pictures show a coloured circle with the first letter of the city displayed in the middle and the phrase ‘[location] today’ down the side (Figure 20). An additional group can be added to those observed within the naming analysis: there are two sets of account that use the term ‘daily’ but have observable differences in terms of their visual presentation- in the daily1 group, all accounts are set on a background of a coloured photograph that indexes aspects related to the location indicated in the name, while in the ‘daily2’ group, all accounts show part of a photograph of a location set at an angle, with the location and ‘daily news’ written over the photo and the white background.

The impact that these groupings may have on authentication processes will be discussed in more detail later.



Figure 20: The 'today' group



Figure 21: The 'online' group



Figure 22: The 'daily1' group



Figure 23: The 'daily2' group



Figure 24: The 'top news' group

All pictures of people result in a default identification of the account with particular types of race and gender, while one picture signals race without gender via the words ‘Black to Live’. As acknowledged above, in cases where race and gender are presented, I further categorised what race/what gender was being indexed. The distributions of different genders and races will be discussed in more detail later. As evident in Table 16, while the majority of more popular

accounts use profile pictures that index location and news, pictures depicting people are also present.

In the data set of less popular accounts, profile pictures primarily identify the accounts with people (see Table 17). Similarly to the other sample, the use of a picture of a person also identifies the account in terms of gender and race. The only exceptions are two spoof accounts where in the first case, race is presented without gender (@Jeblary2016¹⁴), and in the second, gender is presented without race (@Jihadist2ndWife¹⁵). The impact that such accounts may have on authentication processes will be discussed in section 5.6.1.

Table 17: Identity aspects presented in profile pictures in less popular dataset

Identity aspect in profile picture	Number of instances	Proportion of sample representing each identity sample (%)
Location	7	30%
News	3	13%
Person/people	15	65%
Race	14	61%
Gender	14	61%
Topic of interest	1	4%
Political affiliation	2	9%
Animal	0	0%
Inanimate object	0	0%
Religion	1	4%
Organisation	1	4%

In contrast with the more popular account sample, only two of the accounts include both news-related and locational information. In the less popular account sample, then, there is a tendency for profile pictures to index more personally identifying identity aspects through pictures of humans, as opposed to less personally identifying ones indexed through collective identification e.g., geographical location.

Across all samples (including the five main accounts), there are performances of both male and female identities, as well as white and Black identities. While there are of course various other genders and races outside of these binary categorisations, these are the identities made

¹⁴ The profile picture showed a photoshopped image of Jeb Bush and Hillary Clinton merged as one person. Consequently, it was not possible to label gender, but both of these people are white, so race was still signalled.

¹⁵ The profile picture showed a person wearing clothing associated with Muslim women (Piela 2013).

relevant within the data and in previous research into IRA accounts (Arif et al 2018; Freelon et al 2022:572). It is important to highlight that there are some cases, such as in @Pamela_Moore13, where the use of filters can make racial identification slightly more difficult. Despite some levels of ambiguity in these cases, race has still been coded. This decision was in line with Freelon et al's (2022) suggestion that IRA accounts can be categorised as Black and non-Black. They emphasise the role that 'digital blackface'¹⁶ played in IRA identity performances, where photographic filters can be used to make people 'appear "Blacker"' (Freelon et al 2022:563). It therefore appears necessary to acknowledge that filters can also be used to make people appear lighter or, more specifically, whiter. Consequently, it is possible to make categorisations on how the accounts appear to be presenting themselves within the racial asymmetries of Black and white established by Freelon et al (2022).

Overall, there is a slight dominance of female identities in the general sample¹⁷. In terms of race, the more popular account sample reflect an even distribution of identification with both white and Black personas, while most personas in the less popular sample are presented as white (10).

5.5.2 Identity performance across metadata units in the more popular sample

To understand how metadata may contribute to the performance and authentication of particular personas, it is necessary to investigate the consistency (or discrepancy) across the identity aspects conveyed through the three metadata units under investigation. This section will focus, specifically, on the more popular sample. Firstly, the distribution of identity aspects within both naming units and profile pictures will be examined. Then, the relationship between eponymous/pseudonymous names with corresponding profile pictures will be explored. Afterwards, the section will analyse presentations of race and gender across metadata units, followed by a further examination of accounts that draw on location and news branding. Section 5.5.3 will investigate the same areas in relation to the less popular sample.

Table 16 in section 5.5 shows that in the more popular accounts, location is the most frequently signalled identity aspect in profile pictures. This pattern is consistent with the findings of the analysis of naming strategies: in the more popular accounts, location is the most

¹⁶ Freelon et al (2022:563) define this term as 'digitally mediated renderings of online selves as Black by non-Black individuals (Dobson & Knezevic, 2018; Green, 2006; Robertson et al., 2018; Stark, 2018).'

¹⁷ In the more popular sample, there are 8 female presentations and 5 male. In the less popular sample, there are 7 female and 6 male.

commonly signalled identity aspect in both handles and usernames. In terms of other identity aspects, there are also similarities between naming strategies and profile pictures: representations of individual persons via profile pictures are the second most frequent identity aspect as are proper names within naming strategies. Similarly, organisations represent the least frequently signalled identity aspect in both profile pictures and naming strategies.

All accounts that use eponymous naming strategies also use profile pictures that depict a person/people. Depictions of people are done through photographs of unknown individuals (i.e., not famous) (11 cases) or through cartoon-style images (3 cases). These cartoon-style images are (i) a Bitmoji cartoon i.e., a type of cartoon avatar that can be personalised in terms of skin, hair and eye colour, clothes, hairstyle, and hobbies; Bitmojis have previously been used successfully to make people recognisable in the offline sphere, suggesting there is a societal expectation for Bitmojis to show physical likeness (Kelly et al 2021:8), (ii) an illustrated, cartoon-style image of an unknown woman, and (iii) a cartoon image of Donald Trump being punched in the face by Batman.

There are only two instances where a pseudonymous name and handle are combined with a profile picture showing a person (@gloed_up 1-800-WOKE-AF and @BleepThePolice Bleep The Police¹⁸). In both cases, the profile pictures depict Black men, identifying the accounts in terms of race and gender in addition to personal attributes /topics of interest indexed within the naming strategies. While neither the usernames nor the handles explicitly indicate any racial aspects, the terms ‘woke’ and ‘gloed up’ in the first instance originate in African American Vernacular English (Bonnette 2015:3; Nembhard 2020; Ramos 2020; Yao 2018). Similarly, ‘bleep the police’ appears to draw on the phrase “fuck tha police”, that surged in popularity following hip-hop group NWA’s single, Straight Outta Compton. The phrase became iconic within discussions of racism in the police, ‘exemplifying discourses both resisting and justifying police violence in poor Black neighbourhoods’ (Edgar 2016:223). While anti-police sentiment is not exclusive to particular racial groups, the presentation of Black men alongside a phrase that evokes anti-police feeling appears to draw on existing discourses linked to anti-Black racism, protest music, and the police in the US. Similar to some of the five main accounts that were qualitatively analysed (@SouthLoneStar, @USA_GunSlinger, @wokeluisa), it appears that the accounts use linguistic terms associated with particular communities, together with images that depict people from these groups, to construct coherent identity performances

¹⁸ This is the only instance in the data set of a group of people being used rather than an individual. The profile picture shows three Black men sitting on a car and has been decontextualised by the removal of a visible background.

that can be taken as valid representations of particular American personas based on their orientation to pre-existing social stereotypes.

In all cases where gender has been coded in handle/username and profile picture, the same type of gender identification appears constant across all metadata units. For example, in the case of @Jenn_Abrams (Jenna Abrams), the handle and username both index female gender and the profile picture also includes a photograph of a woman.

As mentioned in chapter 4, names can also index race (Gaddis 2017; Sweeney 2013). I analysed profile pictures to further investigate the distribution of Black and white identity performances. I find that in all cases but one, where first names index racial distinction, the profile pictures show people of the same race. For example, the account @TheFoundingSon John Davies uses a username with a white racial distinction (the name John resulted in images of white men on Google Images) and uses a profile picture of a white man. Similarly, the account @TrayneshaCole Traynesha Cole uses a name with a Black racial distinction (the name resulted in images of Black women on Google Images) and uses a profile picture of a Black woman. The only case where this did not take place, was where @KaniJJackson Kanisha Jackson used a profile picture that showed a cartoon of Donald Trump being punched in the face by Batman. While the name indexed a Black racial distinction (the name resulted in images of predominantly Black women on Google Images) and the image does not show a Black person, the use of an image that signals disaffiliation with the socio-political right could be understood as a means of evoking race, particularly when considered alongside findings that within IRA accounts, racial asymmetries existed alongside ideological ones, demonstrated via the distinction between ‘Black left wing activists’ and ‘white far-right Americans’ (Freelon et al 2022:572; Arif et al 2018). It therefore appears that accounts tend to present information that can be interpreted as coherent and consistent across metadata units by reinforcing stereotypical expectations relating to names, race, and gender.

In addition to consistency in terms of race and gender, locational identification is also found to be conveyed in a similar manner across the three metadata units under study. Of the 36 profile pictures indexing location, 31 present the same identity aspect and encode the **same** location in their username and/or handle. For example, @TodayMiami indexes the location of Miami in the handle (@TodayMiami), username (Miami Today), and the profile picture (see Figure 20). In the other five instances, location is presented in profile pictures, but not in the username or handle. In these cases, the additional identity aspect of location contributes information that, in conjunction with other metadata units, forms a coherent identity performance. For example, the profile picture used by @NatPolNews (username: National Politics) contains the seal of the

United States which identifies the account in terms of location and signals to the viewer which nation the username and handle refer to.

Similar to naming strategies, location also co-occurs with news branding in many profile pictures. As displayed in Table 16, 27 accounts coded for location also have profile pictures that signal news branding. Of these 27 accounts, all of them have profile pictures that include repetition of their usernames (see Figure 25).



Figure 25: Profile picture that replicates username, St. Louis Online

It therefore appears that within the more popular sample, accounts tend to present identity aspects consistently across metadata units, regardless of the kind of naming strategy used (eponymous or pseudonymous), or the type of account being constructed (news or personal).

5.5.3 Identity performance across metadata units in the less popular sample

This section will focus, specifically, on the less popular sample, concentrating initially on the distribution of identity aspects within naming units and profile pictures, and then on the relationship between these metadata units. Afterwards, presentations of race and gender will be examined, followed by further discussion of accounts that draw on location and news-related branding.

Table 17 shows that in the less popular accounts, pictures that showed people (and thus, race and gender) were the most frequently employed. This pattern is consistent with the findings of the analysis of naming strategies: proper names (associated with individual persons) signalling gender were most common in both handles and usernames. Although political affiliation was the second most frequent identity aspect in handles and usernames, location is more commonly signalled within profile pictures, and political affiliation is one of the least signalled identity aspects.

As displayed in Table 17, the most frequent type of profile picture is one that depicts a person. Of the 15 accounts using profile pictures of people, 12 also have eponymous usernames and/or handles, indicating that more personally identifying identity aspects are conveyed across all three metadata units. In this sample, we also find pictures of people that belong to celebrities or popular culture figures as in the spoof accounts mentioned before (@Jeblary2016 and @Jihadist2ndWife) and in: @JohnCopper16 Marlboro Man where the picture depicts the so-called Marlboro Man who featured in cigarette ads from the 1950s to the 1990s and has been described as embodying traditional American masculinity (Martin and Gnoth 2009); in @LaChristie Resistance Girl where the picture is a piece of artwork that depicts Angela Davis, a prominent political activist associated with the Black panther movement; and in @March_for_Trump March for Trump that shows Donald Trump. Such cases are very rare in the more popular account sample, where, as mentioned in section 5.5.1, only one account includes a cartoon image of Donald Trump. Within the less popular sample then, it appears that pseudonymous names are broadly associated with images that are not people, while eponymous names tend to be presented alongside images of people. However, within the less popular sample, the kinds of people displayed in these images appear markedly different to those displayed in the more popular sample, mostly due to the stronger presence of ‘spoof’ or celebrity imagery.

As mentioned in section 5.5.1, there is a less even distribution of Black and white identities within the less popular accounts than the more popular ones. Of the seven presentations of white personas that use eponymous handles and/or usernames, all of them are used alongside names established as predominantly white in chapter 4. For example, the name Amy returns images of white women on Google, is categorised as a white name by Gaddis (2017) and in my sample, Amy Green @AmandaVGreen uses a profile picture of a white woman. The three profile pictures that show Black identities, were all used alongside names that retrieve results with no racial dominance (@IlikeBIGbuttand Messiah Haynes, @Booth_Prince Prince Booth, @LaChristie Resistance Girl). In terms of gender, all profile pictures were used alongside names that correspond with the gender category established in chapter 4.

In terms of location and news branding, there are fewer accounts indexing these identity aspects than in the more popular sample. Compared to the more popular accounts, location is no more the most frequent identity aspect encoded, and the relationship between location and news branding across metadata units is not as overt. Of the seven profile pictures signalling location, five have usernames and handles that signal political affiliation, indicating that the co-occurrence of location and political affiliation is more frequent compared to the coupling of

location and news branding. Moreover, there is only one instance where a news-branded account can be placed into one of the groups discussed in section 5.5.1; most of the accounts appear to stand alone. It therefore appears that news-branded, location-based accounts that share metadata features with other news-branded, location-based accounts are more popular than those that do not.

Overall, then, there are some notable differences between the two samples in terms of the identification information mobilised across the metadata units. The less popular accounts orient more towards identifications with individual persons, rather than news organisations. They also include instances where the accounts openly present themselves as not belonging to “real” people that exist offline (e.g., Jeblary Bushton or the Marlboro man). Finally, any news-branded accounts do not follow the same patterns as those documented in the more popular data sample: there are fewer, they cannot be placed into groups that share branding, and they have different relationships with location and political affiliation.

5.6 Discussion

The results from the analysis of identity presentations in profile pictures demonstrate that in both samples, disinformation accounts primarily take the form of the two types of accounts presented in chapter 4: news accounts that use location and news-branded imagery, and personal accounts that show people. This section will firstly discuss how profile pictures may have contributed to authentication processes in personal accounts, and secondly, discuss how profile pictures may have contributed to authentication processes in news accounts. This discussion pays particular attention to the potential relationship between authentication processes, adoption of pre-existing social media practices, and orientation to existing stereotypes.

5.6.1 Personal accounts

A key strategy in the construction of a personal account appears to be the simultaneous use of eponyms in handles and usernames and of photos of people in the profile pictures. It has been argued that eponymous naming strategies can authenticate online identity performances by making a person identifiable and thus evoking familiarity with the ‘real’ or offline world (Jäkälä and Berki 2013:6). Similarly, this can arguably be achieved visually by using photos of people, as a photo of a person has identification power in the offline world. As demonstrated within both the analysis of the five accounts and the use of descriptive statistics to interrogate the wider sample, the coupling of photos of individuals with pseudonymous naming strategies can also contribute to the construction of individuals as less anonymous and more identifiable. While

this view of authenticity can be considered essentialist (i.e., that pictures of people contribute to authentication processes by creating a link with the so-called “real” world), it is necessary to acknowledge that this view of authenticity appears relatively established within the context of Twitter.

On Twitter, profile pictures that show photos of individuals appear to be favoured over others (e.g., cartoon characters, animals, landscape, the default Twitter profile picture) that are more readily associated with bad actors online and thus tend to be associated with lower levels of credibility (Barojan 2018; Dotto and Cubbon 2019; Morris et al 2012). Consequently, on Twitter particularly, anonymity tends to be perceived as a means of acting nefariously and without accountability (Calabrese 2020; Lester 2021; Knowles 2022). As a result, accounts that appear less anonymous (e.g., by using photos of human faces) may be more likely to be taken as valid representations of Twitter users, not solely because of the somewhat essentialist ‘notion of familiarity’ (Jäkälä and Berki 2013) that such images may evoke with an offline person, but the sense of familiarity that such identification practices may evoke with what users typically see on Twitter.

As well as an orientation to established identification practices, identity performances within personal accounts demonstrate an orientation to existing social stereotypes. Identity performances as American personas can be constructed either by the use of eponymous names associated with particular racial and gendered identities, or by pseudonymous names that make intertextual references to particular communities. In both cases, the disinformation accounts tend to use profile pictures that reinforce the identity aspects signalled within the handles and usernames, thus contributing to the construction of valid representations of particular types of American personas. Previous research suggests that IRA disinformation accounts drew on stereotypical associations and symbols related to Black culture to engage in digital Blackface, the ‘digitally mediated renderings of online selves as Black by non-Black individuals’ (Freelon et al 2022:563). This analysis elaborates on these findings and reveals that digital Blackface was constructed using both naming strategies and profile pictures. Similarly, representations of white personas, and gendered identities were constructed by exploiting stereotypes related to names and particular types of imagery.

The analysis also suggests that identity performances may orient to the potential attention and engagement particular kinds of identity performances may attract. This appears particularly relevant to the way that images of women were used in profile pictures. Images employed of women tend to be of young and attractive women (Linvill and Warren 2018:7), reflecting established standards of attractiveness and in some cases, showing a degree of sexualisation (cf.

topless picture of @Pamela_Moore13). As acknowledged in section 2.3.2, online presentations of women that fulfil conventional standards of attractiveness and sexualisation result in increased levels of audience attention and engagement (Drenten et al 2019). In a context where popularity is associated with attractiveness, sexualised presentations of conventionally attractive women may therefore contribute to the construction of valid representations of popular online female personas.

Within the less popular sample, there appeared to be a less overt orientation to established Twitter practices, particularly in relation to identifiability. While the more popular accounts demonstrate an orientation to social stereotypes to some degree, within the less popular sample, some accounts exploited social stereotypes to the extent that they created spoofs or parodies; they presented identity performances that framed their Twitter content ‘within the context of fictional universes or stereotypes of their subjects’ (Highfield 2016:2029) (e.g., ‘Jihadist’ Muslim women and traditional, masculine ‘Marlboro’ men). Freelon et al (2022:562) suggest that parody accounts are ‘clearly labelled as inauthentic’ due to their characterisation as fiction. It therefore appears that accounts in the more popular sample generally presented identity aspects that appeared to orient more overtly to established Twitter practices linked to personal accounts, using profile pictures of humans in ways that not only made accounts appear more identifiable, but drew on stereotypical expectations about race and gender to an extent that could be considered valid, but not overtly fictional.

5.6.2 News accounts

The appropriation of pre-existing norms as part of the authentication process of disinformation accounts is also evident in accounts that use news branding. As discussed in section 5.5.2, most news-branded accounts used profile pictures that repeated their usernames. These types of profile pictures draw on pre-existing norms of news sharing accounts, particularly the use of logos when representing a news broadcaster. The similarities between verified Twitter news profiles and ‘news’ disinformation is evident below, by comparing the verified BBC News Twitter profile picture with that used by @StLouisOnline.



Figure 26: Profile picture used by verified BBC News Twitter account



Figure 27: Profile picture used by @StLouisOnline from the 'more popular' data sample of disinformation accounts

Figure 27 illustrates that the profile picture mirrors the practices of verified news accounts, orients to the genre of fact-based pages and, consequently could be taken as a valid representation of an online news sharing account.

The appropriation of practices of verified news accounts is also evident in the practice of news corporations to have different Twitter accounts for different cities. As evident in Figure 28 and Figure 29, the different accounts that belong to CBS, an American commercial broadcast television and radio network, share similar profile pictures, with the same blue background and the same symbol. The only difference between the profile pictures is the locations stated in the text. Their handles also follow similar structures: CBS (news corporation) + location.



Figure 28: Verified CBS News account, @CBSChicago



Figure 29: Verified CBS account, @CBSNewYork

In the more popular data set, these same types of shared features can be seen in both profile pictures and naming strategies, demonstrated via the categorisation of news accounts into groups that shared identification strategies (see Table 14 and figures Figure 20-Figure 24). Figure 30, Figure 31, and Figure 32 exemplify that within the 'online' group, accounts have the same logo-style profile picture, using the same colour scheme and layout along with the same

construction of handle and username, (location + online). It therefore appears that news-branded, location-based accounts appropriate metadata practices of pre-existing, established news-sharing pages. They a) repeat their usernames in their profile pictures, creating a logo style profile picture and b) construct multiple accounts sharing the same features to suggest a network of news accounts rather than standalone profiles.



*Figure 30: @MinneapolisON,
Minneapolis Online*



*Figure 31: @NewOrleansON,
New Orleans Online*



*Figure 32: @PhiladelphiaON,
Philadelphia Online*

According to previous research, “breaking news” accounts (i.e., accounts that brand themselves as news sharing e.g., through their naming strategies or bios) are influential in propelling the spread of false rumours and they ‘gain the trust of Twitter users by mimicking the appearance of legitimate sources of information, often mirroring the look of a trusted media or official account’ (Andrews et al 2016:462). The findings from my analysis can be used to elaborate on this suggestion by demonstrating **how** accounts ‘mimic’ pre-existing news accounts via their profile pictures: by using logo-style images, accounts create a sense of familiarity (cf. Jäkälä and Berki 2013) with what users ordinarily see on Twitter, thus constructing representations of news accounts that can be taken as valid and therefore authentic. In contrast, news-branded accounts in the less popular data sample presented identity aspects differently: they did not appear in a network of accounts that shared the same logo, and their naming strategies appear not to orient to the kind of news-branding shown in profile pictures. Consequently, it could be argued that this deviation from the norms of the news genre results in representations that can be considered less valid and less authentic, potentially explaining their less popular status in the sample.

5.7 Conclusion

The results from the analysis of profile pictures suggest that popular disinformation accounts use profile pictures for identification in ways that evoke notions of familiarity both with what

users ordinarily see on Twitter, and with pre-existing social stereotypes. In personal accounts, images of humans are predominantly used in ways that make accounts more personally identifiable, while news accounts draw on location and news branding by using logos that appropriate established identification practices of news accounts online. In comparison, accounts with a lower number of followers oriented to established social media norms and practices less consistently. Previous research demonstrates that on social media, performances of authenticity are situated within specific contexts, where representations are taken as valid based on how closely they orient to practices established as normative within a particular community (Marwick 2013; Whitehead 2015). I therefore argue that there is not a specific set of visual markers, nor a certain type of profile picture that contributes to performances of authenticity on disinformation accounts. Instead, the accounts can be taken as valid representations of “real” news organisations or people by fulfilling expectations related both to how different types of accounts typically identify themselves on Twitter, and how broad social stereotypes can be associated with particular identity aspects.

On Twitter specifically, it has been suggested that users make credibility judgements of accounts based on the information available to them ‘at-a-glance’, e.g., their handle, username, profile picture, and tweet content (Morris et al 2012:444). I have shown that disinformation accounts present identity aspects in their handles, usernames, and profile pictures in ways that contribute to the construction of valid representations of different types of Twitter accounts and personas. To further understand how disinformation accounts authenticated particular personas on Twitter, it is therefore necessary to extend exploration to the other information that users are exposed to when they see a tweet, i.e., tweet content.

Chapter 6 Authenticating Identity Performance on Twitter

Disinformation Accounts: Co-telling News Stories

6.1 Introduction

Chapters 4 and 5 examined how the five key disinformation accounts create different ‘types’ of personas using options provided by the site of Twitter, such as naming strategies and profile pictures. While these chapters provide an insight into the ways in which resources facilitated by the platform are deployed in authenticating particular personas and identities, it is also necessary to look at content produced **on** the platform. This chapter shifts the focus to narrative discourse and authentication of particular identities in the disinformation accounts’ tweets, as narratives are the ‘prime vehicle’ for ‘expressing and negotiating individual and collective identities’ (De Fina 2015:351).

On Twitter specifically, users can signal ideological and political positionings when engaging in the storytelling of specific events (Georgakopoulou 2014:526). According to De Fina (2015:365), relations between tellers and stories are influenced by the many options available to tellers to identify themselves online. De Fina (2015:365) draws on Jäkälä and Berki’s (2013) categories of eponymity, nonymity, and polynymity, as explored in chapter 4 of this thesis, to demonstrate that identities are ‘always embedded and conveyed within specific storytelling practices’ (De Fina 2015:365). It is therefore necessary to extend the study of identity performance facilitated by the site of Twitter, as discussed in chapters 4 and 5, by acknowledging how the accounts perform identity when telling stories. This chapter aims to explore how each of the five main accounts narrates the events related to the Charlottesville Unite the Right rally and how this narrative discourse contributes to authenticating the personas constructed in the specific accounts. The analysis focusses on Charlottesville due to the prominent discussion of the event as it unfolded on Twitter.

Firstly, I will discuss why storytelling is an appropriate area to investigate when aiming to understand authenticity. Then, I will present the different parts of the story of Charlottesville that are made relevant within the data analysed in this study. Following that, I will discuss how Georgakopoulou’s (2015) small stories heuristic will be used in this analysis. Then, guided by Georgakopoulou’s (2015) framework, I will analyse the interrelated levels of sites, tellers, and ways of telling within the disinformation accounts’ storytelling of the Charlottesville Unite the Right rally.

6.2 Storytelling and authenticity on social media

Stories have long been acknowledged as providing insights into an individual's interpretation of an event or experience, with people using stories to 'make sense of themselves and the surrounding world' (Page 2018:1). On social media, storytelling is pervasive; people share stories of personal experience, comment on other people's stories, and engage with breaking news stories every day (Dayter 2015; Georgakopoulou 2014, 2017). Given that stories told online are characterised by self-disclosure, emotional, and intimate talk, online storytelling 'seems to be intimately connected with "authentic" identity performances' (De Fina 2015:364) and the study of online storytelling enables a 'deeper understanding of identity processes' (De Fina 2015:365). The way that people tell stories online also has the potential to unite people or drive them apart into polarised, opposing groups (Page 2018:1). What part of a story is told, or omitted, how events are described, and who people engage with when telling stories can all contribute meaningfully to a user's identity performance as a member of a particular social group.

Georgakopoulou's (2014) study into the online circulation of a news story involving MPs illustrates how users' assessments of events and characters can signal ideological or political positionings. By positively evaluating a character's actions, for example, users position themselves as affiliated with the character and their politics (Georgakopoulou 2014:526). The way that users position themselves in relation to topics and characters has also been linked with performances of authenticity online. As discussed in Chapter 2, Maragh (2018:604) investigates performances of racial authenticity on Twitter and finds that Black users avoid certain topics and adopt stances that align with their peers in order to avoid "acting white" and instead to be perceived as authentically Black by other users. In other words, performances of authenticity appear to be influenced by the degree to which users adhere to positions and stances that follow normative expectations associated with their social group. As shown in chapters 4 and 5, metadata units can be strategically employed to identify accounts with aspects of identity related to macro-social factors like age, gender, race, and national identity. These kinds of identity aspects, also likened to 'extrasituational context' by Page (2012:14), arguably contribute to the formulation of normative expectations associated with different social groups. In the context of Maragh's (2018) research for example, it could be argued that the presentation of a Black identity via profile pictures and/or naming strategies may lead to the adoption and avoidance of certain topics that would sustain an identity performance that can be interpreted as authentic on Black Twitter.

The analysis of the disinformation accounts' storying of a particular event will enable an understanding of their positioning in relation to the people and social groups involved. By adopting concepts and heuristics from the broader field of narrative analysis, the chapter aims to explore a) how the accounts perform identity when telling the story of Charlottesville, b) how these identity performances relate to identity aspects signalled in the metadata, and c) how the storying of Charlottesville may thus contribute to the authentication of identity performances as certain types of personas.

6.3 The story: The Unite the Right Rally, Charlottesville

On 12th August 2017, the Unite the Right rally took place in Charlottesville, Virginia. The rally was being held as a protest in response to the suggestion that Confederate monuments should be removed from public spaces after a white supremacist murdered nine Black churchgoers in a racially motivated attack in 2015. In Charlottesville specifically, it was proposed that the statue of Confederate general, Robert E. Lee, should be removed from a park. The rally attracted hundreds of far-right protesters, including neo-Nazis, white nationalists, and Ku Klux Klan members, in 'one of the largest white supremacist rallies in the US in decades' (BBC News 2019). The rally also drew large crowds of counter-protesters. During the rally, a car was purposely driven into a crowd of counter-protesters, killing one person, and injuring 35 others. The driver, James Alex Fields, Jr., was a neo-Nazi and white supremacist, and was convicted on numerous hate crime charges, including murder.

The Unite the Right Rally at Charlottesville has been chosen as a focal event for the current analysis, partly because it was one of the events that happened while all five disinformation accounts were active, and partly because it was widely discussed on Twitter as it unfolded. Previous research suggests that as live news events unfold online, it is often hard for information to be verified; sometimes, entire news stories can unfold on Twitter before any facts have been verified by other sources, leading to uncertainty and the dissemination of inaccurate information (Andrews et al 2016; Starbird et al 2014; Vis 2013). For disinformation accounts whose aim is to create chaos and exacerbate social divisions through the spread of divisive and inaccurate information (Horton 2020; Linvill and Warren 2020), such events that are associated with uncertainty and the spreading of inaccurate information are particularly useful. For that reason, the Charlottesville story was selected as a focal area to research when trying to gain an understanding of how disinformation accounts performed identity.

Another reason for choosing the specific event for closer investigation is that the Charlottesville attack is representative of the type of events that appear in disinformation tweets

that attract attention. Among the five accounts, popular tweets focus on topics relating to acts of violence and criminality. For example, the list of top ten most retweeted tweets of @TEN_GOP included references to a mass shooting in Cleveland, a video of a child being physically attacked, and the murder of a GOP committeeman. Such patterns of engagement can be explained by the tendency for stories identified as newsworthy to contain elements of conflict and violence (Bennett 1996; McManus 1994; Price et al 1997). Within disinformation accounts more specifically, Linvill and Warren (2020) found that IRA accounts posted divisive, partisan material and, as a result, many could be categorised as being ‘right’ or ‘left’ trolls. The events at Charlottesville were clearly linked to distinct groups with different political ideologies, making the study of storytelling of this event particularly important when aiming to understand how collective identities were performed.

To analyse the way that the accounts told the overarching story of Charlottesville, it is necessary to acknowledge not only the event itself, but also other events that preceded or followed what happened on the 12th August 2017, such as the background and context, as well as the aftermath and relevant subsequent events.

6.3.1 Background and context

As mentioned above, the removal of the Robert E. Lee statue in Charlottesville was prompted by a white supremacist terror attack where nine Black churchgoers were killed. To understand why and how these events are linked, it is necessary to acknowledge the history and contention associated with Confederate statues more broadly. As mentioned in Chapter 5, confederate imagery has long been the centre of contentious public debate due to its inherent connection with the Civil War, and thus, slavery. According to Graham (2016), many Confederate statues in the US were built during periods of history defined by racial division (e.g., the Jim Crow era, and the Civil Rights movement in the 1950s and 60s). As a result, it has been argued that the statues were built to ‘perform cultural resistance to black equality’ (Graham 2016), rather than to symbolise a preservation of American heritage as suggested by Trump and others in favour of maintaining Confederate statues (BBC News 2017b, 2020b). As demonstrated by the Unite the **Right** rally, the desire to protect Confederate statues has been associated with the socio-political right. It thus appears that the Unite the Right rally is embedded within debates surrounding politics, white supremacy, and racism.

6.3.2 The aftermath

Not only did the event result in continuing debates surrounding Confederate statues, racism, and right-wing terrorism in the US, but Trump's response to the event received extensive backlash after he suggested that there was blame 'on many sides' for the violence (Mcauliffe 2017). Trump's response was widely discussed both in mainstream and social media in the weeks following Charlottesville, evidenced in Trump's criticism of the 'dishonest media's' reporting of Charlottesville at a rally in Phoenix on 23rd August 2017 (Smith 2017). The prominence of Charlottesville as a point of discussion in public discourse led to the event being intertwined with multiple rallies and protests in the following weeks such as Trump's Phoenix Rally mentioned above, and a Free Speech Rally in Boston the week after Charlottesville. In addition to the Boston Free Speech Rally, terrorist attacks that took place in Barcelona have also been included in the discussion in the aftermath of Charlottesville. A brief presentation of these events is provided below.

The Boston Free Speech Rally was held one week after Charlottesville, and invited 'libertarians, conservatives, traditionalists, classic liberals, Trump supporters or anyone else who enjoys their right to free speech' (Sanchez 2017). While organisers said people from across the political spectrum were welcome, speakers were largely controversial figures associated with the socio-political right. Counter-protesters also attended the event and were described as 'a coalition of mostly left-leaning groups and activists' who had 'occasional clashes and shouting matches' with Donald Trump supporters (Sanchez 2017). As a result, the event appears to be associated with similar socio-political divides as the Charlottesville rally. The association with Charlottesville was explicitly acknowledged by Boston's Police Commissioner who stated they 'didn't want what happened in Virginia to happen here' (Sanchez 2017). As a result, tweets that refer to the Boston Free Speech rally are entangled with the narration of Charlottesville.

Additionally, five days after the Charlottesville rally and the associated car incident, on 17th August 2017, Younes Abouyaaqoub drove a van into a crowd in Barcelona, killing 13 people and injuring at least 130 others. Also, on 18th August, five men drove a car into pedestrians in Catalonia and then stabbed a woman to death. The attacks were attributed to the same terror cell that was reportedly linked to ISIS. The temporal closeness of these events to Charlottesville, along with the use of a vehicle as a weapon, resulted in this event being mentioned in discussions of the Charlottesville rally.

The analysis of the ways in which the overarching story of Charlottesville is constructed in the five main accounts and the types of positions such accounts take in performing particular identities will refer back to the events explained above, where relevant.

6.4 Telling stories on social media: Small stories heuristics

On the one end of the continuum of narrative dimensions introduced by Ochs and Capps (2001:20), narratives contain ‘an active teller, highly tellable account, relatively detached from surrounding talk and activity, linear temporal and causal organization, and certain, constant moral stance’ (Ochs and Capps 2001:20). On the other end of the continuum, we find ‘smaller’ (Georgakopoulou 2015), atypical stories that document unfolding, mundane events that can be detached from their situated context and be co-constructed by the teller and audience members.

On social media, in particular, small stories take precedence over “typical” or “big” stories’ (Giaxoglou 2021:51). According to Georgakopoulou (2016:269), the proliferation of small stories on social media is ‘no accident’, because platform affordances play a role in making traditional long narratives unlikely online. On Twitter for example, the sharing of fragmented, unfolding events is facilitated by an ‘always on’ model that enables users to update others on events as they happen, while the length of these updates is constrained by a 140-character limit per tweet¹⁹. As a result, ‘rather than seeing coherent and whole stories published on Twitter, individual fragments are posted’ (Sadler 2017:3275), contributing to narratives being created across multiple tweets and users (Page 2012, 2018).

Georgakopoulou (2015) also argues that small stories are used for identity performance and presents a type of small stories heuristic as an analytic tool for investigating such stories. Georgakopoulou (2015:256) describes small stories research as ‘a model *for* not a *model of* narrative analysis’, thus enabling the investigation of narrative without imposing the kind of stringent boundaries associated with older analytical frameworks of so-called “big” stories (e.g., in a strict adherence to Labov’s (1972) six-part structure of oral narratives).

Georgakopoulou’s (2015:258) small stories model explores the connections of three ‘separable but interrelated levels of analysis: ways of telling, sites, tellers’. Ways of telling are defined as ‘the communicative how’, including but not limited to features like a story’s plot, the kinds of events that are narrated, and the intertextual links with other stories. Sites refer to ‘the social spaces in which narrative activities take place’ (Georgakopoulou 2015:258). Finally,

¹⁹ Until November 2017, tweets had 140-character limits. At the time of writing (2022), tweets have 280-character limits.

tellers encompass the participants involved in a communicative activity in terms of their role within the narrative, their identity as members of certain social groups, and as individuals with particular beliefs, values, and so on. In the following sections, I explain how I applied these three levels as a heuristic for the analysis of the storying of Charlottesville by the five disinformation accounts under investigation.

6.4.1 Tellers

The chapter will firstly analyse the Charlottesville narrative and the ways in which the accounts identify with certain personas and positions through the lens of tellers. In order to do so, I will firstly analyse what forms of tellership are mobilised by the disinformation accounts when storying Charlottesville: single or multiple. Then, the analysis will focus on the type of actors involved in instances of co-tellership, identifying actors based on their roles within the communicative context of Twitter and their political affiliation. Such an analysis will enable an insight into the positions the disinformation accounts take in relation to other voices and tellers they engage with when telling the story of Charlottesville.

As acknowledged within the Data and Methodology chapter, the analysis of co-tellers in instances of multiple tellership involved three types of coding including the extent to which the other tellers were associated with ‘verified’ accounts or not; the professional background of these tellers; and their political affiliation, if known. Firstly, the co-tellers were labelled in terms of whether their accounts were ‘verified’²⁰ by Twitter or not. Within the ‘attention economy’ of social media that views attention and engagement as valuable currency (Zulli 2018), verified users can be associated with having high levels of influence and social capital due to their status as ‘notable’, and their higher numbers of followers compared to ordinary users (Twitter 2021a). Within an attention economy, where visibility appears key to achieving ‘success’ through engagement (Marwick 2015; Seo et al 2019), it may be expected for disinformation accounts to engage with popular actors, as higher visibility may fulfil better their aim to exacerbate social tension and create chaos.

If the accounts of co-tellers were verified, they were further categorised based on the professional domain that the accounts were associated with. To find this information, I searched

²⁰ ‘Verified’ status is shown on Twitter by the addition of a blue tick next to a profile’s username. Such status is attributed to accounts that are deemed ‘authentic, notable, and active’ (Twitter 2021a). As defined by Twitter, the notability of an individual’s account is assessed on the basis of high numbers of followers, high numbers of mentions, increased search activity on Google trends, a stable Wikipedia article about the individual, ties with official organisations, and expertise in matters of public interest.

each person online and used a combination of self-categorisations (e.g., their Twitter bios) and external sources (e.g., news articles). In cases where the other Twitter user(s) did not have verified accounts, I categorised them as ‘ordinary person’. Examples of the labelling criteria are shown in Table 18.

Table 18: Broad categories of co-tellers

Category	Co-tellers belonging to this category
US political actors/government officials	The President, politicians, political aides, political strategist
International political actors/government officials	Polish Prime Minister, Iranian parliament, UK politician
Political commentators	Individuals famous for sharing their political viewpoints in the media e.g., Tucker Carlson, Tomi Lahren, Dinesh D’Souza
News organisations/Journalists	News organisations and the journalists affiliated with them. Journalists were defined as those who are known for reporting news rather than solely giving their opinions about politics e.g., Fox News, CNN, Jim Acosta. Also includes websites dedicated to spreading ‘alternative facts’ e.g., NewsPunch
Celebrities	Those who are famous for something other than politics e.g., actors, sportspeople, gameshow presenters.
Activists	Those who campaign for certain causes e.g., civil rights activist, pro-choice activist
Ordinary people	Those who do not have verified Twitter profiles and are not famous e.g., ‘Trump supporter’, ‘crowd’, ‘counter-protesters’, as well as other Internet Research Agency accounts that pose as ordinary persons.

Finally, co-tellers were labelled based on their political affiliation. Political affiliation appeared to be a relevant area of analysis because: (i) the event’s name explicitly indexed political affiliation (Unite the **Right**); (ii) the mainstream media reported the event as ‘a stark demonstration of the political divide in the US’ (BBC News 2017c); and (iii) previous research into Twitter responses to the event finds ‘stark distinctions’ between right and left-affiliated communities and their evaluations of the violence and Trump’s response (Tien et al 2020). These categorisations were again made based on publicly available information. If co-tellers were verified, information pertaining to political affiliation was easily accessible through party membership (in the case of politicians), self-categorisations provided by political commentators and celebrities, and previous research or shared public knowledge regarding news organisations and journalists. In the case of non-verified users, political affiliation was identified using

resources like the Wayback Machine. For example, in the bios of the accounts during August 2017, some users self-categorised as being Republican or Democrat, or showing support for particular political figures. Where no such information was available, the categorisation was based on the way the actor was being spoken about by the user in question. For example, if a person was labelled an ‘alt-left thug’, they would receive the labels: ordinary person, left. As a result, the coding strategy includes the following political affiliations: political left, political right, and unknown in cases where it was impossible to retrieve any relevant information. It is important to acknowledge that there exists a spectrum of political ideologies, from moderate to extreme, within labels like political left or right. By forming broad categories, I am not attempting to suggest that all members of each category have the same views, but rather that their socio-political affiliations can be associated with either the right or left, regardless of where each individual’s personal beliefs may fall on the spectrum of the corresponding political ideology.

6.4.2 Sites

The chapter then analyses the tweet formats that tellers used when identifying with other users and positions through the lens of ‘site’. To do this, the analysis will focus on the social media site and its affordances for engaging in different forms of tellership. According to Page (2018:199), Twitter affords a ‘spectrum of [tellership] possibilities’ through numerous tweet formats, where not all discursive contributions are ‘equal’. For example, users can occupy the role of single, main teller by posting original tweets that they have authored (see Figure 33). Alternatively, users can engage in co-tellership by reposting other users’ tweets (retweets, see Figure 34), reposting others’ tweets while adding their own comments (quote tweets, see Figure 35) or replying directly to another user’s post (replies, see Figure 36).

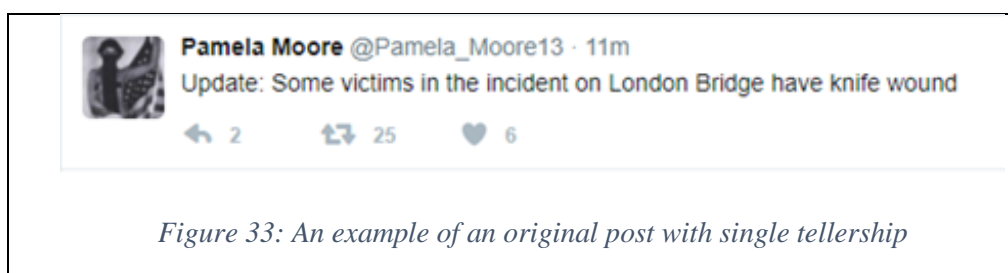


Figure 33: An example of an original post with single tellership



Figure 34: An example of a retweet, where @TEN_GOP retweets Donald Trump



Figure 35: An example of a quote tweet, where @USA_GunSlinger reports another user's tweet and adds their own commentary



Figure 36: An example of a reply, where @TEN_GOP replies to another user

All these tweet formats afford different possibilities for tellers in terms of their narrative roles, the distribution of turns, the reciprocity format, and control of the text. Table 19 provides an overview of the different tweet formats and their impact on the aforementioned narrative aspects. The different possibilities for co-tellership and their impact on the telling of the Charlottesville events will be discussed in more detail in section 6.6.

Table 19: A summary of tweet formats and their associated narrative potentials

Tweet format	Tellership	Narrative roles	Control of text	Turn distribution	Reciency format
Original	Single	Tweet writer is the main teller	Main teller	One turn	One to many
Retweet	Multiple	Person being retweeted: initial teller Retweeter: story introducer	Initial teller	Retweeter gives turn to initial teller	One to many
Quote tweet	Multiple	Person being quoted: initial teller Quoter: Story introducer, potential problematiser ²¹ , and potential co-teller	Quoter	Two turns compressed into one	One to many
Reply	Multiple	Person being replied to: initial teller Person replying: potential problematiser, potential co-teller	Shared	Sequential, separate turns	One to one; one to many

It is important to acknowledge that users also have the option to ‘mention’ other users within original tweets e.g., by including ‘@user’ as a form of address. In line with Dayter’s (2016:114) suggestion that the use of this feature is typically one-sided and is distinguished from ‘most interactive tweets’ by being ‘closer to the monological end of the communicative continuum’, tweets that contain mentions can be viewed as having one author, and thus, constituting single tellership.

In addition to these platform-afforded formats, another category emerged when conducting the analysis: the category of ‘original (reported speech)’ for tweets that reported the speech of someone else. In these cases, the format of the tweet is original (the profile owner is the only Twitter account involved in the telling), but the content being shared has been attributed to an external initial teller. For example:

Example 1: Original (reported speech)

@Pamela_Moore13: .@POTUS: "We condemn in the strongest possible terms this egregious display of hatred, bigotry and violence on many sides, on many sides."

²¹ Ochs and Taylor (1992) use the term problematiser to refer to a person who evaluates and/or questions a narrator’s actions

Tweets that employ reported speech can include no added commentary, as evident in Example 1. But they can also contain positive or negative evaluations (see Example 2 and Example 3, respectively), or comments that are indirectly linked with what is being reported (see Example 4).

Example 2: Original (reported speech) with positive evaluation

@TEN_GOP: Crowds chant "CNN sucks" after Trump calls their ratings "pathetic". This is priceless! #PhoenixRally

Example 3: Original (reported speech) with negative evaluation

@TEN_GOP: CNN's Wolf Blitzer: #Barcelona attack 'Copycat version of what happened in Charlottesville' This is just appalling

Example 4: Original (reported speech) with added commentary

@TEN_GOP: Trump supporters at #PhoenixRally chanting "Peace and Love!" Hmm 🙄 I can't seem to find any coverage on CNN or MSNBC...

The practice of reporting another person's speech is common within storytelling. According to Bakhtin (1981:337), talk produced by others is 'one of the most widespread and fundamental topics of human speech' and it is therefore not uncommon for such speech acts to constitute the actions of a story. Within narratives, reported speech can serve numerous functions including, but not limited to, invoking the words of an authority, discrediting original sources, summarising attitudes of others, dramatising points, and giving evidence for positions (Buttny 1997:478). Reported speech has also been described as 'constructed dialogue', as the report is influenced by the purpose of the communication and the context in which the interaction is (re)embedded (Tannen 2007:112). As a result, reported speech can be used by narrators as 'an effective positioning device, allowing current speakers to display alignment with or opposition to the recipients' and 'invoke particular membership categorization devices' (Griswold 2016:75;84).

In terms of narrative roles, original (reported speech) facilitates the possibility for tweet authors to occupy the role of story introducer, putting another teller's tale 'on the table' (Ochs and Taylor 1992:309) and giving message recipients the opportunity to view and potentially evaluate the reported speech in their own quote tweets, replies, and so on. Where original

(reported speech) tweets are accompanied by evaluations by the tweet author, authors occupy both the roles of story introducer and problematiser, introducing and evaluating the words of, for example, Phoenix rally crowds and Wolf Blitzer (see Example 2 and Example 3). The inclusion of another's voice within original (reported speech) tweets arguably makes them more dialogic than original tweets with single tellership.

The format of original (reported speech) differs from other forms of co-tellership facilitated by Twitter due to the lack of active participation from the co-teller in question. In the case of retweets, quote tweets, and replies, the co-teller's contribution is available on Twitter for others to see, while in original (reported speech), the entire contribution is shaped and controlled by the reporter. While reported speech is by no means a co-tellership strategy afforded solely by Twitter, it is proven to be a means through which co-tellership can be achieved even in original tweets that appear as single tellership on the surface. Consequently, original (reported speech) will be investigated alongside other tweet formats previously mentioned.

6.4.3 Ways of telling

The 'ways of telling' will be analysed in this study by investigating how characters, events, and co-tellers are evaluated in tweets (re)posted by the disinformation accounts. The analysis of such evaluations in the narrative discourse can be revealing of the ideological and political positions that are projected by the disinformation accounts in such tweets (cf. Georgakopoulou 2014:526). According to Bortoluzzi (2016:152), positioning can be understood as a means of (dis)identification, i.e., an 'active and contextual process whereby people are called to define themselves, locating themselves vis-à-vis known others'. In other words, the discursive navigation of social group membership can play a central role in identity performance. Consequently, this analysis will investigate how users told the story of Charlottesville to position themselves towards others and thus perform (dis)identification with wider social groups. Guided by Georgakopoulou's (2015) small stories heuristic, ways of telling will be explored by investigating what parts of the overarching story of Charlottesville were included and/or silenced in instances of co-tellership, and how events, characters, and co-tellers were described and assessed. In addition to the way that users constructed events, characters, and co-tellers, the analysis of the 'ways of telling' will also focus on intertextual links with other discourses and the wider context within which characters and events are positioned (cf. Georgakopoulou 2015:258; Page 2012:202).

By drawing on the three separate but interrelated levels of tellers, sites, and ways of telling, the analysis will focus on the following areas: a) the type of (co)tellership and the types of

(co)tellers disinformation accounts engage with when telling the story of Charlottesville; b) the tweet formats disinformation accounts employ when (co)telling the story of Charlottesville; and c) the ways in which disinformation accounts (dis)identify with different people and groups in the telling of the story of Charlottesville. By addressing these areas, the chapter aims to provide an insight into how disinformation accounts used content production practices facilitated by Twitter to authenticate ‘types of personas’ that have also been invoked in the accounts’ metadata.

6.5 (Co-)tellership and dis/identification in the storying of Charlottesville

This section presents the tellership analysis for each of the five accounts focussed on throughout the thesis. Firstly, the analysis will show the kind of tellership being used by each account: single or multiple. Then, the analysis will explore the types of co-tellers that each disinformation account interacts with when storying Charlottesville.

6.5.1 Tellership formats used in the storying of Charlottesville

Table 20 shows the distribution of single and multiple tellership in each account.

Table 20: Single and multiple tellership tweets across the five disinformation accounts

Disinformation Account	Single tellership tweets (n)	Multiple tellership tweets (n)	Total (n)
@TEN_GOP	131	141	272
@Pamela_Moore13	72	112	184
@wokeluisa	49	4	53
@SouthLoneStar2	36	6	42
@USA_GunSlinger	73	59	132
Total	361	322	683

As shown in Table 20 all accounts engage in both types of tellership: at times they take the role of a single teller of any events reported and express their stances towards such events (see Example 5) and at times they engage with others by retweeting events reported by other users in the network (see Example 6, for instance; other formats of multiple tellership are discussed in section 6.6).

Example 5: Single tellership

Tennessee @TEN_GOP	Whatever your political view is, you have to condemn the violence, regardless of who's committing it. #Charlottesville
--------------------	--

Example 6: Multiple tellership

Tennessee @TEN_GOP	RT @Cernovich: ANTIFA, a left wing organization, threw urine on a left wing media outlet.
--------------------	---

Nevertheless, Table 20 reveals some variation in the tellership patterns noted across the key accounts. While @TEN_GOP, @Pamela_Moore13, and @USA_GunSlinger engage in both single and multiple tellership, @wokeluisa and @SouthLoneStar2 show a stronger preference for tweets where they appear as the only teller, and thus, retain control over the texts being produced. The lower number of tweets posted by @wokeluisa and @SouthLoneStar2 appear to reflect wider patterns within the data; both @wokeluisa and @SouthLoneStar consistently posted less tweets than the other accounts in the months preceding Charlottesville.

Considering that co-tellership enables users to position themselves towards and, thus, (dis)identify with other tellers, the following section will examine what kinds of actors the accounts interact with. It is through such interactions with other co-tellers that disinformation accounts perform particular personas. As shown in chapters 4 and 5, four of the five accounts (@TEN_GOP, @Pamela_Moore13, @USA_GunSlinger, @SouthLoneStar2) construct the persona of an American affiliated with the socio-political right. Of the accounts associated with the socio-political right, @Pamela_Moore13 and @USA_GunSlinger create white female personas, while @SouthLoneStar2's profile picture presents an identity that evokes connotations with traditional white masculinity and conservatism. On the other hand, @wokeluisa's profile picture indexes a Black female persona. Based on previous research into performances of authenticity in relation to social group membership (Maragh 2018), it may be expected for @TEN_GOP, @Pamela_Moore13, @SouthLoneStar2 and @USA_GunSlinger to identify with co-tellers from the socio-political right and to narrate Charlottesville events from stereotypical right-wing perspectives. In contrast, it may be expected for @wokeluisa to identify with those from the left and adopt normative left-wing discourses, with a particular focus on race; previous research suggests that the Black community were specifically targeted by 'left trolls' during the IRA's disinformation campaign (Freelon et al 2022; Linvill and Warren 2020) and that IRA accounts presented Black and white Americans as 'binary groups that are internally homogenous with respect to politics' (Arif et al 2018:13).

6.5.2 Co-tellers involved in the storying of Charlottesville

Table 21 indicates the range of other Twitter users with whom the disinformation accounts appear to engage in formats of co-tellership, as well as the frequency (n) in which they appear in each account. As shown in the table, disinformation accounts engage both with powerful actors in the Twittersphere and in wider society, such as politicians/government officials, political commentators, news organisation/journalists, celebrities and activists, and with ordinary users of the platform (i.e., people with non-verified profiles).

Table 21: Types of co-tellers involved in storying Charlottesville across the five accounts

Disinformation Account	Political actors/ Government officials		Political Commentators	News organisations / Journalists	Celebrities	Activists	Ordinary People	Total ²²
	US	non-US						
@TEN_GOP	44	4	32	13	8	4	40	145
@Pamela_Moore13	20	8	28	10	12	5	30	113
@wokeluisa	3	0	0	1	0	0	0	4
@SouthLoneStar2	1	1	0	2	1	1	1	7
@USA_GunSlinger	20	2	10	1	0	2	24	59
Total	88	15	70	27	21	12	95	328

As shown in Table 21, all accounts draw on fewer ordinary voices than those with verified Twitter accounts and some level of recognisability as political figures, journalists, celebrities, or high-profile activists. This preference may be due to the greater levels of social capital associated with the latter; within the ‘attention economy’ of Twitter, engagement is ‘valuable currency’ (Zulli 2018) and ‘the communicative dynamics of Twitter are well placed to boost celebrity status and reinforce the distinction between elite and nonelite identities’ (Page 2012:96).

According to previous research on politically charged breaking news events, news organisations, journalists, and celebrities play a key role in relaying convincing news updates to Twitter users (Hu et al 2012). While this creates the expectation that disinformation accounts are likely to engage equally with these three groups in their (co-)telling of the Charlottesville events, this is not the case in the sample. Instead, political commentators and political actors are drawn on very frequently in all cases.

²² In each table, the total may add up to more than the instances of multiple tellership because some tweets contain more than one co-teller.

Further analysis of the types of the US government officials and political commentators engaged with by @TEN_GOP, @Pamela_Moore13, and @USA_GunSlinger demonstrates that most of these actors are affiliated with the socio-political right. As a result, these three accounts engage less with traditional news media and more with alternative sources for news, particularly in the form of co-tellers who are politically engaged and affiliated with the socio-political right.

@SouthLoneStar2 also signals identity aspects that can be associated with the socio-political right (see sections 4.3.4, 4.4.1, and 5.4), but differs from the other three right-wing accounts: the account engages with a range of co-tellers that include but are not limited to government officials and political commentators. In terms of political orientation, five of the actors @SouthLoneStar2 engages with are affiliated with the socio-political left, while two have unknown political affiliations. Unlike the other three accounts, @SouthLoneStar2 engages with co-tellers who appear to belong to a different social group than is evoked in the account’s metadata. In a similar vein, @wokeluisa also appears to engage with co-tellers with socio-political affiliations that differ from identity aspects signalled in the account’s metadata; for example, the account engages with President Trump, a co-teller that does not fall in the same political area represented by the word ‘woke’ (i.e., left wing). These discrepancies will be further scrutinised in the next section that analyses how the accounts engaged with the actors discussed above, especially in terms of the tweet formats mobilised in each case.

6.6 Tweet formats and dis/identification in the storying of Charlottesville

As mentioned in section 6.4, platform affordances offer Twitter users a range of tweet formats that establish different patterns of co-tellership (see Table 19).

Table 22 shows that the accounts use a range of tweet formats when engaging in co-tellership, mostly engaging in communication where they occupy Goffman’s (1981) role of ‘animators’ at some level, sharing content that has been attributed to an external initial teller such as through retweets (Page 2018:141) or original (reported speech).

Table 22: Distribution of tweet formats across the five disinformation accounts

Disinformation Account	Original (reported speech)	Retweet	Quote tweet	Reply	Total
@TEN_GOP	53	63	3	22	141
@Pamela_Moore13	69	43	0	0	112
@wokeluisa	3	1	0	0	4
@SouthLoneStar2	6	0	0	0	6

@USA_GunSlinger	31	22	3	3	59
Total	162	129	6	25	322

Based on Table 22, it appears that these disinformation account holders primarily employ formats of multiple tellership that enable them to retain control over the narrative being constructed. In both original (reported speech) and retweets, turns are arguably less dialogic than in quote tweets and replies, where narratives are actively negotiated across turns between multiple users. The preference for original (reported speech) over the format of quote tweet may be related to the advantages of this format, as it enables users to share content appearing in other formats (e.g., speech from a press conference) and to control and/or obscure the original contribution by the co-teller. Having identified the range of tweet formats mobilised in the disinformation accounts, I will examine how each format can be used as a potential identification mechanism by paying attention to how the respective accounts engage with others in each format.

First, retweets have been presented by previous literature as an affiliative practice that ‘obviously align [the retweeter] with the stance expressed in the original tweets’ (Gruber 2017:6), due to their ability to signal agreement and symbolise an act of friendship, loyalty or homage (boyd et al 2010). While @TEN_GOP, @Pamela_Moore13, and @USA_GunSlinger retweet a range of actors that span all actor types discussed in section 6.5.2, most co-tellers are associated with the socio-political right (n: 55, n: 35, n: 16, for each account accordingly). By retweeting these co-tellers, @TEN_GOP, @Pamela_Moore13, and @USA_GunSlinger amplify the voices of those from the socio-political right. At the same time, the absence of left-wing voices (n: 1, n: 0, n: 0, in each account accordingly) further supports the argument that retweets are used to signal affiliation with right-wing co-tellers and their narratives.

In terms of the other tweet formats, @TEN_GOP, @Pamela_Moore13, and @USA_GunSlinger also appear to share right-wing voices (e.g., through reported speech in original tweets). They also often positively evaluate right-wing actors e.g., through positively loaded evaluative terms like ‘nailed it’ and ‘sweet’ in the second and last examples in Table 23, and through displays of gratitude and politeness features such as ‘thank you for everything’ in the third example in Table 23.

Table 23: Examples of tweet formats used when engaging with right-wing co-tellers

Original (reported speech) no added comments	@Pamela_Moore13	Nigel Farage: "American leftists are more concerned with tearing down statues than identifying who the real enemy of the West is."
--	-----------------	--

Original (reported speech) with positive evaluation	@TEN_GOP	Mike Huckabee nails it: If President Trump “shot” Charlottesville driver between the eyes, he’d still be criticized.
Reply that complements initial contribution	@USA_GunSlinger	@realDonaldTrump Thank you for everything, Mr. President!
Quote tweet with positive evaluation	@TEN_GOP	Sweet! Get that ass ready CNN. [Original: @charliespiering: Steve Bannon returned to Breitbart News as Executive Chairman of Breitbart News and chaired our evening editorial meeting]

On the other hand, contributions from left-wing co-tellers that may appear in other tweet formats (e.g., original/reported speech, replies, quote tweets) are primarily problematised (e.g., through analogies between Antifa and Al Qaeda, see the first example in Table 24) and negatively evaluated (e.g. through negatively loaded vocabulary, such as ‘fakest’, ‘brainless’, ‘bigot’, see the second and third examples in Table 24). In the case of @Pamela_Moore13, in particular, any left-wing voices are presented only through original (reported speech) tweets, enabling the account holder to fully control the representation of such voices on the account. For example, in Figure 37, the running headline that appears in the screenshot of the TV news programme is reformulated in the tweet by foregrounding part of the message (peace through violence vs. peaceful movement).

Table 24: Examples of tweet formats used when engaging with left-wing co-tellers

Original (reported speech) with negative evaluation	@Pamela_Moore13	MSM says Antifa are a peaceful movement. Remember what they said about Al Qaeda? [Accompanied by an image-see Figure 37]
Reply that problematises initial contribution	@TEN_GOP	@Acosta You are the fakest of the #FakeNews
Quote tweet with negative evaluation	@USA_GunSlinger	And here it is, the face of the Left. Sexist brainless bigot.



Figure 37: Image accompanying Table 24

While section 6.5.2 shows that @TEN_GOP, @Pamela_Moore13, and @USA_GunSlinger engage with primarily right-wing actors, this section demonstrates that this engagement takes different formats: retweets amplify narratives of right-wing co-tellers, but other formats also achieve the same effect through positive endorsements of such voices and, at times, negative evaluations of voices from the left. Through these means, the storying of Charlottesville by these three accounts appears to be shaped by right-wing perspectives, thus identifying them with right-wing positions and fulfilling stereotypical expectations associated with the political affiliations evoked in metadata units. Based on the consistent conflation of normativity and authenticity in identity and social media research (Kytölä and Westinen 2015; Maragh 2018), it could be the case that such orientation to normative right-wing positions may contribute to the construction of valid and thus authentic representations of right-wing American personas.

As mentioned in section 6.5.2, @SouthLoneStar2 does not engage with any right-wing voices, despite the fact that their profile metadata invokes connotations with the socio-political right, similar to the three aforementioned accounts. In fact, the account engages instead with actors from the socio-political left or with unknown political affiliations. Table 22 shows that these voices appear in the form of original (reported speech), with three out of these six tweets including evaluations that problematised the ‘initial teller’s’ contribution (i.e., CNN, George Soros, and ‘progressives’). It therefore appears that @SouthLoneStar2 primarily uses multiple tellership as a means of disaffiliating with and criticising the socio-political left, which according to social identity theory (Tajfel 1982), may result in indirect alignment with and endorsement of the socio-political right.

The use of **disalignment** as a means of identity performance is also reflected in @wokeluisa’s use of tweet formats. As shown in Table 22, @wokeluisa engages in co-tellership using retweet and original (reported speech) tweet formats. While the retweet amplifies the voice of a CNN journalist which is in line with the left-wing persona construed in the profile metadata, the original (reported speech) tweets include speech produced by Donald Trump, together with negative evaluations. Similar to @SouthLoneStar2, then, @wokeluisa also employs formats that allow them to retain control of any other voices and their narration of the Charlottesville events, criticising and disaligning themselves from politically oppositional co-tellers. As a result, @wokeluisa’s engagement with co-tellers from the socio-political right potentially indirectly aligns with the persona created in the profile metadata.

Overall, @TEN_GOP, @Pamela_Moore13, and @USA_GunSlinger reinforce their identification as right-wing personas in their metadata by primarily amplifying and positively evaluating voices from the right and, at times, by disidentifying with voices from the left in

their tweets. In contrast, @SouthLoneStar2 and @wokeluisa reinforce identity aspects signalled in metadata through **dis**identification with the socio-political left and right, respectively, as evident in the tweet formats mobilised in their storying of the Charlottesville events. A closer analysis of their ways of telling should enable a more detailed understanding of how they positioned themselves in relation to co-tellers and wider social groups.

6.7 Ways of telling and dis/identification in the storying of Charlottesville

In this section, the overarching story of Charlottesville created on each account will be analysed by exploring how fragments of relevant events were narrated. To understand how accounts used different ways of telling to (dis)identify with different social groups when storying Charlottesville, it is necessary to acknowledge the chronological timeline of the events which are both directly and indirectly linked with the overarching story of the Charlottesville Unite the Right Rally. Figure 38 is a visualisation of the relevant events and their temporal order from February 2017, when the removal of a Confederate statue was voted in the city council, until late August 2017, when the Boston Free Speech Rally took place. The focus in Figure 38 is on the events that were made relevant within the data.



Figure 38: Chronological timeline of Charlottesville events

The analysis will now examine how users (dis)identified with other users and wider social groups when narrating the events shown above, with a specific focus on how the indexing of community membership may contribute to the construction of valid representations of particular personas.

6.7.1 The crash

Within the story of Charlottesville, the car crash arguably constitutes the key complicating action, the story's climax that constitutes the peak of a narrative arc (Boyd et al 2020:1). As a result, it may be expected that the disinformation accounts' storying of Charlottesville will draw heavily on the event of the crash, along with the characters involved (the driver and the victims). However, the analysis reveals that only @TEN_GOP used co-tellership to narrate the complicating action of the crash.

In Example 7 that illustrates the type of tweets produced by this account, @TEN_GOP retweets an unverified ‘ordinary’ Twitter account that has since been suspended. Despite the account not being verified and thus being categorised as ‘ordinary’ in section 6.5.2 (see Table 21), my research revealed that this account was a popular source of ‘racist hate and disinformation fed by 4chan’ (O’Brien 2017) with over 93,000 followers.

Example 7: @TEN_GOP retweets an 'ordinary' Twitter user

Tennessee @TEN_GOP	RT @Twitter user: CPR being done on someone after Antifa caused a terrible accident.
-----------------------	--

The tweet mentioning the crash is posted alongside a video that shows the aftermath of the incident. Out of the different characters that may be involved in the Charlottesville events, only two are mentioned; namely, the victim of the crash, Heather Heyer, and Antifa. Unlike the victim who remains nameless (‘someone’), Antifa are named and positioned as the agent responsible for ‘causing’ the ‘terrible accident’ that injured the victim. According to Georgakopoulou (2014:527), the way that an event is described in online storytelling can ‘project specific affiliations on the teller’s part’ through a process she refers to as ‘narrative stancetaking’. Narrative stancetaking involves the use of conventionalised devices (e.g., the use of past tense ‘caused’, use of evaluative devices ‘terrible’) to indicate that an activity involves events and characters whose actions and speech can be assessed, thus suggesting that there is a story ‘in the making’ (Georgakopoulou 2017:276). In her study into the resemiotization of the assault of two female MPs by a male MP on a Greek TV breakfast show, Georgakopoulou (2014:527) finds that descriptions of the assault as a ‘punch’ rather than a slap, for example, cast the assault as more severe and index an affiliation with the victim. In this case then, removing agency from the driver, who is not mentioned at all in the tweet, and instead attributing responsibility to Antifa, arguably projects an affiliation with the driver (who is not at fault) and disaffiliation with Antifa (who are the cause of something terrible).

Considering that Antifa (a loose movement of anti-fascist activists) is associated with the socio-political left, particularly in the discourse of prominent actors like President Trump (Bogel-Burroughs and Garcia 2020), the tweet also implicitly criticises the left as responsible for the tragic events in Charlottesville, a practice that is also noted in other @TEN_GOP tweets where Antifa is presented as throwing urine at reporters. As a result, the event of the car attack is constructed as an accident caused by Antifa, situated in the context of other ongoing incidents of left-wing wrongdoing.

According to the concept of the ideological square, distance between social groups can be exacerbated by emphasising negative behaviours about ‘them’, while suppressing negative behaviours about ‘us’ (Mattu and Lubbe 2007:407). It therefore appears that when narrating the event of the crash, @TEN_GOP used narrative stancetaking to negatively assess and disidentify with left-wing actors, thus projecting an affiliation with the socio-political right which was also achieved by engaging with users from that area (see section 6.5.2). Such practices that shape the content of the account end up reinforcing identity aspects indexed in the profile metadata.

This kind of narration of the car crash, where the right is distanced from responsibility by right-wing accounts, is further reinforced by the narration of a different terrorist incident involving a car attack, i.e., the terrorist incident in Barcelona.

6.7.2 Terrorist incident in Barcelona

As shown in Figure 38, five days after Charlottesville, a car was driven into pedestrians in Barcelona. The incident in Barcelona was mentioned by some of the disinformation accounts, particularly in relation to Charlottesville, where the two events were contrasted and potential connections between them were challenged.

As illustrated in Example 8, on 17th August @SouthLoneStar2 reports on a CNN news broadcast, accompanied by a URL documenting the same event in the Washington Free Beacon (conservative news organisation). The tweet uses reported speech in which there are references to two car attacks against pedestrians: one in Barcelona and one in Charlottesville.

Example 8: @SouthLoneStar2 reports the speech of CNN

Texas Lone Star @SouthLoneStar2	CNN is saying Barcelona terrorist got the idea to run over people from Charlottesville. THIS IS INSANITY https://t.co/G4xnWRciVi
------------------------------------	--

While @SouthLoneStar2 refers to both attacks, they are constructed differently. Firstly, the characters in the events at Barcelona and Charlottesville are identified differently. In the case of Barcelona, the driver of the car is described as a terrorist, leading to the portrayal of the car crash as an act of terror. In the case of Charlottesville, the driver and his actions are omitted from the narration, constructing Charlottesville as an agentless event with no obvious link to terrorism. Example 8 thus implicitly refers to events at Charlottesville, assuming that what happened there is shared knowledge among the Twitter audiences. The construction of the Barcelona event as a terrorist attack in juxtaposition with the unspecified events at Charlottesville thus implies that the two events do not have grounds for comparison. This idea is reinforced by the evaluation of what is conveyed in the reported speech as ‘INSANITY’.

In Example 8, the identification of the reported speaker as ‘CNN’ generalises the comments made by one speaker, Wolf Blitzer, to the whole news organisation. The evaluation of ‘insanity’ is thus also applied to CNN more broadly. The negative evaluation of CNN aligns with wider right-wing discourses in relation to the media (cf. Trump’s criticism of CNN as ‘fake news’ (Farkas and Schou 2018:306)). @SouthLoneStar2’s inclusion of a link to Washington Free Beacon, an American conservative political journalism website, appears to construct the Washington Free Beacon as a preferable source of news over the ‘insanity’ of CNN. Arguably, this action is also in line with Trump’s broader discourse that portrays mainstream media as ‘corrupt, liberally biased, systematic liars and in need of replacement’ (Farkas and Schou 2018:307). Example 8 thus shows how disinformation accounts used both character-focussed and event-focussed narrative stancetaking to assess characters’ reported speech, and question the accuracy of reporting, respectively. The use of both of these evaluative tools disidentifies @SouthLoneStar2 with tellers associated with the socio-political left.

As mentioned in section 6.7.1, most of the accounts omitted the complicating action of the car crash from their storying of Charlottesville. In comparison, the Barcelona event, together with its complicating action (e.g., car running over people) which was framed as a ‘terrorist attack’, was included in the tweets posted around the Charlottesville attack by most accounts that present identity aspects relating to the socio-political right in their metadata. As a result, for these disinformation accounts, identification with the socio-political right was achieved by suppressing representations of any right-wing wrongdoing at Charlottesville, denying links with terrorism, and foregrounding terrorist incidents carried out by ‘others’.

6.7.3 Trump’s response to Charlottesville

6.7.3.1 Disidentification with Trump

As evident in Figure 38, on the day of the Unite the Right rally, President Trump held a press conference where he said ‘We condemn in the strongest possible terms this egregious display of hatred, bigotry and violence on many sides. On many sides’ (Philips 2017). His comments were met with backlash by the media, other politicians, and the public for not explicitly condemning racist groups like white supremacists and neo-Nazis, who were involved in violent clashes at the rally and with whom the driver of the car was affiliated (Smith 2017). Trump’s response was mentioned in all the disinformation accounts examined in this chapter.

In the case of @wokeluisa, where all instances of multiple tellership were negative evaluations of Trump’s reported speech (see section 6.6), Example 9 illustrates how fragments

of Trump’s speech were employed in the account’s tweets. The tweet foregrounds this particular fragment of speech without naming Trump, indicating that the quote is embedded within the context of ongoing debate surrounding Trump’s comments.

In Example 9, the speaker of the words (Trump) is omitted from the tweet, while ‘violent bigotry, white supremacists and nazi salutes’ implicitly refer to the events at the Unite the Right rally. The reference to characters involved in the rally combined with the temporal closeness to the event appears to assume that the audience will have prior knowledge both of Trump’s comments and of the ‘violent bigotry’ that @wokeluisa refers to.

Example 9: @wokeluisa reports the speech of Donald Trump

Luisa Haynes @wokeluisa	"On many sides." Three words that protected violent bigotry, white supremacists and nazi salutes!
----------------------------	---

As found in Georgakopoulou’s (2014) study, the reproduction of recognisable speech from a high-profile event can be used as a means of performing character-focussed narrative stancetaking. In Example 9, @wokeluisa negatively evaluates Trump’s speech by juxtaposing the terms ‘protected’ and ‘violent’, thus emphasising the implied absurdity of Trump’s response; following an act of violence, it is expected that victims rather than perpetrators will be protected. The identification of white supremacists as the characters responsible for the events that Trump attributes to ‘many sides’ functions to problematise the reported words of Trump. @wokeluisa’s attribution of blame to far-right characters distances other groups, such as those oriented to the socio-political left, from the violence of Charlottesville and challenges Trump’s suggestion that groups from ‘many sides’ were equally responsible.

By focussing on a fragment of Trump’s speech that neither condemns the perpetrators nor acknowledges the victim, @wokeluisa constructs Trump as having an affinity with unexpected and undesirable characters (white supremacists). This implication links to widespread criticism by the socio-political left towards Trump’s affinities with racist ideologies, not just following Charlottesville, but throughout his presidency (La Ganga 2016; Shrikant and Sierra 2021:6). @wokeluisa’s construction of Trump therefore appears in line with the account’s identification as ‘woke’ in the handle because it points to a persona that rejects racist ideologies and is thus ‘alert to racial or social discrimination and injustice’ (Oxford English Dictionary 2022). More specifically, the focus on racism as a means for (dis)identification in this context arguably recycles and reinforces normative expectations associated with the previously defined phenomenon of Black Twitter. Previous research suggests that race and race relations are discussed significantly more by Black users rather than white users (Spruce and Leaf 2017:51),

with Black Twitter often being used as a platform to bring attention to social justice issues relating to race and gender (Williams 2015:343). Consequently, disidentification with Trump on the grounds of racism within the storying of Charlottesville may be considered as normative, especially in combination with the identity aspects of race and ‘wokeness’ presented in the account metadata.

Example 9, together with other similar tweets posted in the period analysed, demonstrates how @wokeluisa used normative ‘woke’ discourse to engage in character-focussed narrative stancetaking. By reproducing and negatively evaluating Trump’s comments, @wokeluisa positioned herself as disaffiliated with Trump, his politics, and the wider group of white supremacists that he inappropriately ‘protected’ (cf. Georgakopoulou 2014:526).

6.7.3.2 Identification with Trump

While @wokeluisa negatively evaluates Trump’s response, other disinformation accounts amplify and positively evaluate his comments. In Example 10, @USA_GunSlinger quotes and reshares a different fragment of Trump’s response to Charlottesville (shown in Figure 39). In contrast to Example 9, @USA_GunSlinger foregrounds an arguably less controversial statement. Trump’s words in Figure 39 appear to act as a coda within the overarching story of Charlottesville, providing a statement that ‘returns the temporal setting to the present’ (Labov 2010:547) and alludes to what the country should do next/‘now’.

In Example 10, @USA_GunSlinger positively evaluates the closure that Trump appears to offer to the Charlottesville events. The tweet also orients to dominant right-wing discourses and, more specifically, to the ‘Soros Myth’, a set of far-right (and often Anti-Semitic) conspiracy theories that view George Soros’²³ financial support for what the right views as ‘left-wing causes’ as attempts to interfere with world-wide politics, specifically in the United States and Europe (Kalmar et al 2018:330; Stewart 2020:1210). The amplification and positive evaluation of different fragments of Trump’s response(s) to Charlottesville via retweets and/or original (reported speech) were also found in the other right-wing disinformation accounts under examination.

Example 10: @USA_GunSlinger complements Donald Trump's response to Charlottesville

<p>GunSlinger Girl @USA_GunSlinger</p>	<p>Wise words, Mr. President! We must stand united as one nation under God. Soros-paid protesters won't divide us! #Charoettesville [Accompanied by an image- see Figure 39]</p>
--	--

²³ George Soros is a Jewish multi-billionaire philanthropist.



Figure 39: Image accompanying Example 10

In Example 10, @USA_GunSlinger positively evaluates Trump by referring to his response to Charlottesville as ‘wise’, while ‘Soros-paid protesters’ are portrayed as divisive figures and as the agents responsible for the ‘kind of violence’ Trump refers and responds to. The contrasting assessment of characters suggests that there is a story in the making that positions Trump as a hero trying to fix the problems caused by a villain, Soros. As a teller, then, @USA_GunSlinger adopts specific stances towards these characters and through such stance-taking acts, the account is indexically positioned as affiliated with right-wing actors and politics and disaffiliated with other figures, such as Soros.

The construction of protestors as paid by George Soros omits any right-wing involvement in Charlottesville and shifts responsibility to a character that, according to the Soros Myth, is associated with the socio-political left²⁴ (Kalmar et al 2018:330; Kalmar 2020:191). By positioning Soros as ‘controlling’ protestors, @USA_GunSlinger alludes to the shared story that Soros controls the world and signals an ‘insider status’ within a community that believe Soros controls political events: the American far-right. References to the Soros Myth in relation to Charlottesville are also present on other right-wing disinformation accounts in this sample.

The reference to ‘one nation under God’ alludes to the US Pledge of Allegiance, and further demonstrates an orientation to normative expectations associated with the American socio-political right, particularly in relation to religion and patriotism. According to Claassen (2015:1-5), the so-called ‘God gap’ in American politics has led to the construction of Republicans as the ‘pious party’ trying to stop the ‘Godless’ Democrats from declaring ‘war on religion’. More specifically, this particular fragment of the Pledge of Allegiance has been a source of contentious political debate in recent years, with President Trump warning Evangelical Christians that the Democrats have dropped the four words from the Pledge (Pengelly 2020). It

²⁴ Alongside Soros’ support for causes that are generally viewed as ‘left-wing’ by the right, more specific accusations include Soros engaging in ‘behind-the-scenes manipulation of the 2016 US elections in favour of Hilary Clinton’ (Kalmar 2020:191)

could therefore be argued that the use of ‘one nation under God’ shows an endorsement of a stereotypically Christian and Republican narrative. Similar discourses that make reference to God and Christianity have also been found in tweets posted by the accounts that index right-wing identities in profile metadata (e.g., @TEN_GOP, @Pamela_Moore13, and @SouthLoneStar2).

Beyond narrative stance-taking, the amplification and positive evaluation of Trump’s speech as a means of projecting affiliation with his politics is also evident in tweets with non-narrative stancetaking. According to Georgakopoulou (2014:526), non-narrative stancetaking relates to comments where users ‘detach themselves from the story and provide general socio-political commentary’, positioning themselves as citizens affiliated with particular political parties rather than as (co-)tellers of a story. This kind of stancetaking is illustrated in Example 11, where @TEN_GOP’s reply to Trump does not engage with the narrative of Charlottesville but positions @TEN_GOP explicitly as pro-Trump (‘we’re so proud’). Non-narrative stancetaking tweets may not appear immediately relevant to the storying of Charlottesville but play a significant role in showing ideological positioning and authenticating particular positions as, in this case, a Trump-supporter. Similar tweets are also found across the other right-wing disinformation accounts under investigation.

Example 11: @TEN_GOP replies to President Trump

Donald J. Trump @realDonaldTrump	Leaving for New York City and meetings on military purchases and trade
Tennessee @TEN_GOP	@realDonaldTrump We're so proud! The American people support you!

6.7.3.3 Disidentification with critics of Trump’s comments

While right-wing disinformation accounts identified with Trump by sharing and evaluating his comments, they also engaged with his response to Charlottesville by positioning themselves in opposition to critics of his comments. In Example 12, @Pamela_Moore13 uses a series of screenshots to report and problematise a tweet written by Republican Senator John McCain, who criticises Trump’s response to Charlottesville. McCain has been characterised as ‘one of the most outspoken critics of Mr. Trump from the very start of his candidacy’ (BBC News 2018a) and as a result, became an enemy of the Trump-supporting right (Jaffe and Johnson 2018).

Example 12: @Pamela_Moore13 problematises contribution from @SenJohnMcCain

Pamela Moore @Pamela_Moore13	@SenJohnMcCain You openly supported Ukrainian Neo-Nazis. [Accompanied by images- see Figure 40, Figure 41, and Figure 42]
---------------------------------	--

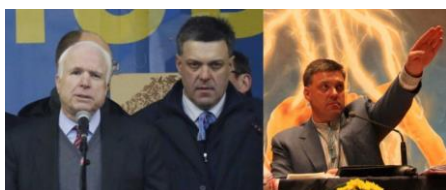


Figure 40: First image accompanying Example 12

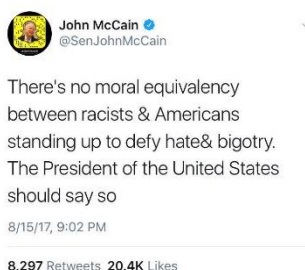


Figure 41: Image accompanying Example 12



Figure 42: Image accompanying Example 12

In Example 12, @Pamela_Moore13 evaluates McCain’s criticism of Trump by introducing a second story that refers to McCain’s past actions, evident in the photo that features him next to a Ukrainian politician. While McCain’s initial tweet focuses on Trump, @Pamela_Moore13 omits Trump from her tweet that focuses on McCain’s actions, thus creating distance between Trump and the ongoing debate surrounding responsibility for Charlottesville and affiliations with neo-Nazis. Instead, McCain is constructed as the person who ‘openly supported’ neo-Nazis, i.e., the kind of characters that he suggests Trump should condemn.

While the accompanying images in @Pamela_Moore13’s tweet appear to function as ‘evidence’ within the story, the term ‘neo-Nazis’ is explicitly used by @Pamela_Moore13. The selection of this characterisation acts as a key link between McCain’s actions and the Charlottesville events: the driver who killed Heather Heyer was characterised as a neo-Nazi (BBC News 2019). As a result, @Pamela_Moore13 suggests an affinity between McCain and the social group that the Charlottesville driver belonged to. The juxtaposition between McCain’s initial statement and the actions attributed to him by @Pamela_Moore13 serve to negatively assess McCain’s speech and actions by implying that McCain is a hypocrite, thus using character-focussed narrative stancetaking to project disaffiliation with McCain.

By focusing on McCain and his statement as a tellable event, rather than statements of many other critics of Trump, @Pamela_Moore13 appears to draw on wider social discourses of Trump-supporters that view McCain in a particularly negative light (Jaffe and Johnson 2018). Similar to the introduction of Soros in Example 10, @Pamela_Moore13 introduces McCain in her narrative of the Charlottesville events and recycles assessments that are embedded within wider right-wing discourses outside of the immediate context of Charlottesville. As a result, it

seems that within the narration of Trump’s response to Charlottesville, right-wing disinformation accounts introduce the negative actions of ‘others’ as a means of reframing widespread criticisms of Trump and attributing such negative events not to Trump, but to his commonly perceived enemies.

6.7.4 Boston Free Speech Rally

The analysis in section 6.7.3 foregrounds the mobilisation of Trump’s response as a relevant sub-story through which right-wing disinformation accounts (dis)identify with key actors and project particular personas. An orientation to Trump’s response and, particularly, his words that there was blame ‘on many sides’, is also evident in the way that disinformation accounts narrate other events in the aftermath of Charlottesville, such as the Boston Free Speech Rally.

Example 13: @TEN_GOP reports the speech of a free speech rally counter-protester

Tennessee @TEN_GOP	Alt-left thug screams at black police officer: "I'll spit on you black b*tch, you're suppose to be on our side."
-----------------------	--

Posted alongside a video showing crowds of people clashing with police at the Boston Free Speech rally, Example 13 foregrounds a moment from an alleged interaction between a (counter-)protester, referred to as ‘alt-left thug’, and a police officer who is identified as ‘black’. The reference to the police officer’s race both before and after the quotation marks appears to highlight race as a key element of the story. The Black police officer is presented as a passive recipient, the victim of aggressive threats from the agent, the ‘screaming and spitting’ alt-left thug. At the same time, the use of the term ‘thug’ to characterise the counter-protester negatively assesses the actions and the speech of the character. The foregrounding of negative incidents perpetrated by the socio-political left at rallies is also present in tweets by @Pamela_Moore13 and @USA_GunSlinger.

In Example 13, both the noun ‘thug’ and the verb ‘screamed’ evoke volume and aggression and thus signal that there is a story of violence in the making. The quote strengthens @TEN_GOP’s presentation of the left as aggressive; @TEN_GOP uses the quote of aggressive and threatening language to show rather than simply tell how the counter-protester was behaving like a thug (Buttny 1997:486). Through character-focussed narrative stancetaking then, @TEN_GOP disaffiliates with the speaker and (re)associates racist and violent behaviour with the socio-political left, rather than the Unite the Right rally protesters. The foregrounding of this event in the aftermath of Charlottesville, together with the lack of any references to right-wing wrongdoing, arguably contributes to an unfolding narrative where violence at rallies is

typical of the left. Considering the account’s identification with right-wing politics in the profile metadata, this kind of construction symbolises the left as the out-group.

Accounts also projected disaffiliation with the left by foregrounding negative behaviours in tweets that use non-narrative stancetaking. In Example 14, @USA_GunSlinger reports and evaluates the speech of a Democratic Senator. Similarly to Example 12, @USA_GunSlinger draws on terms (bigots, murderers, Nazis) that were widely used to criticise Unite the Right rally protestors, the neo-Nazi driver who killed Heather Heyer, and Trump, more generally. In doing so, @USA_GunSlinger reframes condemnations of right-wing actors and targets these negative judgements at members of the socio-political left. As a result, @USA_GunSlinger constructs Chappelle-Nadal both as violent and hypocritical; as discussed in section 6.7.3, the construction of Trump and the right as bigoted was widespread amongst the socio-political left.

Example 14: Non-narrative stancetaking to disaffiliate with the left

GunSlinger Girl @USA_GunSlinger	Missouri Senator Maria Chappelle-Nadal: "I Hope Trump Is Assassinated!" This is the modern left: bigots, murderers and nazis.
------------------------------------	--

Through the fragmented presentation of the Charlottesville events and depictions of violent behaviours of left-wing characters in the aftermath, the four, right-wing disinformation accounts construct an unfolding narrative where the left are responsible for the ‘terrible accident’ at the Unite the Right rally and other negative incidents. In doing so, accounts disidentify with the violent left and strengthen their identity performances as affiliated with the blameless socio-political right.

6.8 Summary

The analysis shows that throughout the different stages of the story - the key event/complicating action of the attack, Trump’s response, and other rallies - the accounts (dis)identify with different social groups through evaluations of key characters and allusions to dominant discourses. Across all disinformation accounts, transgressions of the ‘other’ are emphasised, while negative behaviours of their own social groups are omitted and suppressed. In all accounts, the storying of Charlottesville draws on characters from the public sphere, while characters like the driver and the victims are largely omitted. The prevalence of public figures and politicians within the accounts’ storying of Charlottesville complements their practice of co-telling (i.e., retweeting, replying, etc) the story with actors that are associated with explicit socio-political positions (see section 6.5.2). In all cases, the accounts narrate the Charlottesville-related events in ways that reinforce normative expectations associated with the identity aspects

signalled in account metadata, thus contributing to the construction of @TEN_GOP, @Pamela_Moore13, @SouthLoneStar2, and @USA_GunSlinger as valid representations of right-wing personas, and @wokeluisa as a valid representation of a left-wing persona.

6.9 Concluding discussion

In this chapter, I have explored how disinformation accounts use content production practices afforded by Twitter to co-tell the story of Charlottesville, perform social group membership, and consequently authenticate identity performances as American right and left-wing personas. I showed that the accounts utilised a range of tweet formats to narrate small, fragmented stories where the key characters and other co-tellers are being constructed and evaluated in ways that emphasise in/out-group membership, with a specific focus on socio-political (dis)affiliation.

Research into narrativity, both on and offline, suggests that the investigation of (co-)tellership roles can provide insights into social relations between multiple tellers, and as a result, into identity performance (Bortoluzzi 2017; Ochs and Taylor 1992; Page 2012). The first stage of this analysis showed that when storying Charlottesville, the accounts engaged in co-tellership with a range of actors, particularly those associated with explicit socio-political positions. The prominence of this type of co-teller arguably suggests that politics played a key role in the stories told on the accounts, corroborating previous research that argues that disinformation accounts tend to share politically divisive, partisan content (Linville and Warren 2020).

The next stage of the analysis revealed that co-tellership offered a means of (dis)identifying with different co-tellers: the accounts utilised a range of tweet formats to problematise tales and co-tellers in ways that contributed to the construction of socio-political out-groups, on the one hand, and to amplify and complement stories by co-tellers falling within their assumed socio-political in-groups, on the other. Out-groups were constructed by positioning ‘them’ (those from different positions on the political spectrum) as agents responsible for negative actions, by identifying ‘them’ with negative referential strategies, and by generalising negative characteristics and actions to ‘their’ broader social groups. In contrast, in-groups were signalled by first-person inclusive pronouns, performances of emotional connectedness, and the construction of shared, insider knowledge. The prevalence of socio-political affiliation within in/out-group formation led to the storying of Charlottesville on each account being associated with explicit socio-political positions, potentially explaining the prevalence of politically affiliated co-tellers evident in the first stage of the analysis.

Similar to Twitter practice associated with everyday uses and ordinary users, disinformation accounts manipulate various tweet formats to manage (co-)tellership. Previous research (e.g., Conover et al 2011) suggests that the retweet function can be interpreted as a means of signalling agreement and alignment with the stance being offered, and thus, with the user being retweeted. This analysis showed that, in most accounts, retweets were used to amplify messages that aligned with the narrative created in the disinformation accounts' own tweets. The idea that retweets were used in an affiliative manner is supported by the absence of this type of tweet on @SouthLoneStar2's profile; this account engaged in co-tellership primarily through disidentification and did not retweet any other users.

In terms of the negative assessment of characters and actions in my sample, previous research suggests that 'messages with negative sentiment on Twitter are key to popularity' (Ott 2017:62), with the posting of negative messages on Twitter not only being associated with higher levels of engagement, but with greater success in political campaigns (Duncombe 2019:410). According to Duncombe (2019:409), even accounts affiliated with governments, state leaders, and policy makers target individuals with inflammatory messages. It could be argued that this kind of behaviour is motivated by the normalisation of negative content on the site of Twitter; the targeting of individuals in a negative way has become somewhat expected by other users (Duncombe 2019:417). While disidentification as an identity performance strategy may therefore appear to orient to normative practices associated with Twitter use, the potential for such messages to 'lead to debates that are integrated into offline political outcomes' (Duncombe 2019:410) may arguably be related more specifically to the aims of disinformation accounts, i.e., to exacerbate social tensions and create chaos.

In addition to retweets, this study also foregrounds the use of reported speech in the accounts' tweets as a form of co-tellership that results in amplifying, as well as problematising, the stance conveyed in the reported discourse. The use of original tweets embedding reported speech appeared prominent within all analysed accounts and constituted nearly all instances of co-tellership in some accounts (e.g., @wokeluisa, @SouthLoneStar2). As shown in the analysis, such tweets do not include a link to the initial teller's original contribution. As a result, the embedded utterance is completely controlled by the teller, can be (re-)constructed in numerous ways, and played a key role in accounts' performances of disidentification.

As mentioned in section 6.3, previous research suggests that the uncertainty invoked by breaking news stories means that inaccurate information in relation to unfolding stories is not uncommon (Andrews et al 2016; Starbird et al 2014; Vis 2013). While it may therefore be expected for disinformation accounts to share information relating to the unfolding events, this

analysis showed how accounts primarily contributed evaluations to the overarching story of Charlottesville, offering up assessments of responsibility and causality as opposed to discussing what happened (i.e., the complicating action) in the incident of the car attack.

The use of Twitter to offer broad commentary and evaluation of news stories appears to orient to ordinary usage of the platform, where discussions of popular topics not only function to tell people what is going on in the world but communicate what other people think about these events (Zappavigna 2011:789). Within the specific context of breaking news, Ausserhofer and Maireder (2013:2) state that any single message ‘may include both information and commentary on an event’. This analysis shows how accounts primarily offered commentary in relation to Charlottesville, with a particular focus on blame and responsibility. The analysis demonstrated how agency was removed from characters involved in the key event (like the driver and the victim) and was instead attributed to political figures and broad socio-political groups. The evaluations offered by the disinformation accounts often alluded to dominant discourses related to socio-political communities, indexing ‘insider status’ and reinforcing performances of socio-political (dis)identification.

According to previous research, perceptions of authenticity as realness are influenced by the extent to which performances fulfil normative expectations associated with certain identity aspects and demographics (Coupland 2001; Enli 2017; Kytölä and Westinen 2015; Maragh 2018; van Leeuwen 2001). In each of the accounts above, socio-political affiliation was indexed to some degree in the account metadata, either explicitly, such as in @TEN_GOP, or implicitly through the use of terminology or imagery associated with certain demographics, such as in @wokeluisa and @SouthLoneStar2, respectively. In relation to political affiliation, previous research suggests that disinformation accounts constructed performances of personas by showing ‘shallow’ knowledge of political topics (Xia et al 2019:1656). The prominence of political co-tellers and mobilisation of political discourses in the tweets analysed support Xia’s (2019) findings, but my analysis also highlights the role of consistency in performing authenticated identity performances. My research suggests that while accounts may indeed have constructed cultural competence by broad discussions of ‘shallow’ talking points, they do this consistently within their tweets and, more importantly, in ways that link coherently with identity aspects signalled in their account metadata. In each account, collective identity performances in the storying of the Charlottesville events correspond with identity aspects presented in the account metadata, contributing to the creation of consistent identity performances. The identity aspects presented across metadata units and tweet content orient to normative expectations related to American socio-political positions and can thus be taken as valid representations of

“real” American conservative and/or liberal personas when assessed ‘at-a -glance’ in a user’s news feed (cf. Morris et al 2012:444).

Chapter 7 Concluding Discussion

7.1 Introduction

Through a mixed method approach, this thesis has investigated how authenticated identities were performed on disinformation accounts, with a specific focus on how the design of the platform shaped identity performance strategies employed by users. In line with previous approaches to authenticity (Coupland 2001; van Leeuwen 2001), I approached authenticity as a dynamic process that can be performed and negotiated rather than a static quality of unquestionable origin. More specifically, I examined how users' management of both metadata information and content produced through tweeting contributed to this dynamic process. To enable this kind of investigation, I have used both qualitative methods and quantitative approaches to coding to conduct a linguistic analysis of handles and usernames, a visual analysis of profile pictures, and a narrative analysis of the storying of a high-profile news event. While research on identity performance on Twitter disinformation accounts appears to primarily attend either to metadata and platform design from computer science perspectives or to linguistic content of tweets from discourse analysis perspectives, this study brings together the study of metadata and tweets in disinformation contexts from a discourse analytic perspective.

In this chapter, I provide a summary of my findings guided by the three main research questions of this thesis (see chapter 2). I then discuss how the methodologies used in this thesis can contribute to existing research on identity performance and disinformation research. Next, I discuss the findings from this study in relation to wider debates around issues of authenticity. In particular, I suggest that an approach to authenticity as performed 'at-a-glance' may contribute to the broader study of identity performance on social media. Finally, I critically reflect on the methodological processes of this thesis and the potential directions for future research.

7.2 Summary of findings

7.2.1 Handles and usernames as means for authentication

On Twitter, identity performance is made possible not only by the practice of tweeting, but also through identifiers facilitated by the site itself. According to Jones (2020:718), the process that social media sites use to 'compel people to disclose information about themselves' can be referred to as 'pretexting'. Twitter handles and usernames can be viewed as 'pretexts', as they

constitute pre-set textual formats that users need to populate in order to have a Twitter account. In this section, I will focus on how these pretexts in the form of handles and usernames, determined by the platform itself, can be used by disinformation accounts for authenticating identity performances.

As mentioned in Chapter 4, a username is considered ‘the key piece of your identity on Twitter’ (O’Reilly and Milstein 2012:21). The data analysis showed that disinformation accounts use handles and usernames to present various identity aspects via two key strategies: eponyms, i.e., personal or ‘proper’ names that we use as identifiers in everyday life (e.g., first names like Aurora), and pseudonyms, i.e., any type of identifier other than what could conceivably be a ‘proper’ or ‘real, correct’ name (e.g., location like Cardiff) (cf. Deseriis 2013; Jäkälä and Berki 2013) . While the analysis showed that accounts draw on various identity aspects (e.g., location, political affiliation, topics of interest), they primarily opt for presenting accounts that appear as either news organisations or individual personas. In the sample that included more popular accounts (i.e., accounts that were followed by more than 10,000 users), disinformation accounts that appear as ‘news’ are found to mirror handle and username practices of pre-existing local news aggregators such as @TodayNYCity (New York City Today), while accounts of individual persons are broadly identified through eponyms in their handle and username, such as @Crystal1Johnson (Crystal Johnson).

These findings appear to corroborate Morris et al (2012)’s research that, based on user surveys, reveals an association between different kinds of naming strategies and varying levels of credibility and trustworthiness, qualities that are consistently conflated with discussions of authenticity both online and offline (Appelman and Sundar 2016; Coupland 2001:414; Cutler 2003; Mena et al 2020; Tolson 2001:455). In addition, as part of their work into automated rumour detection, Andrews et al (2016:462) suggest that accounts that mirror the practices of traditional news distributors (e.g., by including the word ‘news’ in their profiles) are influential in the dissemination of information online because their ‘mimicked’ appearances gain trust from users. On the other hand, in their research into the different possibilities for identity management facilitated by online communication, Jäkälä and Berki (2013:6,9) suggest that eponymous identifiers can be associated with higher levels of trustworthiness and authenticity than pseudonyms.

With respect to the sample of less popular accounts (i.e., accounts with between 5,000 and 9,999 followers), a range of identity aspects were also performed, but appeared to be presented differently to those with comparatively more followers. For example, in the less popular accounts, news accounts did not always orient to pre-existing norms. Instead, the use of rebus

writing and a lack of locational information deviated from the norms found in the usernames and handles of the more popular accounts. Furthermore, the accounts associated with people in this sample used naming strategies that allude to stereotypical ‘characters’ (e.g., @Jihadist2ndWife Jihadist Wife) rather than personal names that are perceived to identify specific individuals in the “real” world (e.g., @Crystal1Johnson Crystal Johnson).

These findings suggest that authentication through handles and usernames is achieved through naming strategies that orient to pre-existing norms associated with the kind of identity they are performing, namely the combination of location and news branding for ‘news’ accounts and the use of first names in accounts posing as ‘people’. In other words, authentication is influenced by norms relating to the specific context of use (e.g., the type of account).

Among the different identity aspects noted in the analysis of both the five accounts and wider samples, the accounts’ handles and usernames, location appeared to be a key identification strategy particularly among the five more popular accounts focussed on throughout the thesis and regardless of the type of account (news or people). The prevalence of location as an identity aspect within the more popular accounts challenges Jäkälä and Berki’s (2013) suggestion that a sense of familiarity between the ‘virtual and real world’ is created primarily by the use of eponyms. In addition to eponyms, pseudonymous naming strategies can also create this sense of familiarity by drawing on identification means that can be perceived as linking with the ‘physical’ world, such as location. More importantly, though, this sense of familiarity is achieved not just through the mere use of location but through the fact that the practice of identification through locational means is widely shared online. Rather than authentication being achieved solely through links with the offline, physical world, it seems that accounts can perform authenticity by orienting to and recycling expected norms that emerge in the specific contexts.

Finally, in both the more and less popular accounts, it was found that the same identity aspects were presented across each account’s handles and usernames, for example @Crystal1Johnson, Crystal Johnson. Previous research into Twitter usernames suggests it is relatively common for handles and usernames to present similar information (Olivier 2014:58-59). As a result, the coherence between naming units documented in most accounts in the dataset may demonstrate an orientation to norms established on Twitter in relation to management of identity presentation in the metadata units. Considering that ‘it is difficult, especially within an online environment, to truly separate a name from other visual information presented’ (Olivier 2014:7), the thesis also investigated profile pictures which represent the other key identifier visible together with tweets in Twitter feeds.

7.2.2 Profile pictures as means for authentication

As mentioned in chapter 5, Twitter profile pictures can be viewed as ‘one of the primary methods of visual presentation of self’, as the selection of a profile picture enables users to use ‘signs and symbols to construct meaning around one’s presence’ (Lough et al 2018:1279). Similar to naming strategies, the kinds of profile pictures associated with social media accounts can influence how credible a profile and its content is perceived to be (Morris et al 2012; van der Land et al 2016). The analysis of profile pictures in disinformation accounts showed that in most cases, profile pictures present the same identity aspects as the ones found in the handles and usernames; for example, @todayinsyria Syria Today uses a profile picture of the Syrian national flag.

Similar to handles and usernames, profile pictures also vary depending on the type of account performed: ‘news’ accounts tend to use logo-style images that mirror established news networks, while accounts belonging to people largely use photos of humans. In some cases, images of humans are used alongside pseudonymous naming strategies (such as @USA_GunSlinger GunSlinger Girl). The visual identification through such images results in partly de-anonymising and personalising an account with a pseudonymous handle or username. In all cases where pictures of people are used (together with either eponyms or pseudonyms), there is some coherence in the identity aspects conveyed across the picture and the naming units. For example, GunSlinger Girl uses a photo of a young woman, reinforcing the gender identity signalled by ‘girl’ in the username. There are no cases, either in the five accounts focussed on throughout the thesis or in the wider samples, where contradictory or conflicting information is presented across metadata units. Consistency and authenticity are concepts that often appear to be interlinked, as evident in previous research both on and offline. Identity performances that are consistent (whether that be consistency across time, across platforms, or with stereotypical expectations about different demographics) tend to be viewed as more authentic than those that are not (Coupland 2001; Enli 2017; Kytölä and Westinen 2015; van Leeuwen 2001). In the case of disinformation accounts, authentication is similarly achieved through a coherent and consistent presentation of self across the different metadata units.

While there largely appears to be coherence between metadata units across samples, there are differences between the ways in which the more and less popular accounts use profile pictures. Within the less popular account sample, there is a presence of ‘spoof’ accounts where the people depicted in images and naming strategies appear to present fictional characters, rather than presenting themselves in ways that could lead to their identification as individual

persons in the ‘real’ (Jäkälä and Berki 2013:6), or offline world. These profiles employ photoshopped images of people who do not ‘really’ exist, such as Jeb Bush and Hillary Clinton photoshopped together to make ‘@Jeblary2016 Jeblary Bushton’. In contrast, these kinds of images were not found in the more popular accounts, where identity performances of individuals predominantly use seemingly unedited images of various persons and/or eponymous naming strategies to reduce anonymity. While spoof and parody accounts can be popular on social media and arguably authentic within their own genre, in this dataset these kinds of identity performances did not attract as many followers as the others. While it is beyond the scope of this study to discuss the role that parody can play in the spread of disinformation more broadly, the lack of spoof images in the more popular accounts may be explained by the association often made between anonymity and nefarious activity (Calabrese 2020; Kozinski 2015:6).

Furthermore, the metadata units of the disinformation accounts appear to be designed in ways that orient to wider social stereotypes. For example, the ‘spoofs’ in the less popular accounts created their identity performances by exploiting extreme social stereotypes (such as an Islamophobic construction of Muslim women as Jihadists). Similarly in the five popular accounts analysed in more detail, the design of handles, usernames, and profile pictures adhered to normative expectations about certain identity aspects, ultimately indexing stereotypical personas, such as ‘traditional’ white cowboys from the south (@SouthLoneStar) or young ‘woke’ Black women (@wokeluisa). These coherent identity performances appeared to largely reinforce normative identity aspects that at some level connoted stereotypical associations with the American socio-political left or right. As acknowledged in chapter 2, previous research suggests that online identity performances can be viewed as inauthentic if they are deemed to deviate from normative expectations associated with social group membership (Kytölä and Westinen 2015). While previous research focusses more specifically on the content posted by online users (e.g., messages, tweets, etc), I suggest that the concept of normativity can also be applied to identity performances that take place via metadata units. For example, the presentation of a Texan cowboy across metadata (@SouthLoneStar) can be viewed as consistent with normative expectations related to conservative identities and beliefs. Such consistency achieves authentication via an orientation to what we broadly accept as ‘normal’, thus constructing valid representations of specific personas (e.g., a conservative persona).

7.2.3 (Co-)tellership in storytelling as means for authentication

The analysis of the storying of Charlottesville further explored the interaction between Twitter metadata and content production in performing authenticated identities on disinformation accounts. As mentioned in chapter 6, studies of online storytelling enable a ‘deeper understanding of identity processes’, with storytelling being intimately connected with ‘authentic identity performances’ (De Fina 2015:364-365). I used Georgakopoulou’s (2015) small stories heuristic to investigate the separate but interrelated levels of sites, tellers, and ways of telling, enabling a further exploration of how Twitter’s platform architecture interacts with content produced by users to ultimately facilitate authenticated identity performances. The analysis revealed that the five accounts focussed on throughout the thesis used content production practices facilitated by Twitter to co-tell stories in ways that contributed to collective identity performances, specifically in regard to political (dis)affiliation.

As mentioned above, Jones (2020:718) uses ‘pretext’ to refer to the strategies used by platforms to compel users to share information about themselves. Jones (2020:718) acknowledges that this idea is not exclusive to digital surveillance or online communication; people respond to pretexts in all interactions, in the form of ‘expectations that text producers and text consumers bring to interaction as a way of negotiating common ground’. Maryns and Blommaert (2002:11) use the phrase ‘conditions of sayability’ to refer to the same concept, with conditions being understood as the ‘practices, competencies and contextual frames that make it possible for certain people to credibly engage in certain kinds of interaction’. In terms of tweet format, users are given various options to engage with other people. As discussed in chapter 6, users can post their own tweets, repost others, reply to other users, or repost other people’s tweets with their own added comments; all these options afford different possibilities for tellers in terms of narrative roles, distribution of turns, reciprocity formats, and control of text and can thus be seen as pretextual.

When engaging in the co-tellership of a shared news event, the disinformation accounts used tweet formats that retain control over the content shared by the Charlottesville events by retweeting or reporting the speech of others. These kinds of actions can be considered less dialogic than replying or quote tweeting, where multiple users are actively engaged in the same interaction and negotiate the content shared. With respect to the identity of co-tellers, the accounts primarily engaged with ‘elite’ users (such as politicians and political commentators with large followings and verified Twitter status) that could be affiliated with explicit socio-political positions. Consequently, these disinformation accounts primarily made use of Twitter

content production practices that enabled them to indirectly signal belonging to socio-political groups, rather than to actively engage in the turn-by-turn co-telling of a story we find in offline and other types of online storytelling (cf. Dayter 2015:22, 2016:75). Engaging with stories on social media as a way of signalling (dis)affiliation with social groups corroborates findings from previous research on social media storytelling, where, as acknowledged in chapter 6, users engage with shared stories as a way of ‘making sense of themselves and the surrounding world’ (Page 2018:1). For example, Page (2018) discusses how users shared affective messages to demonstrate support and solidarity for murdered Labour MP, Jo Cox. Giaxoglou and Spilioti (2020) also examine how users engaged with the story of Alan Kurdi, a child whose body was photographed washed up on a beach after he drowned during the 2015 ‘refugee crisis’, to show sympathy for Alan Kurdi and compassion for refugees.

As acknowledged in section 7.2.2, the handles, usernames, and profile pictures used by the five case studies invoked identity aspects related to the American socio-political left or right, in more (TEN_GOP) or less (wokeluisa) explicit manner. The narrative analysis of the five accounts’ storying of Charlottesville revealed that in all cases the accounts narrated events in ways that reinforced the socio-political associations indexed in the account metadata. Similar to previous research according to which collective identity plays a key role in people’s engagement with disinformation narratives (Marwick 2018:477), I have found that the disinformation accounts’ identity performances appear centred around membership to communities affiliated with the socio-political left or right. More importantly, however, this study shows how disinformation accounts performed collective identities in ways that potentially encourage affiliation and engagement from other users (Marwick 2018:477): a) they emphasise the wrongdoing of actors from socio-political out-groups, b) they distance members of their in-group from participation in or responsibility for wrongdoing, and c) they signal insider status by creating intertextual connections with discourses associated with their socio-political in-group.

As mentioned in section 7.2.2, authenticity is connected with concepts of consistency and normativity: identity performances are more likely to be perceived as authentic when they are consistent across time and they align with normative expectations about what a person’s identity performance ‘should’ be like (Coupland 2001; Enli 2017; Kytölä and Westinen 2015; Maragh 2018; van Leeuwen 2001). My analysis of (dis)identification in tweets posted by disinformation accounts also revealed that the accounts’ storying of Charlottesville contributed to the performance of authenticated identities by narrating events in ways that appeared consistent with the normative expectations set up by identity aspects signalled in account metadata. For

example, ‘woke’ Luisa disidentified with Trump by highlighting his affiliations with racist groups and criticising his response to Charlottesville, while TEN_GOP affiliated with Trump and disidentified with the socio-political left. While previous research on disinformation accounts suggests that presentations of ‘talking-points knowledge’ can index authenticity (Xia et al 2019:1656), I offer a more nuanced perspective by arguing that consistency between the narrative tellings of these talking points and identity aspects signalled in metadata also contribute to performances of authenticated identities online.

The claim about the significance of consistency in studies of authenticity can be challenged by research into identity performances of YouTube vloggers where displays of negative emotions that foreground ‘anxiety, distress, and the performance of emotional vulnerability’ contribute to performances of the ‘true’ self as a troubled, not always coherent, persona (Berryman and Kavka 2018:86-87). In the context of personal vlogging then, it seems that reflections on the fragmented self and acknowledgements of inner conflict serve to strengthen rather than negatively impact performances of authenticity. Nevertheless, authenticity may be achieved through different means in these different contexts. For example, personal vlogs serve to document a person’s life, whereas Twitter accounts (in the form of either news organisations or individual accounts) appear to primarily function as a means of sharing and keeping up to date with news and current affairs (Rosenstiel et al 2015; Twitter 2022d). For this reason, Vrikki (2020:83) describes Twitter as a ‘commentary pit’ for public discussion of ongoing events. While personal vlogs typically have large followings of supporters and subscribers that keep up to date with their content, these Twitter disinformation accounts operate differently. Considering that disinformation narratives do not need to be internalised completely to influence decision-making (see chapter 4), it is likely that the IRA operatives’ primary aim was not to gain loyal supporters but to have their messages viewed by as many people as possible. Consequently, the kind of in-depth, somewhat fragmented, identity performances observed in personal vlogs may not be relevant to the context of politically divisive disinformation accounts, where authenticity may be assessed at-a-glance based on the information available within a tweet: tweet content, name, and profile picture. In this context, a consistent display of orientation to the same political cause appears to contribute to the construction of a valid representation of a persona engaged with either the socio-political left or right.

7.3 Methodological contribution

In this section, I will discuss the methodological contribution that this research makes to the study of identity authentication online and particularly in disinformation contexts. Firstly, my

analysis pays attention to platform design and its affordances for identification. By doing so, this research shows how visual identification strategies can be analysed together with textual content and attempts to fill a gap in existing research by investigating authenticity through a multimodal lens of both platform design and user-generated content. Moreover, I will discuss how qualitative narrative analysis can be employed for research on identification and authenticity in disinformation online. Finally, I will show how qualitative and quantitative approaches can be combined in a disinformation setting.

As acknowledged in chapter 2, previous research on identity performance in disinformation accounts has primarily focussed on user-generated content, particularly in the form of tweets. Linvill and Warren (2020), for example, conduct a large-scale content analysis of IRA accounts in order to establish broad categories of profiles, while Xia et al (2019) focus specifically on the construction of an authentic persona within a particular IRA account. Both studies draw on tweet content to provide valuable insights into the broad themes and topics tweeted about by right-wing personas. In both studies, little attention is paid to the contribution that pre-determined and platform-related options for identification can have upon authenticated identity performances. The profile picture and naming strategies employed by @Jenn_Abrams, for example, are not discussed in relation to her authentic identity performance. Similarly, although Linvill and Warren (2020) highlight that Twitter handles can reveal information that can aid understanding of their broad categorisations, no attention is paid to how visual elements may also contribute to account identification or categorisation.

Although both aforementioned studies acknowledge that disinformation accounts employed various tweet formats such as original tweets, retweets, and replies, tweet formats are predominantly discussed in relation to broad engagement and dissemination patterns, particularly in relation to the growth of followers. It therefore appears that in previous disinformation research, accounts' management of pre-determined co-tellership opportunities is largely understood in terms of their potential for profile boosting rather than identity performance as particular personas. This thesis therefore prioritises areas of Twitter communication that have previously been overlooked in terms of their potential to contribute to identity performances in a disinformation context.

To provide a comprehensive understanding of how the communicative ecology of Twitter constitutes a space for identity performance in disinformation accounts, this thesis focusses on different areas of platform design that enable identification: naming strategies, profile pictures, and co-tellership affordances. In early digital discourse analysis, it was noted that 'computer-mediated discourse may be but is not inevitably shaped by the technological features of

computer-mediated communication systems' (Herring 2004b:343). Herring (2004b:343) argues that 'it is a matter for empirical investigation in what ways, to what extent, and under what circumstances computer-mediated technologies shape communication that takes place through them'. In more recent research, Jones (2020:710) clarifies that the interaction between social relationships, discourse practices, and technological tools has created a 'communicative ecology'²⁵ on social media where nearly all social interactions are engineered to produce data of maximal value for internet companies. In other words, Jones (2020:710) suggests that the technological design of social media sites create certain discourses that influence the way that users navigate the platform and share information about themselves. Consequently, Jones' (2020) approach foregrounds more the relevance of platform affordances upon communication nowadays, compared to the early days of CMDA research.

In this thesis, I follow Jones's approach (2020) and pay attention to aspects of platform design and the ways in which discourses are taken up by users to perform identity. In order to do this, I oriented to aspects of platform design in my study of both discourse produced in the metadata units of a Twitter profile and discourse produced in the tweets appearing on the Twitter feed of the five key disinformation accounts. In chapters 4 and 5, I focussed specifically on how the pretextual metadata units of handles, usernames, and profile pictures contribute to the performance of authenticated identities. In chapter 6, I analysed how platform-facilitated opportunities for co-tellership are managed by users to (dis)identify with other users and wider social communities. This thesis therefore shows how pre-determined elements of a platform's architecture can be taken into account and analysed as integral elements of identity performance on social media, as opposed to being approached as the accompanying and/or external context to users' interactions.

My study also contributes to existing disinformation research by demonstrating how authenticity can be investigated through a multimodal lens. By drawing on models of visual modality and visual authenticity, this thesis proposes a way to further explore and revisit disinformation practices that are currently discussed in the mainstream media and amongst social media users who may be vulnerable to disinformation narratives. In both the mainstream media and academic literature, discussions of profile pictures on social media tend to orient to issues of 'origin'; i.e., a picture is considered inauthentic because it is 'stolen' from another user or is 'anonymous' (Bovet and Makse 2019; Calabrese 2020; Im et al 2020; Starbird 2017). Although I have investigated the objects and persons represented in images and their online

²⁵ Defined by Foth and Hearn (2007:756) as 'the context in which the communication process occurs'

history, I have also explored how images are employed for their potential to evoke sociocultural connotations, and how the identity aspects signalled in the images interact with those signalled in other metadata units. Rather than images being viewed as either anonymous/not, and thus authentic/not, I approach the visual as one among a range of resources that contribute to dynamic identity performances. Models of visual analysis are therefore drawn on to inform an understanding of how images contribute to the construction of valid representations of personas within particular communities (e.g., depending on the extent to which they draw on shared community practices), rather than if they have (in)authentic origins.

The thesis also demonstrates how existing models of narrative analysis can be used to approach processes of identification and authentication on disinformation accounts. Previous studies of disinformation have provided valuable insights into how disinformation accounts polarise debate by using content analysis to identify key themes and hashtags (Arif et al 2018; Linvill and Warren 2020; Xia et al 2019). Although these studies acknowledge broad tweeting patterns of the accounts (Linvill and Warren 2020; Xia et al 2019) and demonstrate how the IRA used tweet formats to build polarised clusters and networks (Arif et al 2018), these patterns are primarily explored in terms of network building, disinformation diffusion, and the targeting of particular communities. In contrast, while collective identity has been highlighted as playing a key role in people's engagement with disinformation narratives (Marwick 2018:477), the potential of investigating co-tellership as a means for (dis)identification with others appears relatively underexplored within the field of disinformation research. In this study, the application of the small stories heuristic enables the investigation of 3 interrelated levels: sites, tellers, and ways of telling. By orienting to the level of sites, discourse analysts can acknowledge and attend to the symbiotic relationship between platform affordances and social practices as highlighted by Jones (2009:116). This is particularly evident in the stage of analysis that focusses on tweet formats and their facilitation of co-tellership, where I focussed on exploring how tweet formats were used to achieve (dis)identification with different communities. The use of narrative analysis to zoom in on instances of multiple tellership thus provides an insight into the identity performance strategies used by disinformation accounts when engaging with other users. By using this narrative lens, this thesis shows how the management of tweet formats can be investigated from a qualitative, discourse analytic perspective, rather than solely through the lens of network mapping and IRA-account coordination.

As acknowledged previously, studies of disinformation have largely used quantitative methods to gain insights into the general practices of disinformation accounts. While

quantitative methods can be fruitful, there is also the chance for some finer details to be missed (Brunner and Diemer 2022:199). On the other hand, qualitative methods can use contextual knowledge to provide detailed insights into patterns, including cases that may not initially appear to be numerically significant (Hinrichs 2016:30). Throughout this thesis, I have shown how both quantitative and qualitative approaches can be used to gain an insight into large social media datasets.

For example, in chapters 4 and 5, I initially conducted a close-up qualitative analysis of the identification strategies used by the five case studies. I then contextualised these findings by using the qualitative insights to identify categories of specific features that were counted and interrogated using descriptive statistics. This combination of methods enabled me to discover additional categories that did not emerge within the initial qualitative analysis, such as the news-branded accounts that appear frequently in the sample of more popular accounts but were not represented among the five case studies. In other words, the use of descriptive statistics in chapters 4 and 5 has enabled me to make more general observations about authenticity on Twitter disinformation accounts, including the realisation that an orientation to pre-existing norms appears to have a relationship with popularity, and thus, potentially, authenticity.

The approach taken in chapter 6 involved an initial frequency analysis of co-tellership practices for a broad overview of tellership patterns, followed by qualitative narrative analysis to gain a more in-depth understanding of disinformation accounts' use of (co-)tellership and (dis)identification patterns in each of the five accounts. By using a qualitative lens, it was also possible for me to probe further into themes highlighted via quantitative methods in other studies. For example, Xia et al (2019) carry out a large-scale analysis of 3,878 tweets by one account. Their study enabled the identification of broad themes that appeared throughout the account's content, specifically, construction of personal authenticity (via performing as an American girl with a likeable personality) and performance of cultural competence (via displays of 'talking points knowledge' of political news topics like the Black Lives Matter movement, feminism, and immigration). In this study, I have used qualitative narrative analysis to provide a detailed insight into how performances of socio-political (dis)affiliation are navigated within the disinformation profiles, showing the various ways that accounts (dis)identify with social groups when engaging in discussion of a particular politically charged news story.

To sum up, this study highlighted a gap in pre-existing methodological approaches to disinformation and identity; while research suggests that both metadata and tweet content can contribute to an account's perceived credibility on Twitter, relatively little research in the field of disinformation brings linguistic and visual analysis together from an identity performance

perspective. By paying attention to the discourses shaped by platform affordances (such as the need to identify oneself and the possibility to interact with other users through specific message formats), this study shows how it is possible to situate research on disinformation, language, and authenticity within the communicative ecology of social media sites. Finally, by combining qualitative methods with quantitative approaches to coding (e.g., via the use of descriptive statistics), this thesis attempts to address some of the challenges of research on large social media datasets. Observations of general patterns can be coupled with detailed case studies to provide a comprehensive understanding of authenticity on social media. This approach could be used to investigate disinformation, language, and authenticity on other social media platforms aside from Twitter.

7.4 Theoretical contribution

In addition to the aforementioned methodological insights, this thesis also attempts to contribute to wider discussions of disinformation and authenticity online. First, I explain how this thesis's findings can be used to inform discussions of disinformation that may be limiting the phenomenon to 'bot' behaviour and automated content not only in academic research but also in its application in the context of media literacy and disinformation education (section 7.4.1). Then, in section 7.4.2, I reflect on how approaches to authenticity can be expanded to take into account the discourse shaped by platform-specific design technologies like identification metadata. Finally, in section 7.4.3, I propose 'authenticity at-a-glance' as a particular kind of authenticity that can be constructed via an orientation to the kinds of identification norms outlined in section 7.4.2.

7.4.1 Identifying disinformation

Within mainstream media, discussions of digital literacy tend to advise users to be cautious of certain features of Twitter profiles that may indicate involvement in the spread of disinformation: 'look out for subtle typos, for tweets being sent out very quickly, or for profile pictures and usernames that appear suspicious' (Allyn 2020). These kinds of discussions often foreground the existence of 'bots', defined as automated activity. Scholarly research has also investigated the tweeting behaviour of automated bot networks, including 'ultra-high posting rates, high retweet-to-tweet ratios, repeating variations of the same tweet over and over, suspicious names, and lack of personally identifiable information' (Hindman and Barash 2018:30). While these findings can be useful when trying to provide users with quick and easy tools for spotting potentially nefarious actors, this thesis highlights that Twitter disinformation profiles do not all revolve around automated accounts that share these kinds of 'obvious' clues.

Instead, disinformation accounts are also found to use photos of humans and ‘identifiable’ names or appear as news accounts using news-branded language and typical logo-style profile pictures.

The accounts’ orientation to Twitter norms, as evident in this study, suggests that approaches to disinformation that focus on identifying inauthentic activity in terms of a specific set of marked features, like repetition of the same tweet and lack of personally identifiable information, may not be as straightforward, or necessarily helpful, as they first appear. While the ability to quickly “spot” bad actors may seem appealing, this thesis demonstrates that disinformation accounts perform identities that can be interpreted as authentic at-a-glance based on multimodal cues that orient to normative expectations relating to Twitter use and wider social stereotypes. Although norms may at certain points in time be connected to the use of certain features, it is important to emphasise that normative expectations are co-constructed, context-dependent, and emergent, while marked features are somewhat static and, thus, may be less useful in the context of mediated communication where technologies are rapidly advancing and communities, together with their practices, are constantly being (re)shaped (Jones et al 2015:2). As a result, what may be important is not what features are utilised by users, but how users utilise them to construct meaning and co-construct normative expectations. While a marked feature may be associated with suspicion or disinformation at one point in time, there is no guarantee that such associations will remain, or be valid in other contexts of disinformation.

Furthermore, the orientation to social stereotypes throughout the disinformation accounts studied in this thesis suggests that accounts and their disinformation activity tend to be situated in specific sociocultural contexts. In order to understand how disinformation accounts authenticate certain personas and spread disinformation, then, we need to examine not only computerised algorithms and automated activity but also the norms and practices of the communities that the disinformation accounts infiltrate and the wider sociocultural reality at a given moment in history. Consequently, rather than users being encouraged to look out for certain features as explicit markers of (in)authenticity on Twitter, it may be more appropriate to invite users to critically evaluate why and how accounts may be constructing valid representations of particular identities, especially in interaction with broad social stereotypes.

This realisation may be particularly useful for media literacy or disinformation education. In fact, Stoddard et al (2021:56) have already pointed out that following the 2016 US election, much attention has been directed at encouraging young people to evaluate source accuracy and to use mainstream media sources over unverified information from the internet. While this may

initially appear helpful and in line with the prioritisation of ‘spotting’ certain features of disinformation actors, Stoddard et al (2021:56) argue that this approach to disinformation overlooks the possibility for harmful stereotypes to be circulated and reinforced within the mainstream media: ‘being factually accurate does not stop the perpetuation of stereotypes (e.g., race, gender) or harmful narratives in these sources (e.g., xenophobia)’. Instead, Stoddard et al (2021) propose an approach to digital media literacy education that encourages the critical analysis of representations of certain groups regardless of their source. This kind of education could limit the harmful impacts of disinformation narratives by prioritising an approach that targets the message over the source, particularly in an environment where news messages tend to trade on appeals of collective identity and group belongingness, as well as emotions and personal beliefs (Freelon and Wells 2020; Rochlin 2017). If users are aware of how stereotypes offer ready-made templates that can be exploited to cause harm even in disinformation contexts, there is hope that the risk of disinformation narratives that mobilise and embed such stereotypes in given sociocultural contexts may also be reduced.

7.4.2 Keepin’ it real on social media

Throughout this thesis, I explore how disinformation accounts perform authenticity on Twitter. In Chapter 2, I discussed various definitions and approaches to authenticity and acknowledged that essentialist approaches to authenticity on social media consider an online persona as authentic if there is a seamless correspondence to an offline embodied person with the same attributes. In line with Kytölä and Westinen (2015)’s recommendation to approach authenticity as a dynamic discursive process, rather than a static concept of ‘unquestionable realness’, I adopt an approach to authenticity as a discursively constructed representation of realness that is considered valid in specific contexts and for specific communities. Drawing on van Leeuwen’s (2001) conceptualisation of authenticity, attention to linguistic and other semiotic cues is required in order to understand how representations may have been ‘taken’ as valid within particular communities. Furthermore, the use of cues that can lead to judgements of validity is associated with the notion of consistency, both in terms of a person’s consistent identity performance over time and in terms of consistency with normative practices and expectations (Coupland 2001; Cutler 2003; Kytölä and Westinen 2015; Thornborrow 2001; Tolson 2001; van Leeuwen 2001).

The findings of this study suggest that the five popular disinformation accounts performed authentic identities by recycling and reinforcing normative expectations associated with the types of accounts and personas they constructed via their metadata and tweet content. I have

argued that by orienting to (stereo)typical expectations, the accounts legitimised and validated their identity performances as ‘authentically’ conservative or liberal. Considering the specific context of Twitter, the accounts also oriented to discursive norms related with two distinct types of Twitter profiles: ‘news accounts’ or ‘personal accounts’. Previous research supports the idea that context plays a significant role in perceptions of authenticity amongst users. For example, based on the ‘outing’ of a Christian mommy blogger as being ‘inauthentic’ due to violating expectations of a ‘truthful representation of the blogger and her story’, Whitehead (2015:141) suggests that the rhetoric of authenticity is productive of communities of practice, where communicative norms play a key role in attributions of (in)authenticity.

In the context of the Twitter disinformation accounts in my sample, politically sensitive communities of practice on Twitter, as well as the established norms and behaviours of such ‘politically active communities’ (Freelon and Lokot 2020:2), have been found to play a role in understanding authenticity. While in some contexts, like personal blogging, criteria for authenticity may be situated in the expectation for ‘honest and complete self-disclosure about everyday life’ (Whitehead 2015:129), in the context of political commentary, authenticity as a discursive construct appears to be related more closely to how consistently an account constructs a persona affiliated with a particular socio-political position. This understanding of authenticity consequently problematises broad claims that certain kinds of profile pictures (e.g., photographs of humans) are deemed as more trustworthy and authentic than others (Calabrese 2020; Dotto and Cubbon 2019). For example, in the context of political commentary, the use of cowboy imagery by @SouthLoneStar was used to index an identity that contributed to the construction of a valid representation of a conservative persona. In other words, whether @SouthLoneStar provides a ‘truthful’ depiction of their physical self appears secondary to whether @SouthLoneStar constructs a persona that appears authentically committed to conservative politics.

The orientation to political affiliation within IRA disinformation accounts is also underlined by Xia et al (2019) who found that @Jenn_Abrams authenticated the persona of a conservative woman by displaying cultural competence. This was achieved through the account’s tweets that focused on points of high-profile discussion in the US (such as immigration and feminism) and adopted dominant conservative discourses. This thesis corroborates and elaborates on these findings by highlighting that accounts can evoke cultural competence not only through their tweets but also through their metadata units. For example, one of the key findings in my work has been that the accounts created unified identity performances as conservatives or liberals by

orienting to politically affiliated identity aspects across handles, usernames, and profile pictures and in ways that oriented to existing socio-political discourses and stereotypes.

7.4.3 Authenticity at-a-glance

Morris et al (2012:444) argue that users' 'ability to judge credibility in practice is largely limited to those features visible at-a-glance in current user interfaces (user picture, username and tweet content)'. Drawing on their claim, I argue that disinformation accounts perform a particular type of authenticity which can be called 'authenticity at-a-glance'. As I have shown, the construction of authentically conservative or liberal personas was achieved primarily through the consistent deployment of features and practices across the units that are visible at-a-glance, i.e., usernames, profile pictures, and tweet content. At the same time, this consistent identification with the political right or left also made sense 'at-a-glance', i.e., based on the audience's (stereo)typical expectations regarding the appearance and behaviour of different social groups, the cues provided in metadata and tweet content contributed to the construction of valid representations of American personas. For example, the presentation of a conservative Texan cowboy taps into an existing stereotype that makes sense to the target audience (Americans). A Twitter user who scrolls through their feed can glance at the name, profile picture, and tweet content proclaiming to support Trump and think 'it could be any one of Donald Trump-supporting Americans' (Hern 2017).

The idea that at-a-glance authenticity may be at play in disinformation contexts is reinforced by the suggestion that disinformation narratives do not need to be internalised to the point of complete trust to have an impact on decision-making (Bastick 2021; Canan and Warren 2018). So, while Xia et al (2019) have shown that authenticity created over time via the creation of a 'likeable persona' who is politically engaged is significant in disinformation contexts, I suggest that at-a-glance authenticity may be equally influential.

Authenticity at-a-glance was also achieved in the disinformation accounts through their orientation to norms associated with ordinary Twitter use, rather than features largely regarded as suspicious like handles using long series of numbers, or default profile pictures (BBC News 2018b; Barojan 2018; Dotto and Cubbon 2019). By constructing types of personas that used Twitter in ways that we expect from ordinary users, the disinformation accounts employed communicative strategies that could be taken as valid within the context of Twitter, ultimately enabling them to achieve their purpose: to exacerbate social tensions by sharing politically divisive information. This idea, that authenticity at-a-glance can be understood in part in terms of how users construct performances that orient to established norms, builds on Jones' (2009,

2013, 2020) approach to ‘sites of display’; I argue that performances of authenticity at-a-glance are situated within the specific contexts of the platform on which they are created, where the affordances and constraints of platforms contribute to what kinds of practices are normalised and accepted (and can thus be taken as valid representations of ‘real’ users of particular sites) and which are not (cf. Jones 2009:116).

As a concept, authenticity at-a-glance applies to Twitter due to the way that the site is engineered and taken up by users; there is a constant stream of information that users can see in their news feeds, in trending topics, and in hashtagged discussions. As mentioned previously, when viewing these tweets, users’ ability to assess tweet credibility is limited to what is visibly accessible in a news feed (Morris et al 2012:444). This suggestion is reinforced by Morris et al’s (2012:444) finding that features that do not appear in a news feed (such as bios) have little impact on credibility assessments. As a result, the Twitter layout directs users’ attention less to an account’s overall profile and more on what users can ‘glance’ at when scrolling through streams of tweets. The overall layout as a constantly updated stream of tweets cannot be viewed in isolation from the marketing of the platform as a means of keeping up to date with what is happening in the world; for example, when visiting the page without an account in 2016, users would see the message ‘Get the latest updates on topics that interest you. And watch events unfold in real time from every angle’ (Twitter 2016). Indeed, Twitter is increasingly used as a news source, with 69% of U.S. Twitter users saying that they get news from Twitter (Mitchell et al 2021). At the same time, the non-reciprocal connections afforded by Twitter (e.g., users not needing to follow one another to access content) facilitate the use of this platform for sharing up-to-date information from various sources, rather than the curation of close-knit personal networks that a person may be familiar with in other contexts. As a result, authenticity performances that are achieved, as well as validated, at-a-glance may be particularly relevant within the specific communicative landscape.

The extent to which authenticity at-a-glance can be applied to other social media contexts remains to be explored. For example, the platform architecture of Facebook as a site that connects ‘friends’ through reciprocal online connections reveals an assumption that networks are made of people who know one another on a more personal level. Moreover, Facebook profiles contain photo albums, as well as extra identification options linked to relationship status, school, and work, so at-a-glance authenticity judgements may be less applicable compared to Twitter, where users have a more limited number of cues. Having said this, one could argue that the notion of authenticity at-a-glance, particularly from the perspective of authenticity claims made quickly by orienting to normative or stereotypical expectations, may

still be relevant to Facebook research. According to Van der Nagel and Frith (2015), ‘Facebook explicitly states in its user policies that people are expected to have one identity and use legal names and photographs of their own body on their profile’. Consequently, the expectation for certain kinds of names and photos to be used lays the foundation for authenticity at-a-glance to remain potentially useful, particularly in instances where users access profiles that they are not already familiar or connected with.

On other platforms, the concept of authenticity at-a-glance may be less readily applied. For example, on sites like Reddit, users tend to be ‘nameless’ and ‘faceless’ and participate in a ‘deep vernacular web’ that is not understandable to non-users (De Zeeuw and Tuters 2020: 215-217). Unlike other sites, reddit users can utilise ‘throw away’ accounts, i.e., accounts that are used for a short period of time and then discarded to maintain anonymity (De Choudhury and De 2014:78) and to discuss sensitive topics. In other words, in sites where anonymity is normalised, authenticity at-a-glance based on identification metadata units and orientations to normative expectations about existing social groups offline may be less relevant.

This discussion highlights that expectations associated with different sites of display (cf. Jones 2009, 2013, 2020) result in different cues that can be taken as valid, and thus, different perceptions of (in)authenticity. Although the notion of authenticity at-a-glance may be more applicable to certain platforms (e.g., those where there are non-reciprocal connections, where ‘verification’ is valued), the overall approach to authenticity in this thesis, as a dynamic process constructed through performances that are consistent with platform and social norms, could be used to study authenticity on other sites.

7.5 Critical reflection: limitations and avenues for future research

This study draws on data that were produced on Twitter in 2016, ten years after the platform was created. At the time of data collection (2018), discussions of online disinformation were beginning to rise into public consciousness and receive increased academic attention. Since then, both the technological architecture of Twitter and public debates about online information sharing have evolved. This section will discuss the limitations of the current study and point to future research avenues in examining disinformation in the ever-changing social media environments. More specifically, I will discuss the limitations of using screen data to investigate authenticity and the implications of focusing on Anglo-American contexts of disinformation. I will also reflect on how the current (2022) context of Twitter may influence the relevance of some of the findings from this study. Finally, I will highlight how the approach to authenticity

adopted in this thesis may be beneficial to domains of disinformation outside of the US election context, such as within discussions of health-related topics.

This study uses screen-based data, i.e., data that is solely produced and collected online without direct interaction with participants (Androutsopoulos 2014:76) to explore performances of authenticity on Twitter disinformation accounts. Although this method provides an insight into how the accounts were presented during their activation, this study limits its exploration of authenticity to what is primarily available on screen. In other words, the analysis focuses on the meaning potential of specific linguistic and visual choices that are accessible through the texts posted by the account holders. Although it is possible to make inferences about the meaning-making potential of these choices and their contribution to constructions of authenticity, the extent to which the followers of these accounts made the same meaning and authenticity inferences is unknown. For that reason, future research into disinformation identity performance could draw on reception studies that involve interviews with ordinary users. Although users may engage in over-reporting or under-reporting how they engage with disinformation accounts (Guess 2015; Guess et al 2019a, 2019b), their interview narratives could still provide useful insights into ‘who takes this as authentic and who does not?’ and on the basis of which cues these judgements are made (van Leeuwen 2001:396). Drawing on user-based data rather than screen data in a disinformation context could also involve interviews with producers of disinformation texts. While such research would be very difficult practically and raise ethical concerns, such interactions could shed light on the intentions of text-producers; while researchers can make inferences about the disinformation dissemination process, the intentions and strategies of disinformation producers cannot be assumed.

As highlighted in this study, disinformation discourse appears to be situated within ‘distinct cultural contexts’, where processes of information production and judgement are ‘socially located and not merely the result of individual decisions’ (Wang and Song 2020:1547; see also Fine 2007). This study investigates data from an Anglo-American context, as the tweets are all written in English and engage with the U.S. socio-political context. In fact, disinformation research is characterised by a western Anglocentric dominance, as the majority of existing studies focus on U.S. and European contexts (Nguyễn et al 2022; Wang and Song 2020). Although research largely focusses on these contexts, disinformation activities are not specific to these countries only. Nguyễn et al (2022) highlight how disinformation is a prevalent issue within diasporic and multilingual contexts in Asian countries (such as Vietnam, Taiwan, China, and India) and argue that an assumption of western generalisability leads to historically

marginalised communities being ‘othered’ and ultimately overlooked within disinformation discussions. Considering that existing research demonstrates that disinformation tactics and prevalent discourses vary across countries and languages even within western contexts (Humphrecht 2019), it is necessary for future research to explore further how disinformation tactics may also manifest in non-western, non-English-speaking, and non-white communities (Nguyễn et al 2022:8). Examining disinformation situated in contexts beyond U.S. and Europe will, firstly, prevent historically marginalised communities from being overlooked in initiatives to understand and tackle online harm and, secondly, enable a more thorough understanding of disinformation processes in ways that will not reinforce an existing colonial bias and legacy within academic disinformation research (Nguyễn et al 2022).

Disinformation narratives and identity performances are situated not only within specific socio-cultural contexts but also within the technological platform of the site on which they are posted. Social media sites are constantly changing in terms of both the technological affordances of the sites and the ways that people use them, as certain platforms become ‘old’ and obsolete and previously common behaviours become outdated and ‘uncool’ (Contreras 2021: para 1). This study focusses on the social media site of Twitter, using data from a disinformation campaign that was largely active between 2016-2017. Since the activity of these accounts several years ago, disinformation has risen into public consciousness, the IRA campaign has received lots of media attention, and Twitter has incorporated technological tools to deal with the spread of disinformation. In response to disinformation relating to the COVID-19 pandemic, for example, Twitter introduced warnings to tweets that contained ‘disputed or misleading information’ (Roth and Pickles 2020: para 1). These warnings ‘provided additional context and information’ to prevent the spread of COVID-19 related disinformation and Twitter relied on ‘trusted partners to identify content that is likely to result in offline harm’ (Roth and Pickles 2020: para 6). Similarly, services like ‘birdwatch’, where users can provide informative context to tweets that they feel are misleading, have been available for limited testing to small groups in the US (Coleman 2021). As a result of the ongoing changes to the platform, it would be useful to examine whether the strategies highlighted in this thesis are still relevant to the site which has recently paid more attention to how news stories are taken out of context for the purpose of exacerbating social tensions.

Some of these recent developments, for example, may have implications for the co-tellership patterns discussed in this thesis. During the 2020 US election, Twitter made changes to the options given to users in terms of engaging with other users and co-telling a shared story. The platform prompted users to quote tweet rather than retweet other people’s content in an attempt

to ‘encourage users to consider why and what they were adding to the conversation’ (Gadde and Beykpour 2020: para 11). These changes led to a 23% decrease in retweets and a 26% increase in quote tweets (Gadde and Beykpour 2020), suggesting that users’ behaviours can be influenced by changes in the platform’s design, especially if taken up by other users. As a result, based on these reports, the (dis)identification practices discussed in chapter 6 may be achieved through different co-tellership patterns, as quote tweets gain reportedly more ground compared to retweets.

Furthermore, at the time of writing (2022), Twitter has recently had a change in ownership. While, as mentioned before, Twitter has implemented changes to information sharing by increasing censorship of hate speech and encouraging users to critically engage with material, there are now proposals to remove these features under the new management. For example, users who have previously been banned from the platform for violating Twitter’s rules, such as sharing hateful content and spreading mis/disinformation, have had their accounts reinstated in the interest of ‘free speech’ (Alfonseca 2022; Ivanova 2022) and Twitter has dropped its COVID-19 misinformation policy, where tweets were censored and labelled when deemed to contain misleading information that could pose harm to public health (Devereux-Evans and Morrison 2022). At the same time, the platform has also changed its approach to profile verification. Currently, it is possible for users to become verified either by using existing methods relating to high-profile status and public interest or by paying for verification via a subscription²⁶. The discussion of potential changes to the platform, in relation to both increased/decreased censorship and attention to disinformation spreading, demonstrates that the focus on platform design in combination with users’ practices remains an important lens for future research on authenticity in disinformation contexts.

Finally, the data analysed in this thesis demonstrates the relevance of the construction of consistent political personas and politically sensitive communities of practice in a specific disinformation campaign, raising the question of what other personas or identities may be discursively relevant in other domains of disinformation. While there has been much research on COVID-19 as a sphere of disinformation activity (Jemielniak and Krempovysh 2021; Nogara et al 2022; Scannell et al 2021), these studies tend to focus on predicting and detecting disinformation, rather than investigating the role that identity performance and authenticity may play in the dissemination of health disinformation narratives. For example, exploring what

²⁶ As previously acknowledged in section 2.5, discussions of identity verification are ongoing at the time of writing. The verification options available to users, along with specific details about how these options work, are constantly evolving and changing.

kinds of strategies are used to construct valid representations of personas with ‘real’ health-related knowledge may reveal the role that expertise plays in health-related disinformation narratives.

7.6 Conclusion

In this thesis, authenticity is approached as a dynamic process that can be performed and negotiated. While in some online contexts, authenticity appears linked with aspects of the person’s offline life (Jäkälä and Berki 2013; Enli 2017; Lovelock 2016), this thesis argues that an approach to authenticity as a reflection of a person’s offline life is less relevant in a politically driven disinformation context and, especially, on a public-facing platform like Twitter. Instead, this thesis suggests that authenticity is more related to how authentically a particular persona with a given socio-political stance and ideology can be portrayed to an audience; whether or not an account seems to mirror an offline identity is potentially less significant than whether an account seems to be authentically committed to a shared political cause (in this case).

I demonstrated that in line with an approach to social media platforms as ‘sites of display’ (cf. Jones 2009, 2013, 2020), performances of authenticity online are embedded within the affordances and communicative norms associated with platform in question. I showed how metadata like naming strategies and profile pictures can contribute to performances of authenticity. The analysis of narrative further highlighted that there tended to be consistency across the identity aspects presented in metadata and in the narratives used to frame a high-profile news event. I therefore suggested that disinformation accounts contributed to authenticated identity performances by creating particular personas that could be interpreted as consistent and therefore authentic ‘at-a-glance’. By continuing to develop our understanding of the ways that disinformation accounts perform authenticity, hopefully we can continue to evolve in our approaches to both researching and responding to online disinformation.

References

- Åberg, E., Koivula, A. and Kukkonen, I. (2020) A feminine burden of perfection? Appearance-related pressures on social networking sites *Telematics and Informatics* 46 pp. 1-11
- Adams, C. (2020) *Not all black people are African American. Here's the difference.* Available at: <https://www.cbsnews.com/news/not-all-black-people-are-african-american-what-is-the-difference/> [Accessed 17/11/2020]
- Alfonseca, K. (2022) *Hate speech increased on Twitter after Elon Musk takeover, study finds* Available at: <https://abcnews.go.com/Business/hate-speech-increased-twitter-elon-musk-takeover-study/story?id=92445797> [Accessed 23/11/22]
- Allyn, B. (2020) *Researchers: Nearly Half Of Accounts Tweeting About Coronavirus Are Likely Bots* Available at: <https://www.npr.org/sections/coronavirus-live-updates/2020/05/20/859814085/researchers-nearly-half-of-accounts-tweeting-about-coronavirus-are-likely-bots> [Accessed 14/7/22]
- Amaral, M. and Monteiro, M. (2002) To be without being seen: Computer-mediated communication and social identity management *Small Group Research* 33(5) pp. 575-589
- Andersen, K., Shehata, A. and Andersson, D. (2021) Alternative News Orientation and Trust in Mainstream Media: A Longitudinal Audience Perspective *Digital Journalism* pp. 1-20
- Andrews, C., Fichet, E., Ding, Y., Spiro, E. S., and Starbird, K. (2016) Keeping Up with the Tweet-dashians: The Impact of 'Official' Accounts on Online Rumoring *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing* pp. 452-465.
- Androutsopoulos, J. (2014) Computer-mediated Communication and Linguistic Landscapes in Holmes, J. and Hazen, K. (Eds.) *Research Methods in Sociolinguistics: A Practical Guide* Sussex: Wiley Blackwell pp. 74-90
- Antony, M. G. and Thomas, R. J. (2010) 'This is citizen journalism at its finest': YouTube and the public sphere in the Oscar Grant shooting incident *New Media & Society* 12(8) pp. 1280-1296
- Appelman, A. and Sundar, S. S. (2015) Measuring message credibility: Construction and validation of an exclusive scale *Journalism and Mass Communication Quarterly* 93(1) pp. 59-79
- Arif, A., Stewart, G. L. and Starbird, K. (2018) Acting the Part: Examining Information Operations Within #BlackLivesMatter Discourse *Proceedings of the ACM on Human-Computer Interaction* 20 pp. 1-27
- Armstrong, C. L. and McAdams, M. J. (2009) Blogs of Information: How Gender Cues and Individual Motivations Influence Perceptions of Credibility *Journal of Computer Mediated Communication* 14(3) pp. 435-456
- Arnesson, J. (2022) 'Endorsing a dictatorship and getting paid for it': Discursive struggles over intimacy and authenticity in the politicisation of influencer collaborations *New Media & Society* 0(0) pp. 1-7

- Association of Internet Researchers (2019) *Internet Research: Ethical Guidelines 3.0* Available at: [IRE 3.0 - final-includes missing reference \(aoir.org\)](#) [Accessed 19/8/22]
- Attenborough, F. T. (2011) Complicating the sexualization thesis: The media, gender and 'scandandy' *Discourse & Society* 22(6) pp. 659-676
- Ausserhofer, J. and Maireder, A. (2013) NATIONAL POLITICS ON TWITTER: *Structures and topics of a networked public sphere* *Information, Communication and Society* 16(3) pp. 291-314
- BAAL (2021) *Recommendations on Good Practice in Applied Linguistics* Available at: [BAAL-Good-Practice-Guidelines-2021.pdf](#) [Accessed 17/6/22]
- Bakhtin, M. M. (1981). *The Dialogic Imagination: Four Essays*. Austin: University of Texas Press. Pp. 269-434.
- Barojan, D. (2018) *How to Identify Bots, Trolls and Botnets* Available at: <https://gijn.org/2018/11/05/how-to-identify-bots-trolls-and-botnets/> [Accessed 18/1/21]
- Barthes, R. (1977) *Image, Music, Text* London: Fontana
- Bastick, Z. (2021) Would you notice if fake news changed your behaviour? An experiment on the unconscious effects of disinformation *Computers in Human Behaviour* 116 pp. 1-12
- BBC News (2017a) *How Russian bots appear in your timeline* 14 November. Available at: <https://www.bbc.co.uk/news/technology-41982569> [Accessed 10/3/2019]
- BBC News (2017b) *Trump defends 'beautiful' Civil War statues* 17 August. Available at: <https://www.bbc.co.uk/news/world-us-canada-40965029> [Accessed 19/4/22]
- BBC News (2017c) *Charlottesville: Virginia governor tells white supremacists: 'Go home'* 13 August. Available at: <https://www.bbc.co.uk/news/world-us-canada-40914643> [Accessed 19/4/22]
- BBC News (2018a) *John McCain: Five times he clashed with Trump* 26 August. Available at: <https://www.bbc.co.uk/news/world-us-canada-45313845> [Accessed 30/6/22]
- BBC News (2018b) *How to spot a bot* 2 June. Available at: <https://www.bbc.co.uk/news/av/stories-44215472> [Accessed 18/1/21]
- BBC News (2019) *Charlottesville attacker apologises as he is jailed for life* 28 June. Available at: <https://www.bbc.co.uk/news/world-us-canada-48806265> [Accessed 4/8/21]
- BBC News (2020a) *Coronavirus: Twitter bans incitement to attack 5G towers* 23 April. Available at: <https://www.bbc.co.uk/news/technology-52395158>
- BBC News (2020b) *Donald Trump orders creation of 'national heroes' garden* 5 July. Available at: https://www.bbc.co.uk/news/world-us-canada-53292585?at_medium=custom7&at_custom1=%5Bpost+type%5D&at_custom2=facebook_page&at_custom3=BBC+News&at_campaign=64&at_custom4=E6AA51E4-BE06-11EA-834E-0FBC923C408C [Accessed 19/4/22]
- Beil, K. (2015) The Myth of Black and White Modernism: Color Photographs and the Politics of Retrojective Looking *Visual Resources* 31(3-4) pp. 127-153

- Belfast Telegraph (2019) *Pro-life campaigners in online 'blackout'* 23 October. Available at: <https://www.belfasttelegraph.co.uk/news/northern-ireland/pro-life-campaigners-in-online-blackout-38620998.html> [Accessed 10/04/2020]
- Bennett, W. L. (1996) *News: The politics of illusion* New York: Longman
- Bennett, W. L. and Livingston, S. (2018) The disinformation order: Disruptive communication and the decline of democratic institutions *European Journal of Communication* 33(2) pp. 122-139
- Berenson, E. (2012) *The Statue of Liberty: A Transatlantic Story* United States: Yale University Press
- Berryman, R. and Kavka, M. (2018) Crying on YouTube: Vlogs, self-exposure and the productivity of negative affect *Convergence: The International Journal of Research into New Media Technologies* 4(1) pp. 85-98
- Bertrand, M. and Mullainathan, S. (2004) Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination *American Economic Review* 94(4) pp. 991-1013
- Beverland, M. B. and Farrelly, F. J. (2010) The Quest for Authenticity in Consumption: Consumers' Purposive Choices of Authentic Cues to Shape Experienced Outcomes *Journal of Consumer Research* 36(5) pp. 838-856
- Billson, A. (2020) Where have all the black cowboys gone? *The Guardian* 11 March. Available at: <https://www.theguardian.com/film/2020/mar/11/where-have-all-the-black-cowboys-gone> [Accessed 11/9/2020]
- Bishop, L. and Gray, D. (2018) Ethical Challenges of Publishing and Sharing Social Media Research Data in: Woodfield, K. (ed) *The Ethics of Online Research* Bingley: Emerald Publishing Limited
- Bogel-Burroughs, N. and Garcia, S. E. (2020) What Is Antifa, the Movement Trump Wants to Declare a Terror Group? *The New York Times* 28 September. Available at: <https://www.nytimes.com/article/what-Antifa-trump.html> [Accessed 03/02/2022]
- Bolinger, D. (1980) *Language- the loaded weapon: The use and abuse of language today* New York: Routledge
- Bonnette, L. M. (2015) *Pulse of the People: Political Rap Music and Black Politics* Pennsylvania: University of Pennsylvania Press
- Bortoluzzi, M. (2016) Voices of Self- and Other-Identification From a Pro-Innocent Community in Leppanen, S., Westinen, E., and Kytola, S. (eds) *Social Media Discourse, (Dis)Identifications and Diversities* New York: Routledge pp. 150-178
- Bovet, A. and Makse, H. A. (2019) Influence of fake news in Twitter during the 2016 US presidential election *nature communications* 10(7) pp. 1-14
- boyd, d. Golder, S. and Lotan, G. (2010) "Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter," *2010 43rd Hawaii International Conference on System Sciences*, 2010, pp. 1-10

- Boyd, R. L., Blackburn, K. G. and Pennebaker, J. W. (2020) The narrative arc: Revealing core narrative structures through text analysis *Science Advances* 6(32) pp. 1-9
- Braman, D. and Kahlan, D. M. (2006) Overcoming the Fear of Guns, the Fear of Gun Control, and the Fear of Cultural Politics: Constructing a better gun debate *Emory Law Journal* 55(4) pp. 569-608
- Brock, A. (2012) From the Blackhand Side: Twitter as a Cultural Conversation *Journal of Broadcasting & Electronic Media* 56(4) pp. 529-549
- Brunner, M. and Diemer, S. (2022) Cross-Platform Analysis. In: Vásquez, C. (Ed.) *Research Methods for Digital Discourse Analysis* London: Bloomsbury pp. 197-218
- Bucholtz, M. (1999) You da man: Narrating the racial other in the production of white masculinity. *Journal of Sociolinguistics* 3(4) pp. 443-460
- Bucholtz, M. and Hall, K. (2005) Identity and interaction: a sociocultural linguistic approach *Discourse Studies* 7(4-5) pp. 585-614
- Buttny, R. (1997) Reported Speech in Talking Race on Campus *Human Communication Research* 23(4) pp. 477-506
- Cai, C., Li, L. and Zengi, D., (2017) Behavior enhanced deep bot detection in social media 2017 *IEEE International Conference on Intelligence and Security Informatics* pp. 128-130
- Calabrese, E. (2020) 5 ways to spot disinformation on your social media feeds Available at: <https://abcnews.go.com/US/ways-spot-disinformation-social-media-feeds/story?id=67784438> [Accessed 19/1/21]
- Caldera, A. (2018) Woke Pedagogy: A Framework for Teaching and Learning *Diversity, Social Justice, and the Educational Leader* 2(3) pp. 1-11
- Gunslinger, n. (2022) *Cambridge Dictionary* Available at: <https://dictionary.cambridge.org/dictionary/english/gunslinger> [Accessed 20/01/2020]
- Catfishing, n. (2022) *Cambridge Dictionary* Available at: <https://dictionary.cambridge.org/dictionary/english/catfishing> [Accessed 21/9/22]
- Canan, N. and Warren, R. (2018) Cognitive Schemas and Disinformation Effects on Decision-making in lay Populations *ICCWS 2018 13th International Conference on Cyber Warfare and Security*
- Carlson, J. (2015) *Citizen-protectors: The everyday politics of guns in an age of decline*. New York: Oxford University Press.
- Cavazos-Rehg, P., Krauss, M., Gruzca, R. and Bierut, L. (2014) Characterizing the Followers and Tweets of a Marijuana-Focused Twitter Handle *Journal of Medical Internet Research* 16(6) pp. 1-12
- Chávez, K. and Cardellino, C. (2017) 5 Flawless Tips to Taking Your Best Selfie *Cosmopolitan* 29 June. Available at: <https://www.cosmopolitan.com/style-beauty/a12378/take-the-best-selfie/> [Accessed 21/10/22]

- Claassen, R. L. (2015) *Godless Democrats and Pious Republicans?: Party Activists, Party Capture, and the 'God Gap'* New York: Palgrave Macmillan
- Coleman, K. (2021) Introducing Birdwatch, a community-based approach to misinformation *Twitter Blog* [Online] 25 January 2021 Available at: https://blog.twitter.com/en_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation [Accessed 9/6/22]
- Conover, M. D., Ratkiewicz, J., Francisco, M., Gonçalves, B., Flammini, A. and Menczer, F. (2011) Political Polarization on Twitter *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* pp. 89-96
- Contreras, B. (2021) Hashtags felt dated and cringeworthy. So why are influencers still using them? *Los Angeles Times* 1 June. Available at: <https://www.latimes.com/business/technology/story/2021-06-01/hashtags-influencers-use-them> [Accessed 1/6/22]
- Coupland, N. (2001) Stylization, Authenticity and TV News Review *Discourse Studies* 3(4) pp. 413-42
- Coupland, N. (2003) Sociolinguistic authenticities *Journal of Sociolinguistics* 7(3) pp.417-431
- Coupland, N. (2007) *Style: Language variation and identity* Cambridge: Cambridge University Press
- Coupland, N. (2010) The authentic speaker and the speech community. In: Llamas, C. and Watts, D. (Eds.) *Language and identities* Edinburgh: Edinburgh University Press. pp. 99–112.
- Cowan, M. and Haydock, L. (2021) *Dating apps scam committed by criminal from inside prison* Available at: <https://www.bbc.co.uk/news/technology-56127488> [Accessed 20/8/22]
- Creedon, P. J. (1994) *Women, media, and sport: Challenging gender values* California: SAGE
- Cutler, C. (2003) “Keepin’ it Real”: White Hip-Hoppers’ Discourses of Language, Race and Authenticity *Journal of Linguistic Anthropology* 13(2) pp. 211-233
- Dawson, A. and Innes, M. (2019) How Russia’s Internet Research Agency Built its Disinformation Campaign *The Political Quarterly* 90 (2) pp. 245-256
- Dayter, D. (2015) Small stories and extended narratives on Twitter *Discourse, Context & Media* 10 pp. 19-26
- Dayter, D. (2016) *Discursive Self in Microblogging : Speech Acts, Stories and Self-praise* Amsterdam: John Benjamins Publishing Company
- De Choudhury, M. and De, S. (2014) Mental Health Discourse on reddit: Self-Disclosure, Social Support, and Anonymity *Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media*
- De Cock, B. and Pedraza, A. P. (2018) From expressing solidarity to mocking on Twitter: Pragmatic functions of hashtags starting with #jesuis across languages *Language in Society* 47(2) pp. 197-217

- De Fina, A. (2015) Narrative and Identities In: De Fina, A. and Georgakopoulou, A. (Eds.) *The Handbook of Narrative Analysis* Sussex: Wiley Blackwell. Pp. 351-368.
- De Zeeuw, D. and Tuters, M. (2020) Teh Internet is Serious Business: On the Deep Vernacular Web and Its Discontents *Cultural Politics* 16(2) pp. 214-232
- Defending Democracy Together (2022) *The Russia Tweets* Available at: <https://russiatweets.com/> [Accessed 20/1/2020]
- Deligiannis, N., Huu, T.D., Nguyen, D.M. and Luo, X. (2018) Deep learning for geolocating social media users and detecting fake news *NATO Workshop*
- Deseriis, M. (2013) Is Anonymous a New Form of Luddism?: A Comparative Analysis of Industrial Machine Breaking, Computer Hacking, and Related Rhetorical Strategies *Radical History Review* 2013(117) pp. 33-48.
- Devereux-Evans, O. and Morrison, C. (2022) 'Free speech absolutist' Elon Musk scraps Twitter's Covid misinformation policy that censored tweets questioning vaccines and lab leak origin theory *Daily Mail* 29 November 2022. Available at: <https://www.dailymail.co.uk/news/article-11481189/Twitter-drops-Covid-19-misleading-information-policy.html> [Accessed 6/12/22]
- Djafarova, E. and Rushworth, C. (2017) Exploring the credibility of online celebrities' Instagram profiles in influencing the purchase decisions of young female users *Computers in Human Behaviour* 68 pp. 1-7
- Dobson, A. S. (2015) *Postfeminist Digital Cultures: Femininity, Social Media and Self-Representation* New York: Palgrave Macmillan
- Döring, N., Reif, A. and Poeschl, S. (2016) How gender-stereotypical are selfies? A content analysis and comparison with magazine adverts *Computers in Human Behavior* 55(B) pp. 955-962
- Dotto, C. and Cubbon, S. (2019) *How to spot a bot (or not): The main indicators of automation, co-ordination and inauthentic activity* Available at: <https://firstdraftnews.org/latest/how-to-spot-a-bot-or-not-the-main-indicators-of-online-automation-co-ordination-and-inauthentic-activity/> [Accessed 19/1/21]
- Drenten, J., Gurrieri, L. and Tyler, M. (2019) Sexualised labor in digital culture: Instagram influencers, porn chic and the monetization of attention *Gender, work & organization* 27(1) pp. 41-66
- Dubey, Y., Wankhede, P., Borkar, A., Borkar, T. and Palsodkar, P. (2022) Framework for Fake News Classification Using Vectorization and Machine Learning. In: Lahby, M., Pathan, A. K., Maleh, Y. and Yafooz, W. M. S. (Eds.) *Combating Fake News with Computational Intelligence Techniques* Switzerland: Springer pp. 327-344
- Duell, M. (2017) 'Who is the real monster?' Internet turns on trolls who criticised 'indifferent' Muslim woman seen walking through terror attack *Daily Mail* 23 March. Available at: <https://www.dailymail.co.uk/news/article-4342438/Trolls-blast-Muslim-woman-seen-walking-attack.html> [Accessed 29/6/22]
- Duffy, B. (2017) *(Not) getting paid to do what you love: Gender, social media, and aspirational work*. New Haven: Yale University Press.

- Dumas, M. J. (2016) Against the Dark: Antiblackness in Education Policy and Discourse *Theory Into Practice* 55(1) pp. 11-19
- Duncombe, C. (2019) The Politics of Twitter: Emotions and the Power of Social Media *International Political Sociology* 13(4) pp. 409-429
- Dyer, R. (1991) A Star is Born and the construction of authenticity. In: Gledhill, C. (Ed.) *Stardom: Industry of Desire* London: Routledge pp. 132-140
- Edgar, A. N. (2016) Commenting Straight from the Underground: N.W.A., Police Brutality, and YouTube as a Space for Neoliberal Resistance *Southern Communication Journal* 81(4) pp. 223-236
- Eisenlauer, V. (2014) Facebook as a third author- (Semi-)automated participation framework in Social Network Sites *Journal of Pragmatics* 72 pp. 73-85
- Eklund, L., von Essen, E., Jonsson, F. and Johansson, M. (2018) To be or not to be on the internet: unpacking online anonymity *Stockholm Internet Research Group Working Paper Series SIRR2018_1*
- El Refaie, E. (2010) Visual modality versus authenticity: the example of autobiographical comics *Visual Studies* 25(2) pp. 162-174
- Eligon, J. (2020) A Debate Over Identity and Race Asks, Are African-Americans ‘Black’ or ‘black’? *The New York Times* 26 June. Available at: <https://www.nytimes.com/2020/06/26/us/black-african-american-style-debate.html> [Accessed 17/11/2020]
- Engs, R. F. and Miller, R. M. (2002) *The Birth of the Grand Old Party: The Republicans’ First Generation* Philadelphia: University of Pennsylvania Press
- Enli, G. (2017) Twitter as arena for the authentic outsider: exploring the social media campaigns of Trump and Clinton in the 2016 US Presidential Election *European Journal of Communication* 32(1) pp. 50-61
- Fallis, D. (2015) What is Disinformation? *Library Trends* 63(3) pp. 401-426
- Farkas, J. and Schou, J. (2018) Fake News as a Floating Signifier: Hegemony, Antagonism and the Politics of Falsehood *Journal of the European Institute of Communication and Culture* 25(3) pp. 298-314
- Farkas, J., Schou, J. and Neumayer, C. (2017) Cloaked Facebook Pages: Exploring fake Islamist propaganda in social media *New media and Society* 20(5) pp. 1850-1867
- Ferrara, E. (2017) Disinformation and Social Bot Operations in the Run Up to the 2017 French Presidential Election *First Monday* 22 (8)
- FiveThirtyEight (2018) 3 million Russian troll tweets Available at: [GitHub - fivethirtyeight/russian-troll-tweets](https://github.com/fivethirtyeight/russian-troll-tweets) [Accessed 20/1/2019]
- Fine, G.A. (2007), “Rumor, trust and civil society: collective memory and cultures of judgment” *Diogenes* 54(1) pp. 5-18.

- Fletcher, R., Cornia, A., Graves, L., & Nielsen, R. K. (2018). Measuring the reach of 'fake news' and online disinformation in Europe. *Australasian Policing* 10(2)
- Florini, S. (2014) Tweets, Tweeps and Signifyin': Communication and Cultural Performance on "Black Twitter" *Television and New Media* 15(3)
- Ford, E. (2018) Becoming the West: Cowboys as icons of masculine style for gay men *Critical Studies in Men's Fashion* 5(1-2) pp. 41-54
- Foth, M. and Hearn, G. (2007) Networked Individualism of Urban Residents: Discovering the communicative ecology in inner-city apartment buildings *Information, Communication & Society* 10(5) pp. 749-772
- Freelon, D. and Lokot, T. (2020) Russian Twitter disinformation campaigns reach across the political spectrum *The Harvard Kennedy School (HKS) Misinformation Review* 1(1) pp. 1-9
- Freelon, D. and Wells, C. (2020) Disinformation as Political Communication *Political Communication* 37(2) pp. 145-156
- Freelon, D., Bossetta, M., Wells, C., Lukito, J., Xia, Y. and Adams, K. (2022) Black Trolls Matter: Racial and Ideological Asymmetries in Social Media Disinformation *Social Science Computer Review* 40(3) pp. 560-578
- Froehlich, T. (2017). A not-so-brief account of current information ethics: the ethics of ignorance, missing information, misinformation, disinformation and other forms of deception or incompetence. *BiD: textos universitaris de biblioteconomia i documentació* 39
- Fryer, R. G. and Levitt, S. D. (2004) The Causes and Consequences of Distinctively Black Names *The Quarterly Journal of Economics* 119(3) pp. 767-805
- Gadde, V. and Beykpour, K. (2020) An update on our work around the 2020 US Elections *Twitter Blog* [Online] 12 November 2020. Available at: https://blog.twitter.com/en_us/topics/company/2020/2020-election-update [Accessed 1/6/22]
- Gadde, V. and Roth, Y. (2018) Enabling further research of information operations on Twitter *Twitter Blog* [Online] 17 October 2018. Available at: https://blog.twitter.com/en_us/topics/company/2018/enabling-further-research-of-information-operations-on-twitter
- Gaddis, S. M. (2017) How Black Are Lakisha and Jamal? Racial Perceptions from Names Used in Correspondence Audit Studies *Sociological Science* 4(19) pp. 469-489
- Gallaher, C. (2021) Mainstreaming white supremacy: a twitter analysis of the American 'Alt-Right' *Gender, Place & Culture* 28(2) pp. 224-252
- Gaskell, G and Bauer, M. W. (2000) Glossary. In: Gaskell, G and Bauer, M. W. *Qualitative Researching with Text, Image and Sound: A Practical Handbook for Social Research*. London: SAGE Publications. 351-368.
- Georgakopoulou, A. (2014) Small stories transposition and social media: A micro-perspective on the 'Greek crisis' *Discourse & Society* 25(4) pp. 519-539

- Georgakopoulou, A. (2015) Small Stories Research: Methods – Analysis – Outreach. In: De Fina, A. and Georgakopoulou, A. (Eds.) *The Handbook of Narrative Analysis* Sussex: Wiley Blackwell. Pp. 255-272
- Georgakopoulou, A. (2016) Small Stories Research: A Narrative Paradigm for the Analysis of Social Media. In: Sloan, L. and Quan-Haase, A. (Eds.) *The SAGE Handbook of Social Media Research Methods* London: SAGE Publications
- Georgakopoulou, A. (2017) Sharing the moment as small stories: The interplay between practices & affordances in the social media-curation of lives *Narrative Inquiry* 27(2) pp. 311-333
- Georgakopoulou, A. and Spilioti, T. (2016) Introduction. In: Georgakopoulou, A. and Spilioti, T. (Eds.) *The Routledge Handbook of Language and Digital Communication* Oxon: Routledge pp. 69-84
- Giachanou, A. and Rosso, P. (2020) The Battle against Online Harmful Information: The Cases of Fake News and Hate Speech *CIKM '20: Proceedings of the 29th ACM International Conference on Information & Knowledge Management* pp. 3503-3504
- Giaxoglou, K. (2021) *A Narrative Approach to Social Media Mourning* New York: Routledge
- Giaxoglou, K. and Spilioti, T. (2020) The shared story of #JeSuisAylan on Twitter Story participation and stancetaking in visual small stories *Pragmatics* 30(2) pp. 277-302
- Gibson, C. (2016) How clothing design and cultural industries refashioned frontier masculinities: a historical geography of Western wear *Gender, Place & Culture* 23(5) pp. 733-752
- Glover, S. (2020) One Thing Newsrooms Can Do: Capitalize “B” When Reporting About the Black Community *New York Amsterdam News* 11 June. Available at: <https://amsterdamnews.com/news/2020/06/11/one-thing-newsrooms-can-do-capitalize-b-when-repor/> [Accessed 26/10/2021]
- Goffman, E. (1959) *The presentation of self in everyday life*. New York: Doubleday.
- Goffman, E. (1981) *Forms of Talk* Pennsylvania: University of Pennsylvania Press
- Golovchenko, Y., Hartmann, M. & Adler-Nissen, R. (2018) State, media and civil society in the information warfare over Ukraine: citizen curators of digital disinformation *International Affairs*, vol. 95, no. 5, pp. 975-994.
- Graham, D. A. (2016) The Stubborn Persistence of Confederate Monuments *The Atlantic* 26 April. Available at: <https://www.theatlantic.com/politics/archive/2016/04/the-stubborn-persistence-of-confederate-monuments/479751/> [Accessed 19/4/22]
- Griswold, O. (2016) Center stage: direct and indirect reported speech in conversational storytelling *Issues in Applied Linguistics* 20(0) pp. 73-90
- Groshek, J., & Tandoc, E. (2016). The affordance effect: Gatekeeping and (non)reciprocal journalism on Twitter. *Computers in Human Behavior*, 66, pp. 201–210.

- Grossetti, Q., Mouza, C.D. and Travers, N., (2019) Community-based recommendations on Twitter: avoiding the filter bubble. In *International Conference on Web Information Systems Engineering* Cham: Springer International Publishing. pp. 212-227
- Gruber, H. (2017) Quoting and retweeting as communicative practices in computer mediated discourse *Discourse, Context and Media* 20 pp. 1-9
- Guess, A. M. (2015) Measure for Measure: An Experimental Test of Online Political Media Exposure *Political Analysis* 23(1) pp. 59-75
- Guess, A. M., Munger, K., Nagler, J. and Tucker, J. (2019a) How Accurate are Survey Responses on Social Media and Politics? *Political Communication* 36 pp. 241-258
- Guess, A. M., Nagler, J. and Tucker, J. (2019b) Less than you think: Prevalence and predictors of fake news dissemination on Facebook *Science Advances* 5(1) pp. 1-8
- Hameleers, M., Powell, T. E., Van Der Meer, T. G. L. A., and Bos, L. (2020) The Effects and Mechanisms of Multimodal Disinformation and Rebuttals Disseminated via Social Media *Political Communication* 37(2) pp. 281-301
- Harvey, O. (2017) HACKERS OF CYBERIA How Russia's web trolls tried to create race hatred with Sun pic during Westminster terror attack *The Sun* 14 November. Available at: <https://www.thesun.co.uk/news/4915727/how-russias-web-trolls-tried-to-create-race-hatred-with-sun-pic-during-westminster-terror-attack/> [Accessed 10/3/2019]
- Helmus, T. C., Bodine-Banon, E., Radin, A., Magnuson, M., Mendelsohn, J., Marcelline, B., Bega, A. and Winkelman, Z. (2018) *Russia Social Media Influence: Understanding Russian Propaganda in Eastern Europe* Santa Monica: RAND Corporation
- Hern, A. (2017) How a Russian 'troll soldier' stirred anger after the Westminster attack *The Guardian* 14 November. Available at: <https://www.theguardian.com/uk-news/2017/nov/14/how-a-russian-troll-soldier-stirred-anger-after-the-westminster-attack> [Accessed 14/7/22]
- Herring, S. (2004a) Assessing Gender Authenticity in Computer-Mediated Language Use *Journal of Language and Social Psychology* 23(4) pp. 424-446
- Herring, S. (2004b) Computer-Mediated Discourse Analysis: An Approach to Researching Online Behaviour. In: Barab, S. A., Kling, R. and Gray, J. H. *Designing for Virtual Communities in the Service of Learning*. Cambridge: Cambridge University Press. pp. 338-376.
- Herring, S. (2007) A faceted classification scheme for computer mediated discourse *Language@Internet* 4(1)
- Herring, S. (2011) Linguistic and Critical Analysis of Computer-Mediated Communication: Some Ethical and Scholarly Considerations *The Information Society* 12(2) pp. 153-168
- Herring, S. and Stoerger, S. (2014) Gender and (A)nonymity in Computer-Mediated Communication. In: Ehrlich, S., Meyerhoff, M. and Holmes, J. (Eds.) *The Handbook of Language, Gender, and Sexuality*, Second Edition, Sussex: John Wiley & Sons Ltd. pp. 567-586

- Hess, A. (2016) Earning the ‘Woke’ Badge *New York Times Magazine* 19 April. Available at: <https://www.nytimes.com/2016/04/24/magazine/earning-the-woke-badge.html> [Accessed 02/07/2020]
- High Level Expert Group on Fake News and Disinformation. (2018) *A multi-dimensional approach to disinformation. Report of the independent High level Group on fake news and online disinformation*. Luxembourg : Publications Office of the European Union
- Highfield, T. (2016) News via Voldemort: Parody accounts in topical discussions on Twitter *new media & society* 18(9) pp. 2028-2045
- Hill, A. (2002) Big Brother: the real audience *Television & New Media* 3(3) pp. 323-340
- Hill, M. L. (2018) “Thank You, Black Twitter” State Violence, Digital Counterpublics, and Pedagogies of Resistance *Urban Education* 53(2) pp. 286-302
- Hindman, M. and Barash, V. (2018) *Disinformation, “fake news” and influence campaigns on Twitter* Available at: <https://www.issuelab.org/resources/33163/33163.pdf> [Accessed 29/09/2022]
- Hinrichs, L. (2016) Approaches to language variation. In: Georgakopoulou, A. and Spilioti, T. (Eds.) *The Routledge Handbook of Language and Digital Communication* Oxon: Routledge pp. 19-35
- Horton, A. (2020) Agents of Chaos: A shocking look at what really happened in the 2016 election *The Guardian* 22 September. Available at: <https://www.theguardian.com/film/2020/sep/22/agents-of-chaos-alex-gibney-trump-russia-film> [Accessed 5/4/2021]
- Hu, M., Liu, S., Wei, F., Wu, Y., Stasko, J. and Ma, K. (2012) Breaking News on Twitter *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* pp. 2751-2754
- Huffman, H. S., Gibbs Knotts, H. and McKee, S. C. (2016) Down with the Southern Cross: Opinions on the Confederate Battle Flag in South Carolina *Political Science Quarterly* 132(4) pp. 719–41.
- Humprecht, E. (2019) Where ‘fake news’ flourishes: a comparison across four Western democracies *Information, Community & Society* 22(13) pp. 1973-1988
- Hunter, A. (2016) Monetizing the mommy: Mommy blogs and the audience commodity *Information, Communication & Society* 19(9) pp. 1306-1320
- Hussain, Z. and Griffiths, M.D. (2008) Gender swapping and socializing in cyberspace: an exploratory study. *Cyberpsychology & Behavior* 11(1) pp. 47-53.
- Hutchby, I. (2006) *Media Talk: Conversation Analysis and the Study of Broadcasting* Maidenhead: Open University Press
- Igwebuike, E. E. and Chimuanya, L. (2020) Legitimizing falsehood in social media: A discourse analysis of political fake news *Discourse & Communication* 15(1) pp. 42-58
- Ilbury, C. (2020) “Sassy Queens”: Stylistic orthographic variation in Twitter and the enregisterment of AAVE *Journal of Sociolinguistics* 24(2) pp. 245-264

- Im, J., Chandrasekharan, E., Sargent, J., Lighthammer, P., Denby, T., Bhargava, A., Hemphill, L., Jurgens, D. and Gilbert, E., (2020) Still out there: Modeling and identifying Russian troll accounts on twitter. *12th ACM Conference on Web Science* (pp. 1-10).
- Isaac, M. (2017) Russian Influence Reached 126 Million Through Facebook Alone *The New York Times* 30 October. Available at: <https://www.nytimes.com/2017/10/30/technology/facebook-google-russia.html> [Accessed 6/9/2021]
- Ivanova, I. (2022) *These formerly banned Twitter accounts have been reinstated since Elon Musk took over* Available at: <https://www.cbsnews.com/news/twitter-accounts-reinstated-elon-musk-donald-trump-kanye-ye-jordan-peterson-kathy-griffin-andrew-tate/> [Accessed 23/11/22]
- Jaffe, G. and Johnson, J. (2018) Republicans' anger at McCain speaks volumes about America's tribal politics *The Washington Post* 27 August. Available at: https://www.washingtonpost.com/politics/republicans-anger-at-mccain-speaks-volumes-about-americas-tribal-politics/2018/08/27/fb7fed5c-aa14-11e8-8a0c-70b618c98d3c_story.html [Accessed 29/8/22]
- Jäkälä, M. and Berki, E. (2013) Communities, Communication and Online Identities in: Warburton, S. and Hatzipanagos, S. (eds) *Digital Identity and Social Media* Pennsylvania: Information Science Reference pp. 1-13
- Jarry, J. (2021) *Tips for Better Thinking: Images Can Be Faked* Available at: <https://www.mcgill.ca/oss/article/critical-thinking/tips-better-thinking-images-can-be-faked> [Accessed 20/8/22]
- Jaworska, S. (2018) Doing Well by Talking Good: A Topic Modelling-Assisted Discourse Study of Corporate Social Responsibility *Applied Linguistics* 39(3) pp. 373-399
- Jemielniak, D. and Krempovych, Y. (2021) An analysis of AstraZeneca COVID-19 vaccine misinformation and fearmongering on Twitter *Public Health* 200 (2021) pp. 4-6
- Jenson, J., Taylor, N., de Castell, S. and Dilouya, B. (2015) Playing with Our Selves: Multiplicity and identity in online gaming *Feminist Media Studies* 15(5) pp. 860-879
- John, N.A. (2012) Sharing and Web 2.0: The emergence of a keyword *new media & society* 15(2) pp. 167-182
- Jones, B. (2015). The Confederate Flag is a Matter of Pride and Heritage, Not Hatred. *New York Times* 22 December. Available at: <https://www.nytimes.com/roomfordebate/2015/06/19/does-the-confederate-flag-breed-racism/the-confederate-flag-is-a-matter-of-pride-and-heritage-not-hatred> [Accessed 9/4/2020]
- Jones, R. (2009). Technology and sites of display. In: Jewitt, C. (ed) *The Routledge handbook of multimodal analysis*. London: Routledge. 114-126.
- Jones, R. H. (2013) Technology and sites of display. In: Jewitt, C. (Ed.) *The Routledge handbook of multimodal analysis* 2nd edition. Routledge, London, pp. 139-151.

- Jones, R. (2020) *Discourse analysis and digital surveillance* In De Fina, A. and Georgalopoulou, A. (eds) *The Cambridge Handbook of Discourse Studies*. Cambridge University Press, Cambridge pp. 708-731
- Jones, R. H., Chik, A. and Hafner, C. A. (2015). Introduction: Discourse analysis and digital practices. In: Jones, R. H., Chik, A. and Hafner, C. A. (Eds.) *Discourse and Digital Practices: Doing discourse analysis in the digital age* London: Routledge.
- Jurgens, D., Finethy, T., McCorriston, J., Xu, Y. T. and Ruths, D. (2015) Geolocation Prediction in Twitter Using Social Networks: A Critical Review Analysis and Review of Current Practice *Proceedings of the Ninth International AAAI Conference on Web and Social Media*
- Kalmar, I. (2020) Islamophobia and anti-antisemitism: the case of Hungary and the ‘Soros plot’ *Patterns of Prejudice* 54(1-2) pp. 182-198
- Kalmar, I., Stevens, C. and Worby, N. (2018) Twitter, Gab, and Racism: The Case of the Soros Myth *SM Society’18: Proceedings of the 9th International Conference on Social Media and Society* pp. 330-334
- Kelly, K., Beis, C. A., Zlotnik, A., and Schlangen, M. E. (2021) People-First Promotion: Rallying Library Workers during COVID-19 and Beyond *Marketing Libraries Journal* 5(1) pp. 4-17
- Knieper, T. and Müller, M. Eds. (2003). *Authentizität und Inszenierung von Bilderwelten*. Cologne: Herbert von Harlem.
- Knowles, T. (2022) Crackdown on social media’s anonymous trolls *The Times* 25 February. Available at: <https://www.thetimes.co.uk/article/crackdown-social-media-anonymous-trolls-online-harms-bill-c6r39fb8x> [Accessed 1/7/22]
- Kosoff, M. (2017) Even Trumpworld Got Tricked into Spreading Russian Propaganda *Vanity Fair* 19 October. Available at: <https://www.vanityfair.com/news/2017/10/even-trumpworld-got-tricked-into-spreading-russian-propaganda> [Accessed 29/6/22]
- Kozinski, A. (2015) The two faces of anonymity *Capital University Law Review* 43 pp. 1-17
- Kress, G. and Van Leeuwen, T. (2002) Colour as a semiotic mode: notes for a grammar of colour *Visual Communication* 1(3) pp. 343-368
- Kress, G. and Van Leeuwen, T. (2006) *Reading Images: The Grammar of Visual Design* 2nd Edition. Oxon: Routledge
- Kytölä, S. and Westinen, E. (2015) “I be da reel gangsta” – A Finnish footballer’s Twitter writing and metapragmatic evaluations of authenticity *Discourse, Context and Media* 8 pp. 6-19
- La Ganga, M. L. (2016) Clinton slams Trump's 'racist ideology' that ushers hate groups into mainstream *The Guardian* 25 August. Available at: <https://www.theguardian.com/us-news/2016/aug/25/hillary-clinton-alt-right-racism-speech-donald-trump-nevada> [Accessed 14/2/22]

- Labov, W. (1972) The transformation of experience in narrative syntax. In: Labov, W. (Ed.) *Language in the Inner City: Studies in the Black English Vernacular* Philadelphia, PA: University of Pennsylvania Press. Pp. 354-396
- Labov, W. (1986) The social stratification of (r) in New York City department stores *Dialect and Language Variation* pp. 304-329
- Labov, W. (2010) Narratives of personal experience. In: Hogan, P.C. (Ed.) *Cambridge Encyclopedia of the Language Sciences*. Cambridge: Cambridge University Press, 546–548.
- Langer, J. (1981) Television’s “personality system” *Media, Culture and Society* pp. 351-365
- Lapidot-Lefler, N. and Barak, A. (2012) Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition *Computers in Human Behaviour* 28(2) pp. 434-443
- Ledin, P. and Machin, D. (2018) *Doing Visual Analysis: From Theory to Practice* London: SAGE
- Lee, J. A. and Eastin, M. S. (2021) Perceived authenticity of social media influencers: scale development and validation *Journal of Research in Interactive Marketing* 15(4) pp. 822-841
- Lee, K., Eoff, B. and Caverlee, J. (2011) Seven Months With the Devils: A Long-Term Study of Content Polluters on Twitter. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media* 5(1) pp. 185-192
- Lehdonvirta, M., Nagashima, Y., Lehdonvirta, V. and Baba, A. (2012) The Stoic Male: How Avatar Gender Affects Help-Seeking Behavior in an Online Game *Games and Culture* 7(1) pp. 29-47
- Lepänen, S. and Piirainen-Marsh, A. (2009) Language policy in the making: an analysis of bilingual gaming activities *Lang Policy* 8(3) pp. 261-284
- Lester, N. (2021) Ban anonymous Twitter accounts to tackle cyber abuse, Government urged *Evening Standard* 23 March. Available at: <https://www.standard.co.uk/news/uk/twitter-government-ban-parliament-westminster-b925785.html> [Accessed 1/7/22]
- Levy, G. (2017) *Senators Push Trump to Condemn White Supremacists, KKK* Available at: <https://www.usnews.com/news/politics/articles/2017-09-06/senators-to-introduce-joint-resolution-against-white-supremacists-kkk> [Accessed 20/8/22]
- Lewandowsky, S., Ecker, U. K. H. and Cook, J. (2017) Beyond misinformation: Understanding and coping with the “Post-Truth” Era *Journal of Applied Research in Memory and Cognition* pp. 353-369
- Lieber, C. (2018) *How and why do influencers make so much money? The head of an influencer agency explains* Available at: <https://www.vox.com/the-goods/2018/11/28/18116875/influencer-marketing-social-media-engagement-instagram-youtube> [Accessed 20/04/20]
- Lindholm, C. (2008) *Culture and Authenticity* Malden: Blackwell

- Linville, D. L. and Warren, P. L. (2018) Troll Factories: The Internet Research Agency and State-Sponsored Agenda Building *Resource Centre on Media Freedom in Europe* Retrieved from: [http:// pwarren.people.clemson.edu/Linville_Warren_TrollFactory.pdf](http://pwarren.people.clemson.edu/Linville_Warren_TrollFactory.pdf).
- Linville, D. L. and Warren, P. L. (2020) Troll Factories: Manufacturing Specialized Disinformation on Twitter *Political Communication* 37(4) pp. 447-467
- Liu, W. and Ruths, D. (2013) What's in a name? using first names as features for gender inference on Twitter *Analyzing Microtext: Papers from the 2013 AAAI Spring Symposium*
- Lough, K., Molyneux, L. and Holton, A. E. (2018) A Clearer Picture: Journalistic identity practices in words and images on Twitter *Journalism Practice* 12(10) pp. 1277-1291
- Lovelock, M. (2016) Catching a Catfish: Constructing the "Good" Social Media User in Reality Television *Television & New Media* 18 (3) pp. 203-217
- Luebke S. M. (2021) Political authenticity: Conceptualization of a popular term. *The International Journal of Press/Politics*, 26(3) pp. 635–653.
- Luebke, S. M. and Engelmann, I. (2022) Do We Know Politicians' True Selves From the Media? Exploring the Relationship Between Political Media Exposure and Perceived Political Authenticity *Social Media + Society* 8(1) pp. 1-11
- Lyons, W., Scheb II, J. M. and Stair, B. (2001) *Government and Politics in Tennessee* Knoxville: University of Tennessee Press
- Machin, D. (2004) Building the World's Visual Language: The Increasing Global Importance of Image Banks in Corporate Media *Visual Communication* 3(3) pp. 316-336
- Magdy, W., Elkhatib, Y., Tyson, G., Joglekar, S. and Sastry, N. (2017) Fake it till you make it: Fishing for Catfishes *IEEE/ACM ASONAM 2017*
- Manzoor, S. (2013) *America's forgotten black cowboys* Available at: <https://www.bbc.co.uk/news/magazine-21768669> [Accessed 11/9/2020]
- Maragh, R. S. (2018) Authenticity on "Black Twitter": Reading Racial Performance and Social Networking *Television and New Media* 19(7) pp. 591-609
- Marwick, A. (2013) They're really profound women, they're entrepreneurs': Conceptions of authenticity in fashion blogging. *7th international AIII conference on weblogs and social media (ICWSM)* pp. 1-8
- Marwick, A. (2015) You May Know Me from YouTube: (Micro-)Celebrity in Social Media. In: Marshall, P. D. and Redmond, S. (Eds.) *A Companion to Celebrity* Hoboken, NJ, USA: John Wiley & Sons
- Marwick, A. (2018) Why do people share fake news? A sociotechnical model of media effects *Georgetown Law Technology Review* 2(474) pp. 474-512
- Maryns, K. and Blommaert, J. (2002). Pretextuality and Pretextual Gaps: On Re/defining Linguistic Inequality. *Journal of Pragmatics* 12(1) pp. 11–30

- Martin, B. A. S. and Gnoth, J. (2009) Is the Marlboro man the only alternative? The role of gender identity and self-construal salience in evaluations of male models *Marketing Letters* 20 pp. 353-367
- Mattu, P. M. and Lubbe, H. J. (2007) Investigating Language and Ideology: A presentation of the ideological square and transitivity in the editorials of three Kenyan newspapers *Journal of Language and Politics* 6(3) pp. 401-418
- Mcauliffe, T. (2017) *Trump Takes Heat for Blaming Charlottesville Violence on 'Many Sides'* Available at: <https://www.nbcnews.com/politics/white-house/trump-politicians-condemn-white-nationalist-rally-charlottesville-virginia-n792096> [Accessed 13/12/2021]
- McElhinny, B. (2003) Theorizing Gender in Sociolinguistics and Linguistic Anthropology in Holmes, J. and Meyerhoff, M. (Eds.) *The Handbook of Language and Gender* Malden: Blackwell Publishing Ltd.
- McLean, C. (2022) *Looking to get blue check mark on Twitter? Here's how to get verified* Available at: <https://eu.usatoday.com/tech/tips/> [Accessed 21/9/22]
- McManus, J. H. (1994) *Market-driven journalism: Let the citizen beware* Thousand Oaks, CA: Sage
- McRae, S. (2017) Get off my internets: How anti-fans deconstruct lifestyle bloggers' authenticity work *Personal Studies* 3(1) pp. 13-27
- Mejias, U. A. and Vokuer, N. E. (2017) Disinformation and the media: the case of Russia and Ukraine *Media, Culture and Society* 39(7) pp. 1027-1042
- Melzer, S. (2009) *Gun Crusaders: The NRA's Culture War* New York: New York University Press
- Mena, P., Barbe, D. and Chan-Olmstead, S. (2020) Misinformation on Instagram: The Impact of Trusted Endorsements on Message Credibility *Social Media + Society* 6(2) pp. 1-9
- Meyerhoff, M. (2016) Methods, innovations and extensions: Reflections on half a century of methodology in social dialectology *Journal of Sociolinguistics* 20(4) pp. 431-452
- Minamore, B. (2017) Woke Has Been Added To The Dictionary, But What Does It Really Mean? *Grazia* 14 June. Available at: <https://graziadaily.co.uk/life/real-life/woke-added-dictionary-really-mean/> [Accessed 17/01/2020]
- Mitchell, A., Shearer, E. and Stocking, G. (2021) *News on Twitter: Consumed by Most Users and Trusted by Many* Available at: <https://www.pewresearch.org/journalism/2021/11/15/news-on-twitter-consumed-by-most-users-and-trusted-by-many/> [Accessed 20/11/22]
- Mitchell, W. (2003) *Me ++ The Cyborg Self and the Networked City*. Massachusetts: MIT Press
- Mohamed, B., Haytam, H. and Abdelhadi, F. (2022) Applying Fuzzy Logic and Neural Network in Sentiment Analysis for Fake News Detection. In: Lahby, M., Pathan, A. K., Maleh, Y. and Yafooz, W. M. S. (Eds.) *Combating Fake News with Computational Intelligence Techniques* Switzerland: Springer pp. 387-400

- Monogarova, A., Shiryaeva, T., & Arupova, N. (2021). The Language of Russian Fake Stories: A Corpus-Based Study of the Topical Change in the Viral Disinformation *Journal of Language and Education* 7(4) pp. 83-106
- Morgan-Lopez, A. A., Kim, A. E., Chew, R. F. and Ruddle, P. (2017) Predicting age groups of Twitter users based on language and metadata *PLOS ONE* 12(8) pp. 1-12
- Morris, M. R., Counts, S., Roseway, A., Hoff, A., & Schwarz, J. (2012). Tweeting is believing?: Understanding microblog credibility perceptions *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work* pp. 441–450
- Morstatter, F., Wu, L., Nazer, T.H., Carley, K.M. and Liu, H. (2016) A new approach to bot detection: striking the balance between precision and recall *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)* pp. 533-540
- Mueller, R. S. (2019) *Report On The Investigation Into Russian Interference In The 2016 Presidential Election* Available at: https://upload.wikimedia.org/wikipedia/commons/3/3d/Mueller_Report_by_DPLA.pdf [Accessed 25/6/2021]
- Nakamura, L. (2008) *Digitizing Race. Visual Cultures of the Internet*. Minneapolis: University of Minnesota Press.
- National Park Service (2015) *Abolition* Available at: <https://www.nps.gov/stli/learn/historyculture/abolition.htm> [Accessed 17/7/2020]
- Nawaz, A. (2022) *How disinformation around Jan. 6 riot has downplayed violence, divided Americans* Available at: <https://www.pbs.org/newshour/show/how-disinformation-around-jan-6-riot-has-downplayed-violence-divided-americans> [Accessed 18/8/22]
- Nembhard, C. (2020) *What are Glo-ups & Why are They All Over Social Media?* Available at: <https://www.highsnobiety.com/p/glo-up-meaning-best-examples-2/#:~:text=Where%20did%20it%20come%20from,%20glo%20up%20one%20day.%E2%80%9D> [Accessed 22/1/21]
- Newman, G. E. and Smith, R. K. (2016) Kinds of Authenticity *Philosophy Compass* 11(10) pp. 609-618
- Newton, C. (2017) *Twitter's latest anti-abuse tools will keep eggs out of your mentions* Available at: <https://www.theverge.com/2017/3/1/14767074/twitter-anti-abuse-tools-filters-muting> [Accessed 3/10/22]
- Nguyễn, S., Kuo, R., Reddi, M. Li, L. and Moran, R. E. (2022) Studying mis- and disinformation in Asian diasporic communities: The need for critical transnational research beyond Anglocentrism *Harvard Kennedy School Misinformation Review* 3(2) pp. 1-12
- Nogara, G., Vishnuprasad, P. S., Cardoso, F., Ayoub, O., Giordano, S. and Luceri, L. (2022) The Disinformation Dozen: An Exploratory Analysis of Covid-19 Disinformation Proliferation on Twitter *14th ACM Web Science Conference 2022* pp. 348-358
- O'Brien, L. (2017) *Twitter Ignored This Russia-Controlled Account During The Election. Team Trump Did Not* Available at: <https://www.huffingtonpost.co.uk/entry/twitter-ignored->

[this-russia-controlled-account-during-the-election_n_59f9bdcbe4b046017fb010b0](#)

[Accessed 10/3/22]

- O'Reilly, T. and Milstein, S. (2012) *The Twitter Book*. 2nd edition. California: O'Reilly Media, Inc.
- O'Sullivan, D. (2018) *American media keeps falling for Russian trolls* Available at: <https://money.cnn.com/2018/06/21/technology/american-media-russian-trolls/index.html> [Accessed 29/6/22]
- Ochs, E. and Capps, L. (2001) *Living Narrative*. Cambridge, MA: Harvard University Press.
- Ochs, E. and Taylor, C. (1992) Family narrative as political activity *Discourse and Society* 3(3) pp. 301-340
- Olivier, J. (2014) Twitter usernames: exploring the nature of online South African nicknames *Nomina Africana* 28 pp. 51-74
- Ott, B. L. (2017) "The Age of Twitter: Donald J. Trump and the Politics of Debasement." *Critical Studies in Media Communication* 34 (1): 59–68.
- Page, R. (2012) *Stories and social media: identities and interaction* New York: Routledge
- Page, R. (2018) *Narratives Online: Shared Stories in Social Media* Cambridge: Cambridge University Press
- Page, R., Barton, D., Unger, J. W., and Zappavigna, M. (2014) *Researching Language and Social Media: A Student Guide* New York: Routledge
- Parada, S. A., DeVine, J. G. and Arrington, E. D. (2009) Celebratory gunfire injury to a United States soldier sustained during Operation Iraqi Freedom (OIF) *Injury Extra* 40 pp. 149-151
- Pariser, E. (2011) *The Filter Bubble: What The Internet Is Hiding From You* New York: Penguin
- Patel, M., Padiya, J. and Singh, M. (2022) Fake News Detection Using Machine Learning and Natural Language Processing. In: Lahby, M., Pathan, A. K., Maleh, Y. and Yafooz, W. M. S. (Eds.) *Combating Fake News with Computational Intelligence Techniques* Switzerland: Springer pp. 127-148
- Patterson, A. N. (2018) "YouTube Generated Video Clips as Qualitative Research Data: One Researcher's Reflections on the Process" *Qualitative Inquiry* 24(10) pp. 759-767
- Pavalanathan, U. and Eisenstein, J. (2015) Confounds and consequences in geotagged Twitter data *arXiv preprint arXiv:1506.02275*
- Peddinti, S. T., Ross, K. W. and Capps, J. (2014) "On the internet, nobody knows you're a dog": a Twitter case study of anonymity in social networks *COSN '14: Proceedings of the second ACM conference on Online social networks* pp. 83-94
- Peddinti, S. T., Ross, K. W. and Capps, J. (2017) User Anonymity on Twitter *IEEE Security & Privacy* 15(3) pp. 84-87

- Pengelly, M. (2020) Trump spreads false claim Democrats dropped God from Pledge of Allegiance *The Guardian* 22 August. Available at: <https://www.theguardian.com/us-news/2020/aug/22/donald-trump-democrats-god-pledge> [Accessed 29/5/22]
- Penn, G. (2000) Semiotic Analysis of Still Images. In: Gaskell, G. and Bauer, M. W. *Qualitative Researching with Text, Image and Sound: A Practical Handbook for Social Research*. London: SAGE Publications. 227-245.
- Pennock-Speck, B. (2016) Violent women in Spanish TV ads: Stereotype reversal or the same old same old? *Discourse & Communication* 10(4) pp. 363-377
- Pennycook, A. (2018). Posthumanist Applied Linguistics *Applied Linguistics* 39(4) pp. 445-461
- Philips, K. (2017) Timeline: How growing anger finally pushed Trump to denounce white supremacists *The Washington Post* 15 August. Available at: <https://www.washingtonpost.com/news/the-fix/wp/2017/08/15/timeline-how-the-growing-anger-finally-pushed-trump-to-denounce-white-supremacists/> [Accessed 6/5/22]
- Piela, A. (2013) I am Just Doing my Bit to Promote Modesty: Niqabis' self-portraits on photo-sharing websites *Feminist Media Studies* 13(5) pp. 781-790
- Pillow D. R., Crabtree M. A., Galvan M. J., Hale W. J. (2018). Not simply in the eye of the beholder: Authenticity as a product of candidate preference and unfettered speech. *Political Psychology*, 39(4) pp. 849–868
- Polyakova, A. and Boyer, S.P. (2018). 'The Future of Political Warfare: Russia, the West and the Coming Age of Global Digital Competition'. *Brookings Institution*. Available at: <https://www.brookings.edu/wp-content/uploads/2018/03/the-future-of-political-warfare.pdf>
- Posetti, J. and Matthews, A. (2018) *A short guide to the history of 'fake news' and disinformation* Available at: https://www.icfj.org/sites/default/files/2018-07/A%20Short%20Guide%20to%20History%20of%20Fake%20News%20and%20Disinformation_ICFJ%20Final.pdf [Accessed 15/7/22]
- Powell, T. E., Boomgaarden, H. G., de Swert, K., & de Vreese, C. H. (2015). A clearer picture: The contribution of visuals and text to framing effects. *Journal of Communication* 65(6) pp. 997–1017.
- Price, V., Tewksbury, D., & Powers, E. (1997). Switching trains of thought. The impact of news frames on readers' cognitive responses. *Communication Research* 24, pp. 481–506.
- Ramaswamy, C. (2019) The more the word 'woke' is used as a slur and a joke, the more we need it *The Guardian* 4 March. Available at: <https://www.theguardian.com/lifeandstyle/2019/mar/04/chitra-ramaswamy-the-more-woke-is-used-as-a-joke-and-slur-the-more-we-need-it> [Accessed 16/01/2020]
- Ramos, J. A. N. (2020) *AAVE and its Presence In and Influence On Rap Music* Available at: https://rodin.uca.es/xmlui/bitstream/handle/10498/23290/Neira%20Ramos_TFG.pdf?sequence=1&isAllowed=y [Accessed 26/1/2021]

- Ravelli, L. J. and Van Leeuwen, T. (2018) Modality in the digital age *visual communication* 17(3) pp. 277-297
- Recuero, R., Soares, F. B., and Gruzd, A. (2020) Hyperpartisanship, Disinformation and Political Conversation on Twitter: The Brazilian Presidential Election of 2018 *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media* 14 pp. 569-578
- Robertson, A. (2017) Revealed: 'Fake news' Twitter account that posted photo of 'Muslim woman ignoring the Westminster terror attack' was run from RUSSIA *Daily Mail* 13 November. Available at: <https://www.dailymail.co.uk/news/article-5077771/Photo-Muslim-ignoring-Westminster-came-Russia.html> [Accessed 10/3/2019]
- Rochlin, N. (2017) Fake news: belief in post-truth *Library Hi Tech* 35(3) pp. 386-392
- Rodham Clinton, H. (2017) *What Happened* New York: Simon and Schuster Inc.
- Rosewarne, L. (2007) *Sex in Public: Women, Outdoor Advertising and Public Policy* Newcastle: Cambridge Scholars Publishing
- Rosenstiel, T., Sonderman, J., Loker, K., Ivancin, M. and Kjarval, N. (2015) *How people use Twitter in general* <https://www.americanpressinstitute.org/publications/reports/survey-research/how-people-use-twitter-in-general/> [Accessed 5/12/22]
- Roth, Y. and Pickles, N. (2020) Updating our approach to misleading information *Twitter Blog* [Online] 11 May 2020. Available at: https://blog.twitter.com/en_us/topics/product/2020/updating-our-approach-to-misleading-information [Accessed 31/5/22]
- Sadler, N. (2017) Narrative and interpretation on Twitter: Reading tweets by telling stories *New Media & Society* 20(9) pp. 3266-3282
- Sanchez, R. (2017) *Thousands march in Boston in protest of controversial rally* Available at: <https://edition.cnn.com/2017/08/19/us/boston-free-speech-rally/index.html> [Accessed 20/4/22]
- Sasahara, K., Chen, W., Peng, H., Ciampaglia, G. L., Flammini, A., and Menczer, F. (2021) Social influence and unfollowing accelerate the emergence of echo chambers *Journal of Computational Social Science* 4 pp. 381-402
- Scannell, D., Desens, L., Guadagno, M., Tra, Y., Acker, E., Sheridan, K., Rosner, M., Mathieu, J. and Fulk, M. (2021) COVID-19 Vaccine Discourse on Twitter: A Content Analysis of Persuasion Techniques, Sentiment, and Mis/Disinformation *Journal of Health Communication* 26 (7) pp. 443-459
- Schau, H. J. and Gilly, M. C. (2003) We Are What We Post? Self-Presentation in Personal Web Space *Journal of Consumer Research* 30(3) pp. 385-404
- Sergeant, P. and Tagg, C. (2016) *The filter bubble isn't just Facebook's fault – it's yours* Available at: <https://theconversation.com/the-filter-bubble-isnt-just-facebooks-fault-its-yours-69664> [Accessed 21/09/22]
- Seidel, J. (2017) *US Congress told how Russia 'weaponised' this photo of a 'Muslim woman'* Available at: <https://www.news.com.au/technology/online/social/us-congress-told-how->

[russia-weaponised-this-photo/news-story/8135ca050976761ab476c04c32d44fd2](https://www.russia-weaponised-this-photo/news-story/8135ca050976761ab476c04c32d44fd2)

[Accessed 17/04/19]

- Seo, Y., Kim, J., Choi, Y. K. and Li, X. (2019) In “likes” we trust: likes, disclosures and firm-serving motives on social media *European Journal of Marketing* 53(10) pp. 2173-2192
- Shah, R. and Tewari, R. (2016) Demystifying ‘selfie’: a rampant social media activity *Behaviour and Information Technology* 35 (10) pp. 864-871
- Shane, S. (2017) The Fake Americans Russia Created to Influence the Election *The New York Times* 7 September Available at: <https://www.nytimes.com/2017/09/07/us/politics/russia-facebook-twitter-election.html> [Accessed 10/3/2019]
- Sharma, S. (2013) Black Twitter? Racial Hashtags, Network and Contagion *New Formations* 78 pp. 46-64
- Shrikant, N. and Sierra, S. (2021) “Nasty Question” and “Fake News”: Metadiscourse as a Resource for Denying Accusations of Racism in Donald Trump’s Presidential Press Events *Howard Journal of Communications* pp. 1-21
- Smith, D. (2017) Trump paints himself as the real victim of Charlottesville in angry speech *The Guardian* 23 August. Available at: <https://www.theguardian.com/us-news/2017/aug/23/donald-trump-arizona-rally-phoenix> [Accessed 31/8/2021]
- Smoleňová, I. (2015) *The Pro-Russian Disinformation Campaign in the Czech Republic and Slovakia: Types of Media Spreading, Pro-Russian Propaganda, Their Characteristics and Frequently Used Narratives*. Prague: Prague Security Studies Institute
- Soares, F. B. and Recuero, R. (2021) Hashtag Wars: Political Disinformation and Discursive Struggles on Twitter Conversations During the 2018 Brazilian Presidential Campaign *Social Media + Society* 7(2) pp. 1-13
- Social Security Administration (2020) *Top Names Over the Last 100 Years* Available at: <https://www.ssa.gov/oact/babynames/decades/century.html> [Accessed 18/2/2020]
- Solberg, L. (2010) Data Mining on Facebook: A Free Space for Researchers or an IRB Nightmare? *University of Illinois Journal of Law, Technology and Policy* 2 pp. 311-343
- Spilioti, T. and Tagg, C. (2016) The Ethics of Online Research Methods in Applied Linguistics: Challenges, opportunities, and directions in ethical decision-making *Applied Linguistics Review* 8(2-3) pp. 163-167
- Spilioti, T. and Tagg, C. (2022) Research Ethics. In: Vásquez, C. (Ed.) *Research Methods for Digital Discourse Analysis* London: Bloomsbury Academic
- Spruce, L. and Leaf, K. (2017) Social Media for Social Justice *Journal of Museum Education* 42(1) pp. 41-53
- Squires, L. (2010) Enregistering internet language *Language in society* 39(4) pp. 457-492
- Sririam, S. K. (2016) A Foucauldian Theory of American Islamophobia *International Journal of Islamic Thought* 10 pp. 47-54

- Starbird, K. (2017) *Information Wars: A Window into the Alternative Media Ecosystem* Available at: <https://medium.com/hci-design-at-uw/information-wars-a-window-into-the-alternative-media-ecosystem-a1347f32fd8f> [Accessed 13/7/22]
- Starbird, K., Maddock, J., Orand, M., Achterman, P., and Mason R. M. (2014) Rumors, False Flags, and Digital Vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombing. In: Kindling, M. & Greifeneder, E. (Eds.), *iConference 2014 Proceedings* (pp. 654–662).
- Stewart, B. (2020) The Rise of Far-Right Civilizationism *Critical Sociology* 46(7-8) pp. 1207-1220
- Stewart, L. G., Arif, A., Starbird, K. (2018) Examining Trolls and Polarization with a Retweet Network *MIS2: Misinformation and Misbehavior Mining on the Web Workshop held in conjunction with WSDM 2018* Vol 70
- Stoddard, J., Tunstall, J., Walker, L., Wright, E. (2021) Teaching beyond verifying sources and “fake news”: Critical media education to challenge media injustices *Journal of Media Literacy Education* 13(2) pp. 55-70
- Stokoe, E. (1998) Talking about gender: the conversational construction of gender categories in academic discourse *Discourse & Society* 9(2) pp. 217-240
- Stroud, A. (2016) *Good guys with guns: The appeal and consequences of concealed carry* Chapel Hill, NC: University of North Carolina Press
- Sunstein, C. (2001) *Echo Chambers* Princeton: Princeton University Press
- Sunstein, C. (2009) *Republic.com 2.0* Princeton: Princeton University Press
- Sweeney, L. (2013) Discrimination in Online Ad Delivery: Google ads, black names and white names, racial discrimination and click advertising *ACM QUEUE* 11(3) pp. 1-19
- Sztompka, P. (1999) *Trust: A Sociological Theory*. Cambridge: Cambridge University Press.
- Tagg, C., Seargeant, P. and Brown, A. A. (2017) Constructing Conviviality in Online Interaction. In: Tagg, C. (Ed.) *Taking Offence on Social Media: Conviviality and Communication on Facebook* Cham: Palgrave Macmillan pp. 46-56
- Tajfel, H. (1982) Social Psychology of Intergroup Relations *Annual Review of Psychology* 33 pp. 1-39
- Takhteyev, Y., Gruzd, A. and Wellman, B. (2012) Geography of Twitter networks *Social networks* 34(1), pp.73-81
- Tandoc, E. C., Jenkins, J., and Craft, S. (2018) Fake News as a Critical Incident in Journalism *Journalism Practice* 13(6) pp. 673-689
- Tannen, D. (2007) *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse* 2nd edition. Cambridge: Cambridge University Press
- Tannen, D. and Trester, A.M. (2013) *Discourse 2.0: Language and New Media* Washington, DC: Georgetown University Press

- Taylor, D. (1994) Inauthentic Authenticity or Authentic Inauthenticity? *The Electric Journal for English as a Second Language* 1(2)
- Texas State Historical Association, Handbook of Texas Online (2019) *Stetson Hats* Available at <http://www.tshaonline.org/handbook/online/articles/lfs01> [Accessed 17/01/2020]
- Thomson, R. and Murachver, T. (2001) Predicting gender from electronic discourse *British Journal of Social Psychology* 40 pp. 193-208
- Thornborrow, J. (2001) Authenticating Talk: Building Public Identities in Audience Participation Broadcasting *Discourse Studies* 3(4) pp. 459-479
- Thurlow, C. (2017) Digital discourse: Locating language in new/social media in Burgess, J., Poell, T. and Marwick, A. (eds) *The SAGE Handbook of Social Media* New York: Sage pp. 135-145
- Thurlow, C. and Mroczek, K. (2011). Introduction: Fresh Perspectives on New Media Sociolinguistics. In: Thurlow, C. and Mroczek, K. *Digital Discourse: Language in the New Media* New York: Oxford University Press. pp. xix - 1.
- Tien, J. H., Eisenberg, M. C., Cherng, S. T. and Porter, M. A. (2020) Online reactions to the 2017 'Unite the right' rally in Charlottesville: measuring polarization in Twitter networks using media followership *Applied Network Science* 5(10) pp. 1-27
- Tiggemann, M., Anderberg, I. and Brown, Z. (2020) Uploading your best self: Selfie editing and body dissatisfaction *Body Image* 33 pp. 175-182
- TinEye (2022) *Industry-leading search by image and computer vision solutions* Available at: <https://tineye.com/technology>
- Tolson, A. (2001) 'Being Yourself': The pursuit of authentic celebrity *Discourse Studies* 3(4) pp. 443-457
- Tolson, A. (2006) *Media talk: spoken discourse on TV and radio* Edinburgh: Edinburgh University Press
- Tornoe, R. (2017) Two popular conservative Twitter personalities were just outed as Russian trolls *The Philadelphia Inquirer* 3 November. Available at: <https://www.philly.com/philly/news/politics/presidential/russia-fake-twitter-facebook-posts-accounts-trump-election-jenna-abrams-20171103.html> [Accessed 17/04/19]
- Tovares, A. (2022) The Role of Theory in Digital Discourse Analysis. In: Vásquez, C. (Ed.) *Research Methods for Digital Discourse Analysis* London: Bloomsbury pp. 19-36
- Trilling, L. (1972) *Sincerity and Authenticity*. London: Oxford University Press.
- Tromble, R. (2017) Lost Umbrellas: Bias and the Right to be Forgotten in Social Media Research. In: Zimmer, M. and Kinder-Kurlander, K. (Eds.) *Internet Research Ethics for the Social Media Age: New Cases and Challenges* New York: Peter Lang Publishers pp. 75-91
- Twitter (2016) *Welcome to Twitter* <http://twitter.com> on Internet Archive <https://web.archive.org/web/20161225015357/https://twitter.com/> [Screen grab accessed 6/12/22]

- Twitter (2019) *About Verified Accounts* Available at: <https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts> [Accessed 04/03/2019]
- Twitter (2020) *Glossary* Available at: <https://help.Twitter.com/en/glossary> [Accessed 16/01/2020]
- Twitter (2021a) *About Verified Accounts* Available at: <https://help.twitter.com/en/managing-your-account/about-twitter-verified-accounts#:~:text=The%20blue%20Verified%20badge%20on,authentic%2C%20notable%2C%20and%20active> [Accessed 29/6/22]
- Twitter (2021b) What's in a name? The case for inclusivity through anonymity *Twitter Blog* [Online] 21 September 2021 Available at: <https://blog.twitter.com/common-thread/en/topics/stories/2021/whats-in-a-name-the-case-for-inclusivity-through-anonymity> [Accessed 7/12/22]
- Twitter (2022a) *Twitter Moderation Research Consortium* Available at: <https://transparency.twitter.com/en/reports/moderation-research.html> [Accessed 19/8/22]
- Twitter (2022b) *Twitter Privacy Policy* Available at: <https://Twitter.com/en/privacy> [Accessed 19/8/22]
- Twitter (2022c) *How to customize your profile* Available at: <https://help.twitter.com/en/managing-your-account/how-to-customize-your-profile> [Accessed 7/12/22]
- Twitter (2022d) How many people come to Twitter for news? As it turns out, A LOT *Twitter Blog* [Online] 12 September 2022 Available at: https://blog.twitter.com/en_us/topics/insights/2022/how-many-people-come-twitter-for-news#:~:text=Twitter%20connects%20people%20with%20news,on%20Twitter%20for%20their%20jobs [Accessed 5/12/22]
- United Kingdom Government Department for Digital, Culture, Media and Sport (2022) *Press Release: New plans to protect people from anonymous trolls online* Available at: <https://www.gov.uk/government/news/new-plans-to-protect-people-from-anonymous-trolls-online> [Accessed 1/7/22]
- United States Census Bureau (2010) *Frequently Occurring Surnames in the 2010 Census: Top 1,000 Surnames* Available at: https://www.census.gov/topics/population/genealogy/data/2010_surnames.html [Accessed 18/2/2020]
- United States Department of Justice (2018) *Grand Jury for the District of Columbia Indictment of 13 Russian Nationals* Available at: <https://www.documentcloud.org/documents/4380529-Internet-Research-Agency-Indictment.html> [Accessed 2/3/2019]
- United States Senate Committee, Subcommittee on Crime and Terrorism (2017) *Extremist Content and Russian Disinformation Online: Working with Tech to Find Solutions*. Available at: <https://www.judiciary.senate.gov/meetings/extremist-content-and-russian-disinformation-online-working-with-tech-to-find-solutions> [Accessed 16/6/22]

- van der Land, S. F., Willemsen, L. M. and Wilton, B. G. E. (2016) Professional Personal Branding: Using a Think-Aloud Protocol to Investigate how Recruiters Judge LinkedIn Profile Pictures *HCI in Business, Government, and Organizations: eCommerce and Innovation, July 17-2, Proceedings Part 1*
- van der Nagel, E. and Frith, J. (2015) Anonymity, pseudonymity and the agency of online identity: Examining the social practices of r/gonewild *First Monday* 20(3) pp. 1-13
- van Leeuwen, T. (2001) What is authenticity? *Discourse Studies* 3(4) pp. 392-397
- Varis, P. and Wang, X. (2015). Superdiversity on the Internet: A Case from China. In: Arnaukt, K., Blommaert, J., Rampton, B. and Spotti, M. *Language and Superdiversity*. New York: Routledge. pp. 218-237.
- Vásquez, C. (2022) Introduction. In: Vásquez, C. (Ed.) *Research Methods for Digital Discourse Analysis* London: Bloomsbury pp. 1-18
- Veum, A. and Undrum, L. V. M. (2017) The selfie as a global discourse *Discourse and Society* 29(1) pp. 86-103
- Vieten, U. M. (2016) Far Right Populism and Women: The Normalisation of Gendered Anti-Muslim Racism and Gendered Culturalism in the Netherlands *Journal of Intercultural Studies* 37(6) pp. 621-636
- Vis, F. (2013) Twitter as a Reporting Tool for Breaking News *Digital Journalism* 1(1) pp. 27-47
- Vrikki, P. (2020) #PeoplesVoteMarch or #LoserseVoteMarch? Tracing the Collective Identity of a Post-Brexit Referendum Movement on Twitter in Bouvier, G. and Rosenbaum, J. E. (eds.) *Twitter, the Public Sphere, and the Chaos of Online Deliberation* Cham: Palgrave Macmillan
- Wang, X. and Song, Y. (2020) Viral misinformation and echo chambers: the diffusion of rumors about genetically modified organisms on social media *Internet Research* 30(5) pp. 1547-1564
- White, R. L. C. and Sevier, J. (1901) The Great Seal of the State of Tennessee *The American Historical Magazine* 6(3) pp. 195-212
- Whitehead, D. (2015) The Evidence of Things Unseen: Authenticity and Fraud in the Christian Mommy Blogosphere *Journal of the American Academy of Religion* 83(1) pp. 120-150
- Williams, J. P. (2006) Authentic Identities: Straightedge Subculture, Music, and the Internet *Journal of Contemporary Ethnography* 35(2) pp. 173-200
- Williams, M. L., Burnap, P., Sloan, L., Jessop, C. and Lepps, H. (2017) Users' Views of Ethics in Social Media Research: Informed Consent, Anonymity, and Harm. In: Woodfield, K. (Ed.) *The Ethics of Online Research* Bingley: Emerald Publishing Limited pp. 27-52
- Williams, S. (2015) Digital Defense: Black Feminists Resist Violence With Hashtag Activism *Feminist Media Studies* 15(2) pp. 341-344

- Woke, adj. (2022) *Oxford English Dictionary* Available at: <https://www.oed.com/view/Entry/58068747?rskey=JCSq7y&result=3#eid> [Accessed 1/4/21]
- Wood, K. (2019) *Best States for Gun Owners* Available at: <https://www.gunsandammo.com/editorial/best-states-for-gun-owners/369075> [Accessed 11/02/2020]
- World Economic Forum (2022) *Global Risk Report* Available at: https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2022.pdf [Accessed 16/5/22]
- World Health Organization. (2020) Speech by the director-general at the Munich security conference. [Online] Available at: <https://www.who.int/dg/speeches/detail/munich-security-conference> [Accessed 1/6/22]
- Xia, Y., Lukito, J., Zhang, Y., Wells, C., Kim, S. J., and Tong, C. (2019) Disinformation, performed: Self-presentation of a Russian IRA account on Twitter *Information, Communication & Society* pp. 1646-1664
- Yafooz, W. M. S., Emara, A. M. and Lahby, M. (2022) Detecting Fake News on COVID-19 Vaccine from YouTube Videos Using Advanced Machine Learning Approaches. In: Lahby, M., Pathan, A. K., Maleh, Y. and Yafooz, W. M. S. (Eds.) *Combating Fake News with Computational Intelligence Techniques* Switzerland: Springer pp. 421-435
- Yamane, D. (2017) The Sociology of U.S. Gun Culture *Sociology Compass* 11(7) pp. 1-10
- Yang, X., Kim, S. and Sun, Y. (2019) How Do Influencers Mention Brands in Social Media? Sponsorship Prediction of Instagram Posts *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining* August 2019 pp. 101-104
- Yao, C. X. (2018) #staywoke: Digital Engagement and Literacies in Antiracist Pedagogy *American Quarterly; College Park* 70(3) pp. 439-454
- Yee, N., Ducheneaut, N., Yao, M and Nelson, L. (2011) “Do Men Heal More When in Drag? Conflicting Identity Cues between User and Avatar.” *Proceedings of ACM CHI '11 Conference on Human Factors in Computing Systems* pp. 773 –776.
- Zannettou, S., Caulfield, T., De Cristofaro, E., Sirivianos, M., Stringhini, G., Blackburn, J. (2019) Disinformation Warfare: Understanding State-Sponsored Trolls on Twitter and Their Influence on the Web *WWW '19 Companion Proceedings of The 2019 World Wide Web Conference* pp. 218-226
- Zappavigna, M. (2011) Ambient affiliation: a linguistic perspective on Twitter *new media & society* pp. 788-806
- Zappavigna, M. (2015) Searchable talk: the linguistic functions of hashtags *Social Semiotics* 25(3) pp. 274-291
- Zappavigna, M. (2016) Social media photography: construing subjectivity in Instagram images *Visual Communication Journal* 15(3) pp. 271-292

- Zappavigna, M. (2019) Ambient affiliation and #Brexit: Negotiating values about experts through censure and ridicule. In: Koller, V., Kopf, S. and Milgbauer, M. (Eds.) *Discourses of Brexit* pp. 48-68 London: Routledge
- Zelizer, B. (2017) *What Journalism Could Be Like* Cambridge: Polity Press
- Zempi, I. (2016) 'It's a part of me, I feel naked without it': choice, agency and identity for Muslim women who wear the niqab *Ethnic and Racial Studies* 39(10) pp. 1738-1754
- Zulli, D. (2018) Capitalizing on the look: insights into the glance, attention economy, and Instagram *Critical Studies in Media Communication* 35(2) pp. 137-150

Appendix

user_id	user_display_name	user_screen_name	user_reported_location	user_profile_description	user_profile_picture
4224729994	Tennessee	TEN_GOP		Unofficial Twitter of Tennessee Republicans. Covering breaking news, national politics, foreign policy and more. #MAGA #2A	
842368501528559616	Luisa Haynes	wokeluisa	New York, USA	APSA. #Blackexcellence. Political science major	
4355797094	Texas Lone Star	SouthLoneStar	Texas, USA	Proud TEXAN and AMERICAN patriot #2a #prolife #Trump2016 #TrumpPence16 Fuck Islam and PC. Don't mess with Texas!	
4332740714	Gunslinger Girl	USA_Gunslinger	Wisconsin, USA	Truth is strong, and sometime or other will prevail!	
4272870988	Pamela Moore	Pamela_Moore13	Texas, USA	Southern. Conservative. Pro God. Anti Racism	
905141386781085696	Kanisha J	KanIJackson	Atlanta, GA	Follow the example set by Mrs Obama; peace, love, acceptance & vigilance #Impeach45 #Resist #GunReformNow	
4437233895	Crystal Johnson	Crystal1Johnson	Richmond, VA	It is our responsibility to promote the positive things that happen in our communities.	
4218156466	ELEVEN GOP	ELEVEN_GOP	Tennessee, USA	This is our back-up account in case anything happens to @TEN_GOP	
4301962823	Black News	BlackNewsOutlet	Black America	Freedom is never given; it is won. #BlackLivesMatter	
4508630900	John Davis	TheFoundingSon	Texas, USA	Business Owner, Proud Father, Conservative, Christian, Patriot, Gun rights, Politically Incorrect. Love my country and my family #2A #GOP #cot #WakeUpAmerica	
2882331822	Jenna Abrams	Jenn_Abrams	USA	Calm down, I'm not pro-Trump. I am pro-common sense. Any offers/ideas/questions? DM or email me jennabrams@gmail.com (Yes, there are 3 Ns)	https://t...
3990577513	Tea Party News	tpartynews			
2912754262	Pigeon Today	PigeonToday	United States	America's Weakest Primetime Lineup Anywhere! Follow Merica's #1 cable news network, delivering you breaking news, insightful analysis, and must-see memes.	http://t...
2570631118	Polittweecs	Polittweecs	United States	Tweets and politics	
1930747698	Невские Новости	nevnov_ru	Санкт-Петербург, Россия	Новости Санкт-Петербурга и Ленинградской области	https://t...
2742327187	Washington Online	WashingtonOnline	Washington, D.C.	Breaking news, weather, traffic and more for Washington, DM us anytime. RT us!	

Screenshot 1: A capture of the spreadsheet of initial dataset provided by CSRI

tweetid	tweet_language	tweet_text	tweet_time	tweet_client_name	in_reply_to_tweetid	in_reply_to_username
877343178537353217	en	This is priceless! #FakeNews #GA06 #Handel #Ossoff https://t.co/Fn1E1Ktqbt	21/06/2017 01:51	Twitter Web Client		
878369321604481024	en	VIDEO: Biker Revs Engine, Drives Through Anti-Trump Activists Laying in the Street for a 'Die-In' RT if you'd buy the biker a beer! https://t.co/IY9jwOzMiR	23/06/2017 21:49	Twitter Web Client		
890352159031611392	en	The Assistant US Attorney overseeing the Awan family investigation is Steven Wasserman, the brother of Wasserman Schultz! https://t.co/iRZsUMJRHO	26/07/2017 23:24	Twitter Web Client		
815736482874859520	en	Trump wants to invite Netanyahu to the inauguration. White House "furious". "Unprecedented disrespect" to Obama https://t.co/m74V50TeyQ	02/01/2017 01:48	TweetDeck		
760982090951208960	en	France is lost. #French police try to enter a muslim no-go zone. We can not allow this in America! https://t.co/JtQyIPMhE1	03/08/2016 23:34	Twitter Web Client		
801219752408457216	en	Clint Eastwood Refuses to Accept Presidential Medal of Freedom From Obama, Says "He is not my president" #MedalofFreedom https://t.co/t60YRf49jz	23/11/2016 00:24	Twitter Web Client		
869674406028734464	en	RETWEET if you think Jeff Sessions should indict Kathy Griffin for threatening the President! https://t.co/UbHg3SkU6w	30/05/2017 21:58	Twitter Web Client		
874019843527368704	en	A high school senior in South Carolina was asked to remove an American Flag from his pickup truck. This was the community response https://t.co/I2O2RqUXxO	11/06/2017 21:45	Twitter Web Client		
871145427613224960	en	I'll just leave this right here #LondonBridge #BoroughMarket #Vauxhall https://t.co/ZyWdIENi08	03/06/2017 23:23	Twitter Web Client		
86246774685302784	en	Obama just took a jet & a 14-car caravan to Italy to give a \$2,5mill lecture about living	11/05/2017 00:41	Twitter Web Client		
826166845124313088	en	I would rather take care of TEN homeless US Veterans, than 50,000 migrants/illegal aliens				
		How About You? #BoycottStarbucks #ReasonsToProtest https://t.co/y10hsjxPCW	30/01/2017 20:34	Twitter Web Client		
		"Hillary Clinton asks judge to toss defamation case filed by parents of ' Benghazi' Hall of				

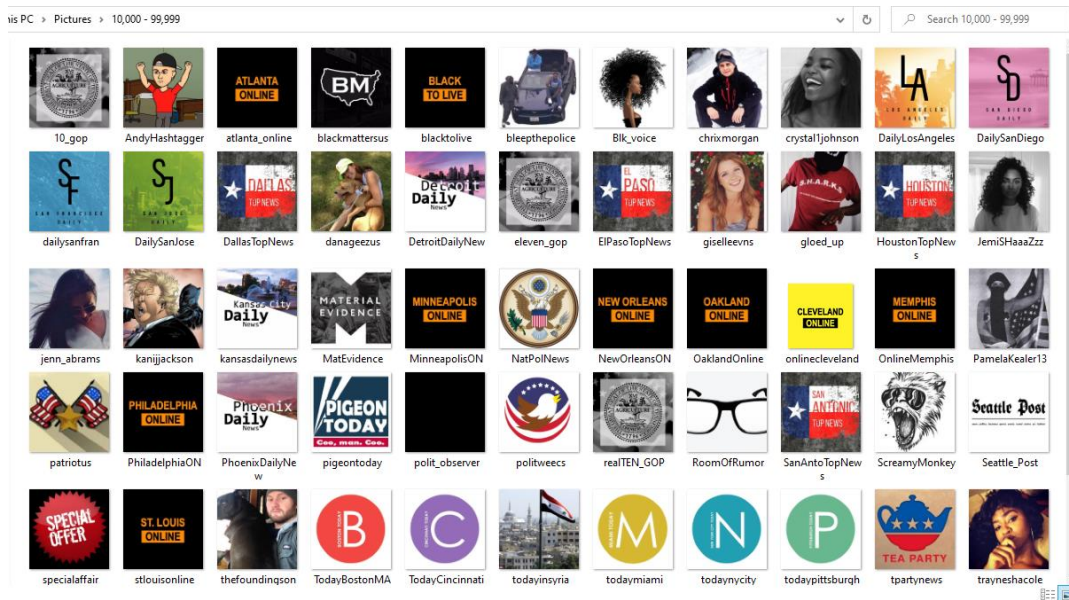
Screenshot 2: A capture of the spreadsheet of tweets created by five main accounts provided by CSRI

Handle	name type - handle	identity aspect - handle	second identity aspect	third additional identity aspect	fourth identity aspect	gender
Jenn_Abrams	eponym	name	gender	race		female
TodayNYCity	pseudonym	location	news			
ELEVEN_GOP	pseudonym	political affiliation				
CrystalJJohnson	eponym	name	gender			female
TheFoundingSon	pseudonym	gender				male
WashingtOnline	pseudonym	location	news			
politweecs	pseudonym	topic of interest				
NewOrleansON	pseudonym	location	news			
todayinsyria	pseudonym	location	news			
KaniJJackson	eponym	name	gender	race		female
WarfareWW	pseudonym	location	topic of interest			
PigeonToday	pseudonym	animal	news			
BlackToLive	pseudonym	race				
KansasDailyNews	pseudonym	location	news			
gloed_up	pseudonym	personality trait				
GiselleEvns	eponym	name	gender			female
DailyLosAngeles	pseudonym	location	news			
SanAntoTopNews	pseudonym	location	news			
TrayneshaCole	eponym	name	gender	race		female
Seattle_Post	pseudonym	location	news			
DanaGeezus	eponym	name	gender	race		female
DallasTopNews	pseudonym	location	news			

Screenshot 3: Example of the data spreadsheet used for the analysis of handles and usernames

Handle	Username	identity aspect - profile picture	second identity aspect	third identity aspect	fourth identity aspect	race
Jenn_Abrams	Jenna Abrams	person	race	gender		white
TodayNYCity	New York City Today	location	news			
ELEVEN_GOP	ELEVEN GOP	location				
CrystalJJohnson	Crystal Johnson	person	race	gender		Black
TheFoundingSon	John Davis	person	race	gender		white
WashingtOnline	Washington Online	location	news			
politweecs	politweecs	location				
NewOrleansON	New Orleans Online	location	news			
todayinsyria	Syria Today	location				
KaniJJackson	Kanisha Jackson	person	gender	race	political affiliation	white
WarfareWW	Warfare Worldwide	topic of interest	location			
PigeonToday	Pigeon Today	animal	news			
BlackToLive	Black To Live	race				Black
KansasDailyNews	KansasCityDailyNews	location	news			
gloed_up	1-800-WOKE-AF	person	race	gender		Black
GiselleEvns	Giselle Evans	person	race	gender		white
DailyLosAngeles	Los Angeles Daily	location	news			
SanAntoTopNews	San Antonio Top New	location	news			
TrayneshaCole	Traynesha Cole	person	race	gender		Black
Seattle_Post	Seattle Post	location	news			
DanaGeezus	Dana Gold	person	race	gender		white
DallasTopNews	Dallas Top News	location	news			

Screenshot 4: Example of the data spreadsheet used for the analysis of profile pictures



Screenshot 5: Example of the image file used in accordance with Screenshot 4 for analysing profile pictures

tweetid	tweet_language	tweet_text	tweet_time	tellership	tweet format	verified
896390383474757632		BREAKING: At least 2 people injured in violent clashes at #EmancipationPark ahead of the #Charlottesville rally https://t.co/1d3ccle1WW	12/08/2017 15:18	single	original	
896397469491372032	en	BREAKING: State of emergency declared for #Charlottesville because of #UniteTheRight rally chaos. https://t.co/136rVeN4gA	12/08/2017 15:46	single	original	
896403638091993088	en	Riot police on scene. #Charlottesville officials have declared the #UniteTheRight rally an unlawful assembly. https://t.co/4lrECMIWw	12/08/2017 16:10	multiple	original: reported speech	
896405393735286784	en	@bakedalaska was badly pepper sprayed during the #UniteTheRight rally. #Charlottesville https://t.co/Vp7iyVFOHj	12/08/2017 16:17	single	original	
896408251553775616	en	ATBroderickGreer Spencer wasn't arrested. He is safe and going to another location https://t.co/gmS5j1XFxg	12/08/2017 16:29	multiple	reply: other	verified
896410764747177985	en	RT @DineshDSouza: "We're socialists...enemies of today's capitalist system of exploitation & we're determined to destroy it"--Bernie, 2017? No, Hitler, 1927	12/08/2017 16:39	multiple	retweet	verified
896410973187264513	en	RT @FLOTUS: Our country encourages freedom of speech, but let's communicate w/o hate in our hearts. No good comes from violence. #Charlottesville	12/08/2017 16:40	multiple	retweet	verified
896413263751237632	en	The violence that's happening in #Charlottesville is getting crazy! Totally out of control. https://t.co/ePFoEgazy	12/08/2017 16:49	single	original	
896416600777248768	en	RT @Timcast: People chanting "Love has already won" in #Charlottesville. Sorry but no, ya'll hate just as much as the other side.	12/08/2017 17:02	multiple	retweet	verified
896418787938979840	en	Whatever your political view is, you have to condemn the violence, regardless of who's committing it. #Charlottesville	12/08/2017 17:11	single	original	
896431640951521280	en	RT @Cernovich: ANTIFA, a left wing organization, threw urine on a left wing media	12/08/2017 18:02	multiple	retweet	verified

Screenshot 6: Example of the spreadsheet used for analysis of (co-)tellership

Table 25: Coding categories for handles and usernames

Identity aspect category	How do handles and/or usernames in this category identify themselves?	Examples of where each code was applied
Geographical Location	Based on a geographical location i.e., one that can be	@todayinsyria Syria Today, @redlanews Red Louisiana

	found on a map on multiple levels e.g., nation, state, city.	News, @HoustonTopNews Houston Top News
News	Using language that either explicitly or implicitly relates to the genre of news sharing.	@DallasTopNews Dallas Top News, @DailySanDiego San Diego Daily, @TodayNYCity New York City Today
Proper Name	Using existing first names that have been established as being used to identify a person in the offline sphere e.g., on baby name websites/in census surveys.	@Jenn_Abrams Jenna Abrams, @Crystal1Johnson Crystal Johnson, @TrayneshaCole Traynesha Cole
Gender	By indexing gender either explicitly or via typical associations with first names.	@TheFoundingSon John Davies [both 'son' and 'John' index male identity], @GiselleEvns Giselle Evans [typically a name given to females]
Hobbies/interests	Drawing on a particular topic of interest, subject, or pastime.	@AndyHashtagger, @politweecs politweecs, @USA_GunSlinger GunSlinger Girl
Race	By either explicitly mentioning a racial identity or using a first name that has a racial distinction established in previous research or via Google Images.	@blackmattersus Black Matters, @PamelaKealer13 Pamela Kealer, @KaniJJackson Kanisha Jackson
Political affiliation	Drawing on partisan political positions (e.g., left or right).	@tpartynews Tea Party News, @ELEVEN_GOP Eleven GOP, @March_for_Trump March for Trump
Personal attribute	By describing the user in terms of attributes, whether these be physical or related to how a user views themselves more generally (e.g., personal distinguishing quality)	@gloed_up 1-800-WOKE-AF, @patriototus Being Patriotic, @wokeluisa
Animals	Using animal terms	@ScreamyMonkey Screamy Monkey, @pigeontoday Pigeon Today
Organisation	Using language that suggests the account belongs to an organised group of people rather than an individual	@SpaceDept Department of Space
Religion	Using language that evokes connotations with religion	@Jihadist2ndWife Jihadist Wife

