

Unpacking Disinformation as Social Media Discourse

Johan Farkas and Yiping Xia

In this chapter, we examine the role of Discourse Studies in social media disinformation research. While currently underrepresented, Discourse Studies can provide key insights into why disinformation gains traction through credibility building, tapping into existing political narratives and stereotypes. Discourse Studies, we argue, can also bring much-needed attention to the constitutive role of antagonism in disinformation and to the connection between political practices, power relations and platform designs; aspects that are often overlooked. Drawing on three empirical cases – revolving around the Russian Internet Research Agency, fake Muslim Facebook pages and far-right conspiracy theories disguised as tabloid news – the chapter aims to provide a clearer view on the application of Discourse Studies (in its various forms) to disinformation.

Introduction

Disinformation on social media has in recent years become a prominent area of both democratic concern and social scientific attention, highlighted by scholars as “*the* defining political communication topic of our time” (Freelon and Wells 2020, 1, original emphasis). The issue rose to prominence following the 2016 U.S. presidential election, when it was revealed that social media was used for foreign interference as well as new forms of micro-targeted advertisements. In relation to the former, a Russian organization with ties to the Kremlin, known as the Internet Research Agency (IRA), operated thousands of fake accounts on Facebook, Twitter, Instagram and other platforms. In terms of the latter, a British-American company, Cambridge Analytica, used personal data on millions of Americans – obtained without consent - to target voters on behalf of the Trump campaign. This company has since been accused of collaborating with the IRA (Palma 2018; Wylie 2019). These revelations coincided with the rise of “fake news” as a near-ubiquitous signifier in political debates – used by journalists and scholars to describe falsehoods in newslike packaging (Tandoc, Lim, and Ling 2018) and by political actors to attack and delegitimise perceived political opponents (Farkas and Schou 2018; Habgood-Coote 2018).

In the wake of these developments, research into online disinformation has grown considerably in the past years, with scholars from a variety of disciplines taking on the topic from different methodological and theoretical perspectives. This has brought forth new insights into deception strategies, aims and the socio-technical dissemination of disinformation in different political landscapes and hybrid media systems. This chapter aims to highlight and discuss one particular research approach to disinformation, namely *Discourse Studies*. Despite remaining underrepresented in disinformation research, Discourse Studies holds potential for providing new insights into crucial aspects of political manipulation on social media.

In this chapter we will first argue that, ontologically, Discourse Studies is relevant to the move beyond a narrow preoccupation with binary characterisations of content, notably the distinctions between true and false, real and fake, or bot and human. By emphasising the semiotic and historical contingency of knowledge, social identities and political struggle, Discourse Studies (in its various forms) emphasises the entanglement of empirical phenomena and broader political contexts. This makes it well suited for studying social media disinformation as an interconnected web: a discursive ensemble built upon the construction and mobilisation of social identities and the amplification of political antagonism, rather than of just collections of truths, half-truths, and falsehoods. Our second main argument is that Discourse Studies helps us attend to techno-discursive dynamics of social media disinformation, i.e. the interplay between social, political and technological relations (Khosravinik 2017, 2018, 2020; Unger, Wodak, and KhosraviNik 2016). As we will explicate later, disinformation scholarship can benefit from such an engagement with Discourse Studies, approaching cases as discursive constructs arising from the entanglement of norms, practices, power relations and platform designs.

A key aim of disinformation is to amplify political contestation, discord, fear and tension between social groups (Benkler, Faris, and Roberts 2018; Marwick and Lewis 2017). This is often done by building credibility within target demographics, for example through a “performance of authenticity” (Xia et al. 2019) via fake personas on social media platforms. Despite the significance of (perceived) credibility, authenticity and antagonism, little research so far has provided in-depth, contextual analyses of these aspects. Indeed, researchers have tended to apply a more or less decontextualised focus on digital content, treating it primarily as a binary issue of true/false or fake/real (Phillips 2020). As we will unfold, this has limited the field.

The political and socio-technical contexts of disinformation campaigns are crucial to their workings. We argue that Discourse Studies can provide important

contributions in this regard. This is not least due to the long-standing emphasis in this tradition on the semiotic “constitution of social identities and social relations” (Jorgensen and Phillips 2002, 9) and "discursive construction of politico-ideological frontiers and the dichotomisation of social spaces" (Laclau 2000, xi). Discourse studies provides a relational and politico-semiotic approach to social phenomena, bringing questions of identity and antagonism to the forefront as well as connections to wider “socio-political and structural context[s]” (Unger, Wodak, and KhosraviNik 2016, 279). In relation to disinformation, this means addressing vital questions such as: How do fear-mongering narratives resonate with or build on existing political discourses? How is “the fabric and ingredients of the content ... strategically designed” (KhosraviNik 2020) on social media? And how do social media platforms play a role in (co-)shaping political struggles and manipulation?

Drawing on findings from three case studies, the chapter aims to discuss the strengths and contribution of Discourse Studies in the context of disinformation. The studied cases span the years 2015 to 2017 and focus – in terms of geography – on the United States and Denmark and – in terms of platforms - on Twitter and Facebook. These revolve around:

1. Fake Muslim Facebook pages in Denmark, sparking thousands of hateful and racist reactions from users by claiming to represent “Muslims” in Denmark taking part in a widespread conspiracy to kill and rape “Danes” (Farkas, Schou, and Neumayer 2018)
2. “Jenna Abrams”, an influential fake Twitter account operated by the Russian IRA, claiming to represent a female, conservative U.S. citizen who decried liberal political figures in language appealing to a conservative community (Xia et al. 2019)

3. Letters to the editor posted on the website of the largest Danish tabloid newspaper, created without editorial supervision by far-right activists to spread racist conspiracy theories disguised as tabloid news on social media (Farkas and Neumayer 2020)

In the following sections, we will first review key literature on online disinformation and Discourse Studies, before turning to the three cases.¹ By reflecting on the role of Discourse Studies in relation to each case, our aim is to highlight the overall merits of this tradition within disinformation research.

Disinformation on Social Media

Disinformation on social media is a growing interdisciplinary research area, studied by scholars from an array of fields, including political science (Golovchenko, Hartmann, and Adler-Nissen 2018), psychology (Lewandowsky, Ecker, and Cook 2017), linguistics (see Block 2019), philosophy (Rini 2017) and media and communication studies (Freelon and Wells 2020; Marwick and Lewis 2017). While political manipulation is certainly not new (Ellul 1965), the rise of digital platforms has sparked increased attention to new forms of deception.

While terminology varies across disciplines and studies, definitions of disinformation are generally predicated on an intentionality to deceive or mislead. For example, Jack (2017, 3) defines it as “information that is deliberately false or misleading”. Similarly, the “High Level Expert Group on Fake News and Online Disinformation of the European Commission” define it as “false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit” (European Commission 2018, 10). These definitions

¹ Longer and detailed versions of these case studies have been published in the journals *Critical Discourse Studies*, *Information, Communication, and Society*, and *Nordicom Review*.

also encompass what has been described as fake news: “viral posts based on fictitious accounts made to look like news reports” (Tandoc et al. 2018, 138). Increasingly, however, scholars have encouraged peers to move “beyond fake news” (Freelon and Wells 2020, 3) due to the term’s ambiguity and increasing politicization (Habgood-Coote 2018; Farkas and Schou 2018).

Although the concept of disinformation is useful for empirical analysis, its usage comes with the risk of over-accentuating binary emphases on true versus false, intentional versus unintentional. As Phillips (2020) notes, research on disinformation often limits “discussions to the basic assertion *that* a particular story is false, rather than encouraging reflection on *why* the story resonates with audiences” (56, original emphasis).

Disinformation campaigns are complex political, cultural, and socio-technical phenomena that rely on a mix of different types of content, platforms and sources. Here, binary analyses of true versus false, intentional versus unintentional fall short, as they fail to explain why certain narratives gain traction in specific contexts, while others do not. Understanding this requires in-depth examinations of wider discourses and social identities.

Since the election of President Donald Trump in 2016, research into disinformation on social media has surged (Freelon and Wells 2020). Scholars have argued that “context collapse” (boyd 2010) and the lack of information gatekeepers on social media propels users to rely more on cultural signifiers – in particular those that speak to their identities – when evaluating messages (Marwick 2018; Tripodi 2018). Political actors have seized these openings in social media spaces, clinging onto the exploitation of partisan polarization (Faris et al. 2017; Xia et al. 2019) and racial identity (Daniels 2009), among other deep-rooted cultural mindsets.

Among online disinformation campaigns, Russia's IRA has garnered the most political and scholarly attention. This campaign spanned multiple countries and all major social media platforms in Europe and the U.S., including Twitter, Facebook, Instagram, and Reddit (DiResta et al. 2018; Lukito 2019). On Twitter, researchers have identified thousands of fake accounts targeting U.S. citizens, impersonating, among others, left-wing users, right-wing users, Black Lives Matter activists and local news outlets (Freelon and Lokot 2020; Linvill and Warren 2019; Bastos and Farkas 2019). In doing so, this campaign clearly intended to captivate U.S. users via identity-conforming messages, tailor-made with a sense of constructed authenticity (Xia et al. 2019).

Beyond the IRA, cases of disinformation similarly revolve around manipulation of socio-cultural identities and feigned authenticity. For example, Daniels (2009) has shown how white supremacists use "cloaked websites" to project cultural legitimacy. Marwick and Lewis (2017) have shown how media manipulators exploit online participatory culture to push racism and misogyny towards mainstream audiences. Scholars of online trolling and harassment have argued that ambivalent identities in internet culture undergird such aggression (Phillips and Milner 2017).

Despite disinformation campaigns being deeply context-dependent, the dominant research paradigm has emphasised binary distinctions between true and false with little emphasis on context. As Hedrick, Kreiss, and Karpf (2018) critically note, research has largely "occurred in a vacuum, often ignoring the deeper political, social, and cultural contexts from which they have emerged" (1059). In such scholarship, disinformation has been treated as an amalgam of decontextualized texts, analysed with the aim of classifying messages according to their truthfulness or likelihood of deriving from bots. Researchers have also used disinformation messages as independent variables to measure their – usually short-term – effect on user beliefs and behaviours (e.g. Bail et al. 2020). As argued by Marwick (2018), such studies shed much-needed light on disinformation, yet

ultimately risk falling short of thoroughly understanding it. This is due to the tendency to see disinformation as a “magic bullet”, thus “disregarding the structural influence of problematic patterns in media messaging and representation” (Marwick 2018, 485). Marwick further points out that researchers, for this reason, “may underestimate the engagement that people have with problematic or ideologically-driven information online” (487).

Emphasis on true/false dichotomies risks trading in-depth research for a simplistic remedy: to save democracy, just eliminate falsehoods. This can legitimise anti-democratic solutions, such as state censorship (Lim 2020; Farkas and Schou 2019). Much of current disinformation research stops short at the point of mapping the ostensible features and scales of falsehoods. Even the studies that “debunk” the power of “fake news” and propaganda (e.g. Bail et al. 2020; Guess et al. 2020) assume that fake or inauthentic *content* itself is of most concern. This outlook neglects the ambivalent dynamics surrounding people’s exposure to online manipulation, chiefly shaped by social identities and political discourses. As Daniel Kreiss suggests,

Our field has been woefully inadequate at addressing the communicative aspects of social identity. Even more, we have largely ignored the ways that identity shapes epistemology — that social identities come prior to what people evaluate as true or false. (Kreiss 2019b)

Research into the mutual shaping of social identities, epistemologies and meaning-making requires contextual and socio-political perspectives, something often lacking in current scholarship. While no theoretical or methodological framework can do a catch-all job, we argue that Discourse Studies is well-suited for addressing key aspects of such an endeavour. This includes studying how social identities, antagonism and political struggles are mobilised in disinformation

campaigns. Instead of viewing disinformation messages as “magic bullets” wreaking havoc on democratic societies, Discourse Studies takes full account of connections to wider narratives, political struggles and social identities. Discourse Studies also rejects binary notions of true and false, emphasising the underlying power relations of all meaning-making. In the following section, we will discuss the application of Discourse Studies in the context of social media, before turning to our three case studies.

Discourse Studies and Social Media

Discourse Studies encompasses a number of overlapping research schools that share an overarching commitment to the “semiotic dimensions of power, injustice, and political-economic, social or cultural change in society” (Unger et al. 2016, 278). They also share a number of fundamental premises: knowledge is historically contingent, political struggles and power relations shape meaning-making, and social identities are relationally constituted (Laclau 2014; Wodak 2009; Potter 1996; Krzyżanowski and Forchtner 2016). Beyond this shared foundation, different schools have “different theoretical models, research methods and agenda” (Fairclough, Mulderrig and Wodak 2011, 394). Two important schools of Discourse Studies are *Critical Discourse Analysis*, associated with Fairclough (1995) and Wodak (2009), and *The Essex School of Discourse Theory*, associated with Laclau and Mouffe (2014). While they diverge in some key areas, they also share many similarities. As noted by Torfing (2005): “when it comes to the actual analysis of social and political discourse, the differences between Fairclough and Laclau and Mouffe are small” (9).

In this chapter, we will emphasise the shared premises and similarities across Discourse Studies and discuss the merits of this tradition overall in the context of disinformation and social media. As will become clear through the discussed

case studies, our own research has drawn respectively on Critical Discourse Analysis (Fairclough 1995, 2003) and the Essex School of Discourse Theory (Laclau and Mouffe 2014; Laclau 2014). This has led us down similar paths, examining the discursive construction of credibility, trust, and authenticity in disinformation campaigns as well as the constitutive role of antagonism. As such, our aim is to illuminate our shared paths and reflect on the overall strengths of Discourse Studies.

Discourse Studies has a long history in media scholarship, especially in research on media's role in shaping knowledge, power relations and political struggles (Hall 1982; van Dijk 1985; Fairclough 1995; Wodak 2009). As Stuart Hall argued in 1982, media systems play a crucial role for meaning-making due to the way they "selectively circulate... *preferred* meanings and interpretations" (Hall 1982, 341, original emphasis). Media institutions help us "not simply to know more about 'the world' but to make sense of it. Here the line... between the 'meaningless' and the 'meaningful,' between the incorporated practices, meanings and values and the oppositional ones, is ceaselessly drawn and redrawn" (Hall 1982, 341). While media systems have changed considerably since Hall's (1982) seminal work, media institutions and technologies still play a crucial role in shaping knowledge, power and identity boundaries - all of which are subject to political struggles and deliberate attempts at manipulation.

Many social practices and struggles over signification have in recent years moved to the realm of social media. Here, a few platforms have grown increasingly hegemonic, most notably U.S.-based Meta (controlling Facebook, Instagram and WhatsApp) and Chinese-based Tencent (controlling Tencent QQ, WeChat and more). These platforms are reshaping diverse aspects of human life, including news consumption, political deliberation, and everyday sociality (van Dijck 2013).

Researchers within Discourse Studies have increasingly examined social media communication. Most prominently, KhosraviNik (2017, 2020) has developed an approach dubbed *Social Media Critical Discourse Studies* (SM-CDS), building on Fairclough and Wodak (1997) and integrating the digital dynamics of discursive production and consumption on social media. This approach emphasises attentiveness towards socio-technical characteristics of digital platforms, including interactivity, intertextuality, and multimodality, as these play key roles in shaping online discourse (KhosraviNik 2017, 2018, 2020). Social media platforms revolve around “techno-discursive design” (KhosraviNik 2018, 440), facilitating the interplay between digital and social practices. Function such as the “like” button as well as algorithmic dissemination of content not only enable communication, but are also part of discursive dynamics (KhosraviNik 2017, 2018, 2020).

A key aspect of social media’s techno-discursive design is the ability to create personal accounts, including fake personas. Social media revolve around decentralised content production, meaning that news, videos or stories can derive from a near-endless number of sources that can be difficult to verify. Powerful and resource-rich actors take advantage of this by manipulating audiences on a wide scale, for example in cases of state-backed propaganda agencies deploying thousands of fake accounts. This poses key challenges for researchers and journalists aiming to critically examine political discourses online.

On social media, boundaries between actors and discursive practices are often difficult to pin down (Farkas, Schou, and Neumayer 2018). It is often challenging to trace the origins of narratives, sources and underlying intentions (Phillips and Milner 2017). At the same time, as highlighted by Krzyżanowski and Ledin (2017), “social media has fostered the rise of various agorae of exchange of views which often escape the traditional norms of political expression by progressively ‘testing’ as well as ‘stretching’ norms of publically-acceptable languages” (567). This has given rise to *borderline discourse* (Krzyżanowski and

Ledin 2017), the deliberate blurring of boundaries between civil and uncivil discourse.

Researchers from Discourse Studies have emphasised the importance of examining the techno-discursive design and layered nature of social media communication (KhosraviNik 2017). On social media, communication encompasses “not only what is said, but also specific information about the profile of the user sending out a message, the users receiving that message, [and] about how users interact with a message” (Langlois and Elmer 2013, 2). A platform like Facebook is not simply a neutral or transparent conveyor of ‘text,’ but a “sociotechnical ensemble[s] whose components can hardly be told apart” (van Dijck 2013, 14). Despite this, however, much social media research has tended to focus on content only, as KhosraviNik (2020) summarises:

In practice, a large body of studies on SMC [social media communication] are predominantly focused on the content... rather than dealing with how the participatory web may have changed the politics of discursive dynamics, the quality of the very content and the overall structure of discursive participation. (2)

Following KhosraviNik (2020), we argue that Discourse Studies must capture the complex interplay on social media between digital technologies, social practices, cultural identities and hegemonic relations. This means capturing both the “socio-political context of users in society” (KhosraviNik 2017, 585) and how “diverse elements and actors (human and non-human, informational, communicational, and political) are mobilized and articulated in specific ways” (Langlois et al. 2009, 416–417).

In this regard, The Essex School of Discourse Theory (also known as simply *Discourse Theory*) offers a powerful framework for studying relational constructions of social identities, political struggle and “the dimension of antagonism that is inherent in human relations” (Mouffe 2000, 15). Through concepts such as

hegemony, sedimentation, reactivation, nodal points, and logics of difference and equivalence, discourse theory provides a vocabulary that brings attention to political struggles and power relations behind meaning-making (Laclau and Mouffe 2014; Laclau 2014). Within the context of disinformation, this framework enables us to study the discursive construction of both credibility and political antagonism.

As we will outline in the following sections through three case studies, Discourse Studies – both in the tradition of Critical Discourse Analysis and Discourse Theory – can serve as a productive approach for studying disinformation. Such research is particularly adept for examining the creation of (perceived) credibility, trust, and authenticity as well as amplification of political antagonism, all of which is key to disinformation despite remaining under-researched. The aim of the following is thus to exemplify how Discourse Studies can be mobilised empirically and to reflect on the merits of such an endeavour.

Fake Muslim Facebook Pages and Discourse Theory

In 2015, a number of fake personas on Facebook managed to attract thousands of angry and racist comments from Danish Facebook users by claiming to represent Muslims living in Denmark. Supposedly, these Muslims were part of a grand conspiracy to overtake the country, raping and killing (white, Christian) Danes in the process (Farkas, Schou, and Neumayer 2018). By (crudely) constructing fake Muslim identities through a tapestry of images, videos, texts and hyperlinks, anonymous culprits spread well-known far-right conspiracy theories about Muslims and immigrants in a disguised form:

I am a proud Muslim. Denmark will be a Muslim country because it is Allah's will. We are the ones who laugh while we take your money, women and houses. :)

(“About” section of Mohammed El-Sayed, 1 July 2015)

Our case study focused on 11 fake Muslim Facebook pages that attracted more than 20 000 comments, many of which expressed hatred towards both the individual (fake) identities as well as Muslims in general. A member of the Danish parliament even shared one of the pages, seemingly believing in the credibility of the source. In order to understand this spectacle of deception and hostility, we examined both the (crude) construction of fake identities on these pages and the antagonism created through dichotomous narratives of “Muslims” versus “Danes.” To do so, our study relied on Discourse Studies, specifically the theoretical vocabulary of the Essex School of Discourse Theory (Laclau 2014; Laclau and Mouffe 2014).

The first step of our analysis examined the discursive construction of both individual and collective Muslim identities on the Facebook pages. On social media, individual identities are established and reproduced through the continuous curation of images, texts, and videos (Hogan 2010) as well as interactions with (other) users. On the fake pages, this curation involved the appropriation of Arabic names, profile pictures, images from protests (for example against the Danish Mohammed cartoons), common Arabic phrases (e.g. ‘*Mashallah*’), “Islamist” symbolism (such as the so-called Black Standard) and hyperlinks to existing [pan-Islamic](#) political organizations (Farkas, Schou, and Neumayer 2018). In order to establish that these individual identities were indeed false, we analysed patterns across multiple pages, highlighting a cyclical use of made-up Arabic names, narratives, images and affiliations.

Through discourse theory, our analysis focused on how the fake personas constructed authenticity and sparked user reactions by tapping into existing fear, xenophobia, and antagonism in Denmark (Hervik 2011). By constructing fake identities whose entire meaning derived from an oppositional relationship to “the Dane”, these Facebook pages projected a common fear-mongering narrative of “us” versus “them” found in European far-right politics – what has been characterised as a “politics of fear” (Wodak 2015, 186).

When studying the comment sections of the Facebook pages, we found a large number of users reproducing the dichotomous identities of “the Muslim” and “the Dane”, expressing general hatred towards Muslims and immigrants (often used interchangeably). We thus found that users seemingly accepted, not only the credibility of the individual fake personas, but also a larger narrative of Muslims being dichotomous adversaries to Danes. In turn, these user responses became part of the continued performance of authenticity of the (fake) Muslim identities, as the comment sections “supported” or “confirmed” both their credibility and antagonism.

The anonymous administrators of the Facebook pages exacerbated the role of user comments by systematically removing comments expressing that the pages were fake. It cannot be excluded that the anonymous administrators also contributed with comments of their own to bolster the dichotomous narrative of “Muslims” versus “Danes”. Considering these aspects, it is clear that Facebook’s techno-discursive design played a key role in shaping both deception and antagonism, enabling the construction of fake identities, anonymity of content creators and selective removal of comments.

By approaching cases of fake identities from a discourse theoretical perspective, we get an in-depth understanding of the construction of credibility around fake personas as well as the political discourses they amplify and hateful reactions they receive. Building directly on existing stereotypes and fear-mongering narratives in Denmark (Hervik 2011), fake personas managed to spark outrage from Facebook users, amplifying existing xenophobia and racism. The analysis shows how manipulation and antagonism can arise from the interplay between the (fake) identities, existing discourses, user practices and what can be described as a “hijacking” of “social media sites’ technical infrastructure” (Matamoros-Fernández 2017, 935). Instead of applying a binary view on disinformation centred solely on deciphering what is true and what is false, this study focused on questions of how and why disinformation operates on a techno-discursive level. Through a

discourse theoretical approach, questions of identity, power and political struggle were brought to the forefront.

Fake Twitter Discourse and Critical Discourse Analysis: Jenna Abrams

With more than 70,000 followers, “Jenna Abrams” (@Jenn_Abrams) was the second most followed English-speaking account from the Russian IRA by the time Twitter suspended IRA-related accounts in 2017. This account impersonated a conservative, white, female U.S. citizen who was highly expressive about political issues and general interest topics. Her tweets managed to stir up online debates about divisive political issues and appeared in over 30 media outlets, including the *New York Times*, the *Huffington Post*, and *CNN* (Collins and Cox 2017).

Our initial research interest in this “sockpuppet” account stemmed from reading a specific tweet: “Use #AskJennaAbrams tag or email to ask me your questions I'll answer on Sunday on <https://t.co/XGTOAyM8rS>”. Knowing that a Russian team operated “Jenna Abrams”, we were surprised by the effort put into interacting with followers, as well as her seemingly large and committed following – enough to hold a Q&A session. Further investigation only brought up more unexpected facts: the Q&A session was not a one-off but was announced at least three times in her Twitter feed; the URL mentioned in the above tweet directed users to a WordPress blog that also belonged to “Jenna Abrams.” Pursuing the blog would lead the reader to “Jenna’s” profiles on other platforms, including Telegram, Medium, and an email address.

That the IRA campaign went beyond Twitter and operated across platforms has since been studied by other researchers (DiResta et al. 2018; Lukito 2019) – attesting to the intertextuality and complexity of the operation. By applying Discourse Studies - in this case Critical Discourse Analysis (Fairclough 1995, 2003)

– to Jenna Abrams’ messages on Twitter and beyond, we were able to shed light on the “intertextuality among textual practices on (potentially) multi-sites” (KhosraviNik 2017, 585).

What we see in Jenna Abrams’s case is a calculated campaign to exploit social media self-branding – establishing a female, white, conservative persona spanning a handful of platforms – for the purpose of optimising a performance of personal authenticity. Here, critical discourse analysis enabled us to unpack this performative manipulation through the lens of multimodality and techno-discursivity. Specifically, the Abrams account took full advantage of Twitter’s affordances, including a profile photo, header banner, personal bio and more to build authenticity. Using a picture of a young, white woman as her profile photo, which turned out to be from a 26-year-old Russian woman (O’Sullivan 2017), Abrams wrote in her bio that she lived in “the USA”. The bio also claimed that she was “pro-common sense”, an implicit attack of liberals who she often denounced as “hypocritical” and “lacking common sense.” Moreover, Twitter’s features such as “@ reply” and hashtags were often used to boost the account’s visibility. Specifically, in the early stages of the operation, Jenna Abrams frequently “@ replied” to prominent political figures and mainstream media outlets and used trending political hashtags such as #ISIS -- both tactics useful for relatively unknown accounts to gain followers. Together, Abrams’s multimodal operation exploited the digital architecture of Twitter in building credibility and amplifying discord among the U.S. public.

As the previous example of Q&As vividly shows, interactivity lies at the heart of such an operation. The Abrams account often signalled to followers her love for American democracy and popular culture, and maintained a conversational, light-hearted style in her interactions. For example, she once tweeted about her attachment to the U.S., her “home country”, saying: “I am afraid that one day incompetent and greedy politicians will ruin the #US, our home! #WhatAreYouAfraidOf.” She also commented on American celebrities to demonstrate

her authenticity as an American and spark user interactions. Such techniques of “authentic” self-expression are no secret for social media strategists and well-studied by Internet scholars (boyd and Marwick 2014), but the importance of performed authenticity in disinformation practices has just begun to be noticed by scholars.

Abrams’s authenticity work would have been pointless but for the IRA’s second set of tactics, embracing familiar conservative narratives and discourses. Let us return to the wider political context – we now know that the goal of the IRA campaign was to “sow discord among the American public” (Grand Jury for the District of Columbia 2018). To this end, the IRA deployed a host of Twitter accounts – Jenna Abrams included – impersonating American conservative users and participating in political discourse on Twitter, while at the same time setting up liberal-leaning accounts to contribute with polarising counter-narratives, notably through accounts that avowedly supported the Black Lives Matter movement (Freelon et al. 2020). The chief task for the Jenna Abrams account, then, was to establish her credibility as a conservative in-group member. The account performed its competence in conservative culture by posting news or right-leaning commentaries on contentious issues, repeatedly loathing liberal’ “hypocrisy” and “lack of common sense”, engaging in conservative insider lingo that mocked and attacked discourses of the American left. For instance, one of the repeated themes was hatred towards then-President Barack Obama. Two tweets exemplify this effort:

Have a good Tuesday, my friends and remember that it’s 416 days till Obama leaves the Office.

It’s almost 8 years of Obama and people don’t know which bathroom to use.

Other political topics of interest to the Abrams account included the Black Lives Matter movement, feminism, and refugee policies. Her commentary on these divisive issues served to evoke a partisan identity (in this case, that of conservatives) that deeply shapes how Americans view politics and the world today (Kreiss 2019a).

The case of Jenna Abrams demonstrated that calculated identity work on social media platforms, when astutely designed to fit in the technical infrastructure, can generate credibility that serves political manipulation. Applying Critical Discourse Analysis to the Jenna Abrams case allowed us to pay detailed attention to the role of intertextuality, multimodality, interactivity, and bringing in wider political contexts, in order to unpack the work of this prominent fake identity.

Fake Letters to the Editor on ‘*Folkets Røst*’

Our last case of disinformation, studied through Discourse Studies, is The People’s Voice (‘*Folkets Røst*’) – a digital platform run by the Danish tabloid newspaper *Ekstra Bladet* from 2010 to 2018 (Farkas and Neumayer 2020). Our study found that prominent far-right activists in Denmark systematically used The People’s Voice to disseminate xenophobic narratives and conspiracy theories, disguised as tabloid news. The 50 most visible letters on social media from The People’s Voice were shared more than 120 000 times on Facebook, including by leading members of the Danish parliament who clearly mistook the letters for professional news articles.

Drawing on the Essex School of Discourse Theory (Laclau and Mouffe 2014; Dahlberg and Phelan 2011), our analysis examined the construction of legitimacy and credibility on The People’s Voice as well as how far-right activists used the letters to manipulate users and spread xenophobic antagonism. Visually, content on The People’s Voice was close to indistinguishable from articles from

Ekstra Bladet's newsroom, having the same overall visual presentation of content. Additionally, when shared on social media, users could only see the source as being "ekstrabladet.dk" - the top-level domain of the well-known tabloid – making the source identical to professional articles from the media outlet. Prominent far-right activists exploited these visual ambiguities and similarities by systematically creating letters to the editor mimicking the tabloid journalistic genre with clickbait headlines, third person writing, hyperlinks, images, and references to their own work as "articles." By blurring already opaque boundaries between user-generated content and tabloid news, these actors successfully managed to disseminate fear-mongering and racist discourses about immigrants and Muslims in a manipulative newslike packaging, supported by an established newspaper driven by a click-for-profit incentive.

The letters contained a carefully constructed mix of cherry-picked, manipulated and false information, promoting well-known far-right conspiracy theories about Muslims, immigrants and "liberal elites." This was done through headlines, such as "German Newspaper: Merkel Will Use the Refugee Crisis to Create the United States of Europe", "The Terrorists are Pouring over the Borders", and "The UN and Goldman-Sachs: The EU Must be Destroyed for Capital Gain." Through discourse theory, we examined both the construction of these antagonistic narratives – building on existing tropes, stereotypes and narratives– and the use of Ekstra Bladet's digital architecture to construct credibility.

Drawing on Krzyżanowski and Ledin (2017), we argue that the letters to the editor represent systematic attempts to disseminate 'borderline discourse' in the form of uncivil narratives packaged in a civil guise. While the narratives on The People's Voice were identical to those on far-right blogs and websites, their packaging as tabloid news produced a veil of credibility and journalistic authority. The letters can thus be seen as part of a larger development, where far-right discourses have increasingly become "spread and effectively 'normalised'... online" (Krzyżanowski and Ledin 2017, 577).

Through the theoretical vocabulary of The Essex School of Discourse Theory, we unpacked how these letters do not simply constitute cases of “false information” or “fake news”. Instead, these letters revolve around systematically blurring boundaries between societal truths and falsehoods, journalism and user-generated content, news and opinion as well as civil and uncivil discourse, all supported by the infrastructure of both Ekstra Bladet and social media platforms. By exploiting the ambivalence of the techno-discursive design of The People’s Voice, the studied letters amplify existing fear-mongering discourses and conspiracy theories, prevalent in Denmark and beyond. This highlights the interplay between platform design and political discourses.

Conclusion

In this chapter, we have argued that Discourse Studies (in its various forms) offers a powerful framework for providing new, critical and contextual understandings of disinformation on social media, moving beyond a preoccupation solely with the falseness of content. This approach emphasises the contingency of knowledge and relational constitution of social identities, enabling in-depth examinations of how and why disinformation campaigns construct and amplify political narratives and identities and play into wider discourses. We have further argued that Discourse Studies is capable of unpacking the techno-discursive dynamics of social media (KhosraviNik 2018): how platforms facilitate the tactical construction of credibility and authenticity, as well as the role of antagonism in disinformation campaigns.

The three cases described in this chapter all revolve around the construction of credibility and authenticity to spread fear-mongering and polarising political messages. Existing research shows that these aspects lie at the heart of most disinformation campaigns (Marwick and Lewis 2017; Daniels 2009), though they

have not received adequate attention in disinformation scholarship (Hedrick, Kreiss, and Karpf 2018). While our cases differ in terms of geo-political contexts and platforms, they all touch on the tactical blurring of boundaries: between authenticity and inauthenticity, visibility and anonymity, civility and incivility, credibility and deceitfulness – all enabled by the techno-discursive design of specific platforms (Phillips and Milner 2017). It is exactly by navigating these techno-discursive boundaries that disinformation thrives, whether deriving from small groups of activists, as in the case of The People’s Voice, or from large-scale actors, as in the case of the IRA.

Content analyses based on binary distinctions such as “fake news” versus “real news” can only get so far, their importance in motivating this field notwithstanding. Disinformation actors like the IRA rely on a complex interplay between discourses, identities, and technological affordances. These techno-discursive elements need to be accounted for in analyses of disinformation campaigns. This requires researchers to search beyond characteristics of textual information alone, considering disinformation messages as organic narratives that portends to join with existing discourses and available media infrastructures.

As noted in the introduction, Discourse Studies has so far remained underrepresented in scholarship on disinformation. This is a shame since this multi-faceted research tradition is uniquely fitted for unpacking questions of how disinformation resonates with citizens based on “already existing fears and doubts” (Farkas and Schou 2018, 309), all the while exploring “who sees what in what context, and how the internal structure of the message is influenced by the intended circulation” (KhosraviNik 2020).

Future research could benefit from Discourse Studies on a range of central questions. This includes how disinformation campaigns construct credible identities on social media and how users relate to them; how questions of race, gender, class and political struggle are mobilised; the role of borderline discourse

(Krzyżanowski and Ledin 2017); and how the technical infrastructure of social media is tactically appropriated to manipulate and disseminate political narratives. Here, new approaches such as Social Media Critical Discourse Analysis (SM-CDS) holds key potential. Finally, Discourse Studies can also be mobilised to critically examine how the very definition of terms such as “fake news” are mobilised politically as part of hegemonic struggles (Farkas and Schou 2018). The tradition can also contribute to researching how the constructed threat of disinformation is used rhetorically to legitimise new forms of legislation, including censorship laws in authoritarian regimes (Lim 2020; Farkas and Schou 2019). Hopefully, future research will provide new insights on these important issues.

References

- Bail, Christopher A., Brian Guay, Emily Maloney, Aidan Combs, D. Sunshine Hillygus, Friedolin Merhout, Deen Freelon, and Alexander Volfovsky. 2020. “Assessing the Russian Internet Research Agency’s Impact on the Political Attitudes and Behaviors of American Twitter Users in Late 2017.” *Proceedings of the National Academy of Sciences* 117, no. 1: 243-250.
- Bastos, Marco, and Johan Farkas. 2019. “Donald Trump Is My President!?”: The Internet Research Agency Propaganda Machine." *Social Media+ Society* 5, no. 3.
- Benkler, Yochai, Robert Faris, and Hal Roberts. 2018. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. New York: Oxford University Press.
- Block, David. 2018. *Post-truth and Political Discourse*. New York: Springer.
- boyd, danah. 2010. “Social Network Sites as Networked Publics: Affordances, Dynamics, and Implications.” In *Networked Self: Identity, Community, and*

Culture on Social Network Sites, edited by Zizi Papacharissi, 39-58. London: Routledge.

Collins, B., and Joseph Cox. 2017. "Jenna Abrams, Russia's Clown Troll Princess, Duped the Mainstream Media and the World." *The Daily Beast*, November 3, 2017. <https://www.thedailybeast.com/jenna-abrams-russias-clown-troll-princess-duped-the-mainstream-media-and-the-world>.

Dahlberg, Leif., and Sean Phelan. 2011. *Discourse Theory and Critical Media Politics*. Basingstoke: Palgrave Macmillan.

Daniels, Jessie. 2009. "Cloaked Websites: Propaganda, Cyber-racism and Epistemology in the digital era." *New Media and Society*, 11, no. 5: 659–683. <https://doi.org/10.1177/1461444809105345>

DiResta, Renee, Kris Shaffer, Becky Ruppel, David Sullivan, Robert Matney, Ryan Fox, Jonathan Albright, and Ben Johnson. 2018. *The Tactics and Tropes of the Internet Research Agency*. New Knowledge.

Ellul, Jacques. 1965. *Propaganda: The Formation of Men's Attitudes*. New York: Vintage Books.

European Commission. 2018. *A Multi-dimensional Approach to Disinformation: Report of the Independent High level Group on Fake News and Online Disinformation*. <https://doi.org/10.2759/0156>

Farkas, Johan, and Christina Neumayer. 2020. "Mimicking News." *Nordicom Review*, 41, no. 1: 1–17. <https://doi.org/10.2478/nor-2020-0001>

Farkas, Johan, and Jannick Schou. 2018. "Fake News as a Floating Signifier: Hegemony, Antagonism and the Politics of Falsehood." *Javnost - The Public*, 25, no. 3: 298-314.

Farkas, Johan, and Jannick Schou. 2019. *Post-Truth, Fake News and Democracy:*

Mapping the Politics of Falsehood. New York: Routledge.

Farkas, Johan, Jannick Schou, and Christina Neumayer. 2018. "Platformed Antagonism: Racist Discourses on Fake Muslim Facebook pages." *Critical Discourse Studies*, 15, no. 5: 463–480.
<https://doi.org/10.1080/17405904.2018.1450276>

Fairclough, Norman. 1995. *Media Discourse.* New York: Bloomsbury.

Fairclough, Norman. 2003. *Analysing Discourse: Textual Analysis for Social Research.* London: Routledge.

Fairclough, Norman, Jane Mulderrig, and Ruth Wodak. 2011. "Critical Discourse Analysis". In *Discourse Studies: A Multidisciplinary Introduction* (2nd Ed.) edited by Teun A. van Dijk, 394–417. London: SAGE Publications.

Freelon, Deen, Michael Bossetta, Chris Wells, Josephine Lukito, Yiping Xia, and Kirsten Adams. 2020. "Black Trolls Matter: Racial and Ideological Asymmetries in Social Media Disinformation." *Social Science Computer Review*, OnlineFirst.

Freelon, Deen, and Tetyana Lokot. 2020. "Russian Twitter Disinformation Campaigns Reach across the American Political Spectrum." *Harvard Kennedy School Misinformation Review*, 1, no. 1.

Freelon, Deen, and Chris Wells. 2020. "Disinformation as Political Communication." *Political Communication*, 37, no. 2: 145–156.
<https://doi.org/10.1080/10584609.2020.1723755>

Golovchenko, Yevgeniy, Mareike Hartmann, and Rebecca Adler-Nissen. 2018. "State, Media and Civil Society in the Information Warfare over Ukraine: Citizen Curators of Digital Disinformation." *International Affairs*, 94, no. 5: 975–994. <https://doi.org/10.1093/ia/iyy148>

- Grand Jury for the District of Columbia. 2018. *Indictment*. Retrieved from <https://www.justice.gov/file/1035477/download>
- Guess, Andrew M., Brendan Nyhan, and Jason Reifler. 2020. "Exposure to Untrustworthy Websites in the 2016 US Election." *Nature Human Behaviour*, 4, no. 5: 472-480.
- Habgood-Coote, Joshua. 2018. "Stop Talking about Fake News!" *Inquiry: An Interdisciplinary Journal of Philosophy*, 62, no. 9–10: 1033–1065. <https://doi.org/10.1080/0020174X.2018.1508363>
- Hall, Stuart. 1982. "The Rediscovery of 'Ideology': Return of the Repressed in Media Studies." In *Culture, Society and the Media*, edited by Michael Gurevitch, Tony Bennett, James Curran, and Janet Woollacott, 52-86. London: SAGE Publications.
- Hedrick, Ashley, David Karpf, and Daniel Kreiss. 2018. "The Earnest Internet vs. the Ambivalent Internet." *International Journal of Communication*, 12, no. 8: 1057-1064.
- Hervik, Peter. 2011. *The Annoying Difference: The Emergence of Danish Neoliberalism, Neoracism, and Populism in the post-1989-world*. New York: Berghahn Books.
- Hogan, Bernie. 2010. "The Presentation of Self in the Age of Social Media: Distinguishing Performances and Exhibitions Online." *Bulletin of Science, Technology and Society*, 30, no. 6: 377–386. <https://doi.org/10.1177/0270467610385893>
- Jack, Caroline. 2017. *Lexicon of Lies: Terms for Problematic Information*. Retrieved from https://datasociety.net/pubs/oh/DataAndSociety_Lexicon-ofLies.pdf

- Jorgensen, Marianne, and Louise J. Phillips. 2002. *Discourse Analysis as Theory and Method*. London: SAGE Publications.
- KhosraviNik, Majid. 2017. "Social Media Critical Discourse Studies (SM-CDS)." In *The Routledge Handbook of Critical Discourse Analysis*, edited by John Flowerdew and John E. Richardson, 582–596. London: Routledge.
- KhosraviNik, Majid. 2018. "Social Media Techno-Discursive Design, Affective Communication and Contemporary Politics." *Fudan Journal of the Humanities and Social Sciences*, 11: 427–442. <https://doi.org/10.1007/s40647-018-0226-y>
- KhosraviNik, Majid. 2020. "Digital Meaning-making across Content and Practice in Social Media Critical Discourse Studies." *Critical Discourse Studies*, online first. <https://doi.org/10.1080/17405904.2020.1835683>
- Kreiss, Daniel. 2019a. "The Fragmenting of the Civil Sphere: How Partisan Identity Shapes the Moral Evaluation of Candidates and Epistemology." In *Politics of Meaning/Meaning of Politics*, edited by Jason L. Mast and Jeffrey C. Alexander, 223-241. New York: Palgrave Macmillan.
- Kreiss, Daniel. 2019b. "Four Provocations for Info Politics Research." *Medium*. <https://medium.com/@dkreiss/four-provocations-for-info-politics-research-5026bc809c29>
- Krzyżanowski, Michał, and Bernhard Forchtner. 2016. "Theories and Concepts in Critical Discourse Studies: Facing Challenges, Moving beyond Foundations." *Discourse and Society*, 27, no. 3: 253–261. <https://doi.org/10.1177/0957926516630900>
- Krzyżanowski, Michał, and Per Ledin. 2017. "Uncivility on the Web Populism in/and the Borderline Discourses of Exclusion." *Journal of Language and Politics*, 16, no. 4: 566–581. <https://doi.org/10.1075/jlp.17028.krz>

- Laclau, Ernesto. 2000. "Foreword." In *Discourse Theory and Political Analysis: Identities, Hegemonies and Social Change*, edited by David Howarth, Aletta J. Norval, and Yannis Stavrakakis. Manchester: Manchester University Press.
- Laclau, Ernesto. 2014. *The Rhetorical Foundations of Society*. London; New York: Verso Books.
- Laclau, Ernesto, and Chantal Mouffe. 2014. *Hegemony and Socialist Strategy: Towards a Radical Democratic Politics*. London: Verso.
- Langlois, Ganaele, and Greg Elmer. 2013. "The Research Politics of Social Media Platforms." *Culture Machine*, 14, 1–17.
- Langlois, Ganaele, Greg Elmer, Fenwick Mckelvey, and Zachary Devereaux. 2009. "Networked Publics: The Double Articulation of Code and Politics on Facebook." *Canadian Journal of Communication*, 34, no. 3: 415–434.
- Lewandowsky, Stephan, Ullrich K. H. Ecker, and John Cook. 2017. "Beyond Misinformation: Understanding and Coping with the "Post-Truth" Era." *Journal of Applied Research in Memory and Cognition*, 6, no. 4: 353–369. <https://doi.org/10.1016/j.jarmac.2017.07.008>
- Lim, Gabrielle. 2020. "Securitize / Counter - Securitize: The Life and Death of Malaysia's Anti-Fake News Act." *Data and Society*. <https://datasociety.net/library/securitize-counter-securitize>
- Linville, Darren L., and Patrick L. Warren. 2019. "Troll Factories: Manufacturing Specialized Disinformation on Twitter." *Political Communication*, 37, no. 4: 447-467. <https://doi.org/10.1080/10584609.2020.1718257>
- Lukito, Josephine. 2019. "Coordinating a Multi-Platform Disinformation Campaign: Internet Research Agency Activity on Three US Social Media Platforms, 2015 to 2017." *Political Communication*, 37, no. 2: 238-255.

- Marwick, Alice E. 2018. "Why Do People Share Fake News? A Sociotechnical Model of Media Effects." *Georgetown Law Technology Review*, 2, no. 2: 474-512.
- Marwick, A. E., & boyd, d. (2011). I tweet honestly, I tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, 13(1), 114-133.
- Marwick, Alice, and Rebecca Lewis. 2017. *Media Manipulation and Disinformation Online*. Retrieved from <https://datasociety.net/output/media-manipulation-and-disinfo-online/>
- Matamoros-Fernández, Ariadna. 2017. Platformed racism: the mediation and circulation of an Australian race-based controversy on Twitter, Facebook and YouTube, *Information, Communication & Society*, 20(6), 930-946.
- O'Sullivan, Donie. 2017. "A Notorious Russian Twitter Troll Came Back, and for a Week Twitter Did Nothing." *CNN Business*, November 19, 2017. Retrieved from <https://money.cnn.com/2017/11/17/media/new-jenna-abrams-account-twitter-russia/index.html>
- Palma, Bethania. 2018. "Zuckerberg: It's 'Entirely Possible' Cambridge Analytica Activities Overlapped With Indicted Russian Firm." *Snopes*, April 10, 2018. Retrieved from <https://www.snopes.com/news/2018/04/10/cambridge-analytica-indicted-russian-firm/>
- Phillips, Whitney. 2020. "You're Fake News: The Problems with Accusations of Falsehood." In *Fake News: Understanding Media and Misinformation in the Digital Age*, edited by Melissa Zimdars and Kembrew McLeod, 55–64. Cambridge: MIT Press.
- Phillips, Whitney, and Ryan M. Milner. 2018. *The Ambivalent Internet: Mischief, Oddity, and Antagonism Online*. New York: John Wiley and Sons.

- Potter, Jonathan. 1996. *Representing Reality. Discourse, Rhetoric and Social Construction*. London: SAGE Publications.
- Rini, Regina. 2017. Fake News and Partisan Epistemology. *Kennedy Institute of Ethics Journal*, 27, no. 2: E-43-E-64. <https://doi.org/10.1353/ken.2017.0025>
- Tandoc, Edson C., Zheng Wei Lim, and Richard Ling. 2018. “Defining ‘Fake News’: A Typology of Scholarly Definitions.” *Digital Journalism*, 6, no. 2: 137–153. <https://doi.org/10.1080/21670811.2017.1360143>
- Torfining, Jacob. 2005. “Discourse Theory: Achievements, Arguments, and Challenges.” In *Discourse Theory in European Politics: Identity, Policy and Governance*, edited by David Howarth and Jacob Torfining, 1–32. Basingstoke: Palgrave Macmillan.
- Tripodi, Francesca. 2018. *Searching for Alternative Facts: Analyzing Scriptural Inference in Conservative News Practices*. New York: Data and Society Research Institute.
- Unger, Johann, Ruth Wodak, and Majid KhosraviNik. 2016. “Critical Discourse Studies and Social Media Data.” In *Qualitative Research (4th edition)*, edited by David Silverman, 277–293. London: SAGE Publications.
- van Dijck, Jose. 2013. *The Culture of Connectivity: A Critical History of Social Media*. Oxford: Oxford University Press.
- van Dijk, Teun. A. 1985. *Discourse and Communication: New Approaches to the Analysis of Mass Media Discourse and Communication*. Berlin: Walter De Gruyter.
- Wodak, Ruth. 2009. “Discourse Analysis: History, Agenda, Theory, and Methodology.” In *Methods for Critical Discourse Analysis*, edited by Ruth Wodak and Michael Meyer, 1–33. London: SAGE Publications.

Wodak, Ruth. 2015. *The Politics of Fear: What Right-Wing Populist Discourses Mean*. London: SAGE Publications.

Wodak, Ruth, and Michael Meyer. 2016. *Methods of Critical Discourse Studies*. London: SAGE Publications.

Wylie, Christopher. 2019. "Whatever this report says, here's what I know: when I was at Cambridge Analytica, the company hired known Russian agents, had data researchers in St Petersburg, tested US voter opinion on Putin's leadership, and hired hackers from Russia - all while Banno." *Twitter*. Retrieved from <https://twitter.com/chrisinsilico/status/1109983298397310976?lang=en>

Xia, Yiping, Josephine Lukito, Yini Zhang, Chris Wells, Sang J. Kim, and Chau Tong. 2019. "Disinformation, Performed: Self-presentation of a Russian IRA Account on Twitter." *Information, Communication and Society*, 22, no. 11: 1646-1664.