

# What Does Fake Look Like? A Review of the Literature on Intentional Deception in the News and on Social Media

Alyt Damstra, Hajo G. Boomgaarden, Elena Broda, Elina Lindgren, Jesper Strömbäck, Yariv Tsfati & Rens Vliegenthart

To cite this article: Alyt Damstra, Hajo G. Boomgaarden, Elena Broda, Elina Lindgren, Jesper Strömbäck, Yariv Tsfati & Rens Vliegenthart (2021) What Does Fake Look Like? A Review of the Literature on Intentional Deception in the News and on Social Media, *Journalism Studies*, 22:14, 1947-1963, DOI: [10.1080/1461670X.2021.1979423](https://doi.org/10.1080/1461670X.2021.1979423)

To link to this article: <https://doi.org/10.1080/1461670X.2021.1979423>



© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 29 Sep 2021.



Submit your article to this journal [↗](#)



Article views: 5943



View related articles [↗](#)



View Crossmark data [↗](#)

## What Does Fake Look Like? A Review of the Literature on Intentional Deception in the News and on Social Media

Alyt Damstra <sup>a</sup>, Hajo G. Boomgaarden <sup>b</sup>, Elena Broda<sup>c</sup>, Elina Lindgren <sup>d</sup>,  
Jesper Strömbäck <sup>c</sup>, Yariv Tsfaty <sup>e</sup> and Rens Vliegenthart <sup>a</sup>

<sup>a</sup>Amsterdam School of Communication Research, University of Amsterdam, Amsterdam, Netherlands;

<sup>b</sup>Department of Communication, Universität Wien Institut für Publizistik und Kommunikationswissenschaft, Wien Austria; <sup>c</sup>University of Gothenburg, Göteborg, Sweden; <sup>d</sup>Department of Journalism, Media, and Communication, University of Gothenburg, Gothenburg, Sweden; <sup>e</sup>Department of Communication, University of Haifa, Haifa, Israel

### ABSTRACT

This paper focuses on the content features of intentional deceptive information in the news (i.e., fake news) and on social media. Based on an extensive review of relevant literature (i.e., political journalism and communication, computational linguistics), we take stock of existing knowledge and present an overview of the structural characteristics that are indicative of intentionally deceptive information. We discuss the strength of underlying empirical evidence and identify underdeveloped areas of research. With this paper, we aim to contribute to the systematic study of intentional deception in the news and on social media and to help setting up new lines of research in which intentionally deceptive news items can be operationalized in consistent ways.

### KEYWORDS

Disinformation; fake news; journalism; social media; linguistics; content features

The recent outbreak of a novel coronavirus disease (COVID-19) vividly demonstrated how the deliberate spread of disinformation may have very real consequences for people's beliefs and behavior. In the early stages of the pandemic, for example, rumors swirled that the rise and spread of COVID-19 could be linked to 5G cellular networks (Ahmed et al. 2020). The panic on social media was said to travel faster than the virus itself (Depoux et al. 2020) and within weeks, dozens of 5G masts were attacked throughout the United Kingdom and elsewhere.<sup>1</sup> The fabrication and spread of intentionally deceptive information is not confined to the case of COVID-19, nor is it typical for British culture. The phenomenon is on the rise across the globe, as illustrated by a Freedom House report finding that 30% more countries were subject to online fake activity in 2017 than 2016.<sup>2</sup> The rise of deliberate disinformation—nowadays often referred to as fake news—has sparked societal as well as academic debates. However, despite this heightened attention, important aspects of the phenomenon are still only marginally understood.

In the field of political communication and journalism, research on deceptive news can be roughly divided into three sub areas. First, there is conceptual work focusing on

theoretical clarification, offering fake news typologies and connecting the recent upsurge to the context of changing media landscapes (e.g., Egelhofer and Lecheler 2019; S  e 2019; Tandoc, Lim, and Ling 2018; Waisbord 2018). Second, a strand of empirical work studies the dynamics, spread, and effects of fake news (e.g., Grindberg et al. 2019; Guess, Nyhan, and Reifler 2018; Pierri et al. 2020; Silverman 2016a), identifying the scope of exposure as well as the most relevant disseminators, and, as far as possible, assessing its impact on citizens' attitudes and beliefs (see also Tfsati et al. 2020). Finally, there is work focusing on ways in which the spread and impact of disinformation could be tackled (e.g., Hamel-eers 2020; Lazer et al. 2017; Walter et al. 2020). Possible remedies that have been suggested are increasing levels of media literacy on the audience level and debunking (i.e., fact-checking) by journalists (Hameleers and Van der Meer 2019; Nyhan et al. 2020). Social media platforms and advertising networks have also been urged to take measures to reduce the prevalence of fake news on their systems as they violate policies against misleading content (Wingfield, Isaac, and Benner 2016).

These works have offered important insights into the nature, spread and scope of fake news, and the varying degrees to which counter measures such as fact-checking activities are successful. However, they offer less guidance when it comes to identifying the more general characteristics of fake news. In fact, we know only little about structural features that may facilitate the identification of disinformation. It may well be that intentionally deceptive information differs from regular news reports in only one respect: The intentions of the source. In that case, identifying intentionally deceptive information based on content characteristics alone is virtually impossible. However, it is reasonable to assume that the content of deceptive information is structurally different compared to fully correct information. For example, work in the field of linguistics suggests that deceptive language contains more tentative, angry, and emotional words than truthful language does (Asubiaro and Rubin 2018).

With this paper, we aim to fill this void by mapping existing knowledge about the structural features of intentional deception in texts. To delimit the scope, we focus specifically on messages that are created and distributed with the aim of influencing political attitudes, behavior, or processes. Two types of communication are distinguished: Intentional deception masqueraded as an entire news article (which often is referred to as fake news), and intentional deception in texts on social media (i.e., Twitter, Facebook). We adopt a systematic approach, reviewing different literatures to identify features that are common in both types of communication. To enhance a structured discussion of the results, we arrange the features by distinguishing between content features on the one hand and linguistic features on the other. The former refers to structural characteristics that deal with the substantive message that is communicated, such as ideological biases. The latter relates to structural characteristics of the language that is used, for example, the use of pronouns or the presence of swear. In sum, our overview offers a set of empirically validated criteria indicating the likelihood of intentional deception in text. Based on our findings, we discuss ambiguities and underdeveloped areas in this research field and suggest possible avenues for future research.

## Intentionally Deceptive Information

The recent upsurge of fake news may elicit the view that we deal with a modern phenomenon. That is however not the case. The complicated relationship between news, politics,

and the truth did not start with the inauguration of President Trump, nor with the advent of the Internet (Schudson 2019). In fact, intentionally deceptive information in the news is all but new; it has been around since the establishment of the press, although not equally present in all periods since (Ortoleva 2019). What makes it such a pressing concern today is that the digital media environment enhances the enormous and rapid spread of information. For example, we know from several studies that Facebook was a key disseminator of fake news stories during the 2016 Presidential election campaign (e.g., Guess, Nyhan, and Reifler 2018). Against this backdrop, some have argued that we have entered “a post-truth era” (Benkler, Faris, and Roberts 2018, 23) which has provoked renewed academic interest in deceptive information—nowadays often referred to as fake news.

Despite this heightened interest, the concept remains contested which is illustrated by the numerous operationalizations and typologies of fake news (e.g., Tandoc, Lim, and Ling 2018). Also from an audience perspective, no unanimity exists as people hold rather diverging views of what the phenomenon entails (Nielsen and Graves 2017). With this paper, we hope to contribute to research in this area by identifying the content features that are indicative of intentionally deceptive information in news-like texts and on social media. Given the conceptual ambiguity and the frantic pace at which this research field has developed over the past few years, it is key to clearly limit this study's scope and to adopt a systematic approach when reviewing the relevant literature. We do so in the following ways.

First, to demarcate the scope of our review, two dimensions are of particular interest: The specific format in which deceptive information is presented and the intentions underlying the creation and spread of deceptive information. Regarding the former, we are interested in the manifestation of intentional deception in written news-like reports and on social media. Studies that focus on intentional deception in other sorts of texts, ranging from emails (Zhou et al. 2004) to financial reports (Humpherys et al. 2011), are not included in the current study. Related to deceptive information in the news, we include studies that focus on deceptive information that masquerades as an entire news article. Regular or truthful news reports that contain one or more false claims thus fall outside this study's scope. Related to the format of social media messages, we include studies that focus on textual communication on social media platforms. Thus, visual messages or hybrid communication in which text is subordinate to the image—such as memes—are not included, neither are videos and so-called “deepfakes”. Regarding intentions, Allcott and Gentzkow (2017) identify two motivations underlying the production of intentional deception in texts: financial and ideological. We are mainly interested in the latter, thus in content that is created and distributed with the goal of promoting particular ideas or people, often by discrediting others. In news-like texts, this is often achieved by taking on some form of journalistic credibility by trying to mimic the look and feel of real news (Tandoc, Lim, and Ling 2018). This implies that satire falls outside this study's scope, as does commercial (native) advertising or clickbait articles that are created out of commercial considerations alone. For social media, we include studies that focus on intentional deceptive information that is mainly produced in order to influence public opinion—and not only to entertain, for example.

It is important to note, however, that the intentions of news sources are typically not known. This is a major challenge to research in this area, as communication scholars generally distinguish disinformation from misinformation and regular news by referring to the intentions of the source—without measuring them. The most common way to work

around this issue is to rely on datasets with validated examples of news that have been proven to be false, containing information with very low levels of facticity (Tandoc, Lim, and Ling 2018). Examples of such datasets are fact-checking websites such as *PolitiFact* or the *Washington Post Fact Checker* (for an overview, see Omezi and Jahankhani 2020). Theoretically, there is a possibility that these news stories identified as false are examples of misinformation (not being meant to deceive) instead of disinformation. The same applies to tweets in which the content is not compatible with reality (low facticity). All this is to say that in some cases, we are forced to make assumptions about the underlying intentions of the source, without being able to substantiate these assumptions with empirical assessments.

Second, we have collected studies in a conscious and systematic way. Key to the validity of our review is that the findings are representative of a fast-growing research area that covers diverse literatures. More specifically, we identified relevant research in two ways. First, we restricted ourselves to scholarly articles. To identify relevant studies, we used Google Scholar to search for articles published in peer-reviewed journals between 2010 and 2021. This time frame was selected because it covers the vast majority (>95%) of all scholarly work dealing with intentional deception in news-like texts and on social media. To select relevant articles, we used the search terms *misinformation*, *disinformation*, and *fake news*. Relevant studies had to have at least one of these search terms in their title, abstract, or as keyword. Book reviews, editorials and other types of content that usually are not peer-reviewed were filtered out, as were articles that only mentioned the search terms in passing or that were published in non-reputable journals.<sup>3</sup> Based on this corpus, we close-read the selected articles and identified the ones—regardless of discipline—that dealt with third-person detection of intentional deception in news-like texts and on social media. Papers dealing with intentional deception in other sorts of content were excluded from the analysis. In addition, we followed up the references in the articles, to make sure we had not missed any relevant studies. If we came across papers that dated from earlier periods (before 2010), we read them and included them if they focused on intentional deceptive information in news-like texts or on social media. This process revealed that most relevant research has been published in the fields of journalism, political communication, and (computational) linguistics.

Altogether, we identified more than 30 relevant studies. Having examined these, we list the characteristics that are found to be indicative of intentional deception in text. We subdivide these features into two broad categories: *content features* and *linguistic features*. These categories are not fully mutually exclusive, and they mainly serve to enhance a more structured discussion of the results. We start off with the content features that have been identified across literatures. For each feature, we discuss some of the most relevant studies, instead of providing an exhaustive overview.

## Content Features

In the examined literature, we observe four structural characteristics that relate to the substantive message that is communicated by intentional deceptive texts. These messages often have an ideological bias in favor of the right, are designed to provoke negative emotions such as anger and fear, contain little verifiable information, and make use of fully packed and sensationalist headlines. We discuss each feature in more detail below.

## *Ideological Bias*

Research suggests that intentionally deceptive information tends to be ideologically biased in favor of the right. The U.S. 2016 presidential election has been a catalyst for academic research on disinformation—quite some influential work focuses on the U.S. context in the years preceding and following this election. For example, Faris et al. (2017) analyze mainstream and social media election coverage. Their data cover over two million stories, published by approximately 70,000 online media sources between May 1, 2015 and Election Day (November 8, 2016). While partisan bias exists on both sides of the political spectrum (Bradshaw et al. 2020), the content receives more amplification and legitimation on the right side of it, especially on Facebook (Faris et al. 2017). In fact, intentionally deceptive information in news-like texts was found to often have a pro-Trump signature (Silverman 2016b), and stories favoring Trump were shared more widely than the ones favoring Clinton (Allcott and Gentzkow 2017; Lazer et al. 2017). In a similar vein, Benkler, Faris, and Roberts (2018) collected and analyzed two million stories published during the 2016 presidential election campaign and 1.9 million stories about the Trump presidency during its first year. In both studies, the spread and reach of disinformation is examined by tracing cross-linking patterns between media sources, including sharing activities on Twitter and Facebook. The insulated right-wing media ecosystem is found to be much more susceptible to disinformation than the other side of the spectrum, spanning from center-right to far-left publications. Similarly, Marwick and Lewis (2017) study media manipulation in the U.S. context by far-right groups in the run-up to the 2016 election. They conclude that most Clinton supporters got news from mainstream sources, while many Trump supporters were surrounded by a far-right network which “peddled heavily in misinformation, rumors, conspiracy theories, and attacks on the mainstream media” (Marwick and Lewis 2017, 21).

Further research is however needed, especially between-country and over-time analyses, to examine whether this partisan asymmetry is a persistent phenomenon, or whether it is a transient artifact elicited by the Trump administration (Freelon and Wells 2020, 152). There is some work examining partisan content of disinformation outside the U.S. context. For example, Pierri et al. (2020) examine Italian disinformation on Twitter and find that most topics concern polarizing arguments related to immigration, crime, and national safety—issues that mirror the agenda of the conservative and far-right political community. Similarly, German disinformation is found to contain mostly right-wing implications “such as skepticism toward the European Union (...) and above all the exclusion of migrants and refugees” (Zimmermann and Kohring 2020, 221).

## *Use and Presence of Emotions*

Research from different fields indicates that the use and presence of emotions may be structurally different in intentionally deceptive information compared to regular news reports. While the latter also are pervaded by subjective language in the form of journalistic appraisals (see e.g., Wahl-Jorgensen 2013, 2019, 2020), the use of emotions in intentional deceptive news items is considerably different: emotions are more visible, more prominent, and more negative.

There are quite some studies that provide empirical support for this observation. Bradshaw et al. (2020) for example, have developed a grounded typology distinguishing different types of news and information that was spread and shared on Twitter during the 2016 U.S. Presidential election and the State of the Union address in January 2018. Based on 21.8 million tweets, they analyze the domains (URLs) that were most often shared. Most relevant to the purpose of the current paper is what they call *junk news*—sources that deliberately publish or aggregate misleading, deceptive, or incorrect information packaged as real news about politics, economics, or culture (Bradshaw et al. 2020, 188). A key feature of this type of deceptive information relates to emotionally driven language; propaganda techniques are used to persuade users at an emotional rather than cognitive level. A similar approach is adopted by Neudert, Kollanyi, and Howard (2017), who examine the spread and reach of intentionally deceptive information during campaign periods in Germany and, later, in Germany, France, and the UK (Neudert, Howard, and Kollanyi 2019). They also find that emotionally charged words are key to these types of messages. Scholars analyzing intentionally deceptive information on *Breitbart's* Facebook timeline also conclude that the content is affective, purposively designed to provoke voter outrage and to elicit fear and disgust (Benkler, Faris, and Roberts 2018; Faris et al. 2017). It is thus important to note that emotive language mainly refers to *negative* emotions, as deceptive news or tweets aim to elicit negative affect among those who consume it (see also Hameleers, van der Meer, and Vliegenthart 2021).

Emotive language and words are also identified as a typical feature in a whole different strand of studies. Deception detection is a thriving research field in computer science and linguistics. To develop methods of automated detection, scholars rely on benchmark labeled datasets (e.g., *PolitiFact*) that contain truthful and deceptive news content that has been fact-checked for its veracity. Based on such datasets, machine learning models are developed to automatically detect disinformation based on linguistic features, with varying degrees of success (e.g., Asubiaro and Rubin 2018; Wang 2017). Horne and Adali (2017), for example, use three separate datasets to study the linguistic features of real news, fake news, and satire. They conclude that intentionally deceptive news typically contains more negative emotive words than truthful news (see also Zhou and Zafarani 2020). Also fabricated headlines score high on emotiveness (Asubiaro and Rubin 2018, see for similar findings Volkova et al. 2017). In relation to social media, Van Der Zee et al. (2018) conduct a linguistic analysis on 447 tweets written by former U.S. President Trump that were fact-checked by *The Washington Post*. That study suggests that truthful tweets contain more positive sentiments, while deceptive tweets have more negative emotions. Research examining Russian disinformation campaigns on social media also find that these messages aim to provoke outrage and anger against oppositional outgroups. A common strategy to achieve this effect is by vilifying political and social adversaries (Freelon and Lokot 2020; Howard et al. 2019).

### Verifiability

Maybe not surprisingly, intentional deceptive information tends to be less verifiable. Journalism scholars have examined examples of journalistic fakery in order to identify salient features. For example, Lasorsa and Dai (2007) compare deceptive news stories with authentic ones and conclude that fake news stories often deal with issues that are conducive to



secrecy, which may conceal the absence of verifiable information. The latter also relates to the use of sources, by which we mean individuals who provide information or quotes for news articles. Intentionally deceptive stories contain more references to sources; however, those sources are rarely presented in a way that makes them traceable. Identifiable names are not provided, as sources are typically anonymous and vague (Bonet-Jover et al. 2021). Others refer to the use of conspiratorial and dubious sources, without fact-checking (Bradshaw et al. 2020; Marchal et al. 2019; Neudert, Howard, and Kollanyi 2019). The use of non-verifiable sources is also observable in linguistic features, such as the greater presence of pronouns instead of specific source names (see below). Furthermore, sources are less often used for direct quotes, as quoted content is found to be lower in fake news compared to legitimate news (Reddy et al. 2020). A lack of verifiable information resonates with work suggesting that source cues are less and less important for how people find and process information about public affairs, which may ultimately enhance the consumption of deceptive messages (e.g., Kalogeropoulos, Fletcher, and Nielsen 2019).

### ***Use of Headlines***

Headlines are a strong differentiating factor between fake and real news (Horne and Adali 2017). In general, the headlines of intentionally deceptive news items tend to be eye-catching, with a propensity for exaggeration, sensationalism, and scaremongering (Chen, Conroy, and Rubin 2015; Potthast et al. 2016; Sahoo and Gupta 2021; Shu et al. 2017; Tucker et al. 2018). In addition, these headlines tend to be longer (Asubiaro and Rubin 2018; Liu et al. 2019), using more capitalized words, proper nouns and verb phrases. By doing so, titles of fabricated news items try to get many points across, while titles of truthful texts most often opt for brief and general summary statements (Horne and Adali 2017, 764). For example, Liu et al. (2019) show in their study on health-related information on Chinese social media that the long headlines of fake news articles often displayed patterns of “click-baiting” measured through the use of imperative idioms such as “(you) must” or “never (do this)”.

### ***Linguistic Features***

In addition, based on the examined literature, a number of features can be identified that relate to the specific language that is used in intentional deceptive news-like texts or social media posts. Summing up, intentionally deceptive texts tend to be characterized by a more frequent use of capitalization, pronouns and informal language or swear. In addition, there are three linguistic features that are often examined but for which no unambiguous evidence is found. For levels of lexical diversity, text length, and the use punctuation the literature offers diverging clues.

### ***Lexical Diversity***

Higher degrees of lexical diversity imply a higher number of unique words used in a text, when this number is adjusted to be independent of the length of the text, it is referred to as the Type-Token Ratio. Lower degrees of lexical diversity in intentionally deceptive communication have been confirmed in a variety of settings (e.g., Fuller, Biros, and Delen



2011; Newman et al. 2003; Zhou et al. 2004). However, applied to news articles, results are mixed. Horne and Adali (2017), comparing the linguistic features of real and fake news content, find that intentionally deceptive news items are characterized by lower levels of lexical diversity than real news articles: “Fake articles need a slightly lower education level to read (...) They seem to be filled with less substantial information, which is demonstrated by a high amount of redundancy, more adverbs, fewer nouns, fewer analytic words, and fewer quotes.” (Horne and Adali 2017 : 763). This observation is supported by Ahmed, Traore, and Saad (2018 ) who find higher usage of verbs and adverbs for fake news, whereas real news is characterized by a higher proportion of nouns and adjectives. In contrast, Abonizio et al. (2020), comparing fake and legitimate news across three corpora (in English, Portuguese, and Spanish), find higher Type-Token Ratio values for the fake news corpora. Similarly, there is some evidence pointing towards higher lexical diversity in deceptive news, specifically in terms of the number of unique verbs used (Zhou and Zafarani 2019, 2020).

### **Capitalization**

There is some evidence that intentionally deceptive news-like texts tend to include excessive capitalization, in the headline as well as in the body of an article, in order to attract attention (e.g., Bradshaw et al. 2020; Marchal et al. 2019; Neudert, Kollanyi, and Howard 2017; Reddy et al. 2020). This may also apply to tweets, where the excessive use of upper-case text (> 70 percent) has been found to be an indicator of false content (Srivastava, Rehm, and Schneider 2017).

### **Use of Pronouns**

The use of pronouns might also be structurally different. Applied to news content, Rashkin et al. (2017) find that first-person and second-person pronouns are more often used in deceptive news texts. The authors explain the difference by referring to journalism practices. Editors of trustworthy sources are possibly quite rigorous about removing language that seems too personal, while such processes do not play a role in the production of fabricated news items (see also the findings about source use). In a similar vein, Asubiaro and Rubin (2018) find that the use of (all types of) pronouns is typically higher in deceptive news. Applied to social media, the aforementioned study by Van Der Zee et al. (2018) finds similar patterns (for third-person pronouns) for tweets by former President Trump that were established to be deceptive by *The Washington Post*. Gupta et al. (2014), who also study communication on Twitter, develop a model that automatically assesses the credibility of tweets during crises. Different from the study by Van der Zee *et al*, they base themselves on a large N dataset, covering tweets from millions of users. A higher use of pronouns is one of the criteria on the basis of which tweets score lower on credibility.

### **Length**

Turning to the length of texts and the difference between deceptive and truthful communication, research has provided some mixed results. This might be due to the type

of text that is considered. Concerning fake news reports, these tend to be shorter than truthful news stories: the number of words as well as the number of paragraphs is significantly lower (Asubiaro and Rubin 2018; Horne and Adali 2017). In tweets, however, the opposite pattern seems to apply as tweets by former President Trump are found to be longer when they contain deceptive information (Van Der Zee et al. 2018).

### ***Informal Language and Swear***

In regular news reports, informal words and language are generally kept to a minimum. However, in intentionally deceptive news items, slang and swear (“inflammatory language”) are found to be rather frequent (e.g., Asubiaro and Rubin 2018; Gupta et al. 2014; Neudert, Howard, and Kollanyi 2019; Rashkin et al. 2017; Zhou and Zafarani 2020). In a recent paper, Hameleers, van der Meer, and Vliegenthart (2021) conduct an extensive content analysis of deceptive statements that were fact-checked by *Politifact.org* and *Snopes.com*. They find that completely false information is most likely to contain hate speech and incivility. As a result, they conclude that hate speech and incivility may be considered indicators of disinformation.

### ***Punctuation***

Finally, several studies examine the use of punctuation as a possible feature of intentionally deceptive information in news-like texts or on social media. Punctuation includes periods, comma, colon, semi-colon, question marks, exclamation marks, and quotes. Results are mixed. There is some evidence that fake news articles use fewer punctuation (e.g., Horne and Adali 2017), but others find more use of punctuation, at least in the headlines (Asubiaro and Rubin 2018; Liu et al. 2019). Further, there may be differences dependent on the specific punctuation signs that are used. For instance, Liu et al. (2019) conclude in their study of health-related information on Chinese social media that fake news articles use more exclamation marks in their headlines, whereas real news articles utilize more question marks.

Summing up, this review shows that there are a number of characteristics that are potentially indicative of intentionally deceptive content. Table 1 summarizes the features, together with references to relevant studies.

### **Evaluation of Features—Solidity Versus Ambiguity**

The list of features that follows from the above review—and that is summarized in Table 1—varies in terms of the strength and unambiguity of the underlying empirical evidence. For some of the features there is considerable empirical support, such as the ideological bias and the use of emotions. Lower levels of verifiability, the use of eye-catching headlines and informal words and language are also recurring observations across disciplines. It must be noted, however, that most of the work focuses on intentionally deceptive information *in the news*, and less can be said about the applicability of these findings to the content that is produced on social media. This may be related to higher individual variation in writing style on social media (Faustini and Covões 2020). For some of the features,

**Table 1.** Overview of identified features indicative of intentionally deceptive texts in the news and on social media.

Content features		
<i>Ideological bias</i>	(Hyper-)partisan bias, often with a right-leaning ideological orientation. Negative references to left-leaning, progressive political actors or issues, positive references to (populist) right leaning political actors or issues.	Relevant studies: Allcott and Gentzkow (2017); Benkler, Faris, and Roberts (2018); Bradshaw et al. (2020); Faris et al. (2017); Lazer et al. (2017); Marchal et al. (2019); Marwick and Lewis 2017; Neudert, Kollanyi, and Howard (2017, 2019); Pierri et al. (2020); Silverman (2016b); Zimmermann and Kohring (2020)
<i>Use and presence of emotions</i>	Greater use of emotive and affective language. Especially negative emotions typically more present. In news, content contains heavy emotional appeal to readers, provoking fear, anger, outrage.	Relevant studies: Asubiaro and Rubin (2018); Bakir and McStay (2018); Benkler, Faris, and Roberts (2018); Bradshaw et al. (2020); Chu, Xie, and Wang (2021); Faris et al. (2017); Faustini and Covões (2020); Freelon and Lokot 2020; Hameleers, van der Meer, and Vliegenthart (2021); Horne and Adali (2017); Howard et al. 2019; Marchal et al. (2019); Neudert, Kollanyi, and Howard 2017, 2019; Van Der Zee et al. (2018); Volkova et al. (2017); Zhou and Zafarani (2020)
<i>Verifiability</i>	Less verifiable facts. More use of anonymous sources. Stories focus on topics that are conducive to secrecy.	Relevant studies: Asubiaro and Rubin (2018); Bonet-Jover et al. (2021); Bradshaw et al. (2020); Lasorsa and Dai (2007); Marchal et al. (2019); Neudert, Howard, and Kollanyi (2019); Reddy et al. (2020)
<i>Headlines</i>	Fake news headlines tend to be relatively long and eye-catching, with a propensity for exaggeration and scaremongering.	Relevant studies: Asubiaro and Rubin (2018); Chen, Conroy, and Rubin (2015); Horne and Adali (2017); Liu et al. (2019); Potthast et al. (2016); Sahoo and Gupta 2021; Shu et al. (2017); Tucker et al. (2018)
<b>Linguistic features</b>		
<i>Lexical diversity /type-token ratio</i>	Information is less (or more) elaborated or complex in terms of word diversity (such as the number of unique words used). Results are mixed.	Relevant studies: Abonizio et al. (2020); Ahmed, Traore, and Saad (2018); Faustini and Covões (2020); Horne and Adali (2017); Liu et al. (2019); Zhou and Zafarani (2019, 2020)
<i>Capitalization</i>	Excessive use of capitalization in order to attract attention.	Relevant studies: Bradshaw et al. (2020); Marchal et al. (2019); Neudert, Kollanyi, and Howard (2017); Reddy et al. (2020); Srivastava, Rehm, and Schneider (2017)
<i>Use of pronouns</i>	High use of personal pronouns.	Relevant studies: Asubiaro and Rubin (2018); Gupta et al. (2014); Rashkin et al. (2017); Van Der Zee et al. (2018)
<i>Length</i>	Articles tend to be short (body of text). Tweets tend to be longer.	Relevant studies: Asubiaro and Rubin (2018); Horne and Adali (2017); Van Der Zee et al. (2018)
<i>Informal words and language</i>	More use of informal words and informal language (slang, swear). Higher likelihood of hate speech and incivility.	Relevant studies: Asubiaro and Rubin (2018); Gupta et al. (2014); Hameleers, van der Meer, and Vliegenthart (2021); Neudert, Howard, and Kollanyi (2019); Rashkin et al. (2017); Zhou and Zafarani (2020)
<i>Punctuation</i>	Results are mixed: less use of punctuation in the body of text, more use of punctuation in the headline.	Relevant studies: Horne and Adali (2017); Asubiaro and Rubin (2018) ; Liu et al. (2019)

research provides mixed results. Our review indicates that different automated detection techniques may lead to different observations, for example in the case of punctuation.

In addition, our systematic review also facilitates the identification of underdeveloped research areas. First of all, comparative research is scarce. A notable exception is provided by Humprecht (2019), who compares (fact-checked) online disinformation across English-speaking (US, UK) and German-speaking countries (Austria, Germany). She observes how

the subject of fake news strongly mirrors national news agendas: topics such as health care (US) and macroeconomic issues (UK) prevailed in the English-speaking countries, whereas fake news produced in German-speaking countries often targeted and vilified immigrants, as the refugee situation was on top of these national agendas. Such an approach is valuable as it sheds light on cross-national differences and similarities. The study indicates that—despite considerable cross-national variation—intentionally deceptive information thrives best on those topics that are most polarized in a given context. More comparative research is needed in this area. Relatedly, the bulk of research focusing on the automated detection of deceptive texts is based on the English language (but see e.g., Abonizio et al. 2020; Chu, Xie, and Wang 2021; Faustini and Covões 2020). Work in which these or similar techniques are developed based on other world languages such as Mandarin and Spanish would be extremely useful as there is no reason to expect that intentionally deceptive information stops at the borders of the Anglophone world. Finally, research that looks into deceptive information on social media is strongly shaped by considerations of data availability. As a result, most of what we know is based on data stemming from Twitter and public Facebook pages. Private Facebook pages and, more importantly, private messaging apps are rarely examined. For many reasons, including ethical ones, studying these sources is a challenging endeavor. However, as a substantive part of our daily communication runs through these channels, academic research into the prevalence of deceptive information is needed.

As intentionally deceptive information is a phenomenon with potentially far-reaching consequences for democratic stability, the question arises how the identified features may facilitate countermeasures such as fact-checking activities. This question is even more relevant, as many scholars engaged with automated deception detection, propose human-in-the-loop solutions to ultimately detect fake news (Bourgonje, Schneider, and Rehm 2017; Conroy, Rubin, and Chen 2015). Linguistic analyses provide valuable tools but cannot fully replace human judgments (Rehm 2017), as automated models may only detect certain types of fake content (Ruchansky, Seo, and Liu 2017) and may be vulnerable to false positives, for example when real news is under-written or deals with specific topics (Zhou et al. 2019). There is a rich literature on the effectiveness of several fact-checking tactics (e.g., Nyhan and Reifler 2015; Nyhan et al. 2020; Walter et al. 2020), that also highlights some ongoing controversies around this rather new and expanding practice, such as accusations of partisanship (Graves and Cherubini 2016; Graves 2017). Our review may be of use to this fast-growing industry, especially in relation to the veracity of information in news-like texts. Identifying features that can draw from substantial empirical support—e.g., use of emotions—may facilitate a more effective selection of material to be scrutinized. The partisan bias identified in this paper is particularly relevant in light of ongoing discussions about the objectivity of fact-checking. As fact-checking is largely motivated by journalistic ideals (Graves, Nyhan, and Reifler 2016), objectivity is a first prerequisite for any fact-checking activity. To prevent any criticism of taking sides, some fact-checking organizations have taken a methodological approach to maintain political balance (Graves and Cherubini 2016). The observation that intentionally deceptive information is *not* equally spread across the political spectrum, may justify a focus on the more right-leaning end of it. Finally, recent work by Clayton et al. (2020) suggests that social media platforms may combat fake news by providing deceptive headlines with a warning flag (“rated false”).

The selection of these headlines may benefit from our review, that identifies long and sensationalist ones as indicative of deceptive content.

## Conclusions

This paper has reviewed research related to content features of intentionally deceptive information and sought to combine scattered knowledge into a single overview. Estimating the likelihood of news being fake—based on content alone and in the absence of additional information such as source cues or network characteristics—is never watertight, and in some cases, it might approach an educated guess. With this paper, we hope to contribute to the “educated” part of this process and to provide more solid ground for informed estimations. Based on a systematic review of the literature and a careful examination of the underlying scientific evidence, we identify several elements as agreed-upon indicators of intentionally deceptive news content: an ideological/partisan bias and the use of negative emotions provoking anger or fear. In addition, low levels of verifiability, the presence of a long sensationalist headline in which a lot of information is packed, and the use of informal language or swear, further increase the probability of dealing with intentionally deceptive news items. Interestingly, the features identified are rather similar for social media. However, as most research thus far has focused on the spread of disinformation *through* social media, less attention has been devoted to deceptive user generated content on the platforms themselves.

For several reasons, we believe that this review is important. First, it may contribute to the development of methods that enable a more reliable detection of deceptive information. This has proven to be a challenging endeavor. While the relevance of such a method clearly crosses the boundaries of a single discipline, there have not been much interdisciplinary collaborations. Second, the list of features may facilitate a more consistent operationalization of fake news, which can be applied in new research settings such as experiments and content analyses. Especially the latter may help scholars to examine the prevalence of the phenomenon, in general and across topics (see also Tucker et al. 2018, 56).

Of course, this paper is not without limitations. First of all, our review is inevitably incomplete. We selected studies by using the search terms *misinformation*, *disinformation*, and *fake news*. Research that engages with these and related phenomena, spanning several disciplines, is growing at a frantic pace with new and insightful studies being published every week. To minimize selection bias, we adopted a systematic approach and followed up on all references, but we nevertheless realize that results might change as new studies are being published. Second, a sole focus on content implies that we disregard other sorts of information that may be indicative of deception. Most notably, the specific source (i.e., website, platform) of an article and the visual material that accompanies it, may have a bearing on the perceived credibility (e.g., Hameleers et al. 2020). Krafft and Donovan (2020), for example, show that visual “evidence” collages are a key strategic element in the formation and spread of disinformation. Related to social media, Zhang and Ghorbani (2020) argue that an analysis of content alone is not sufficient for (automated) deception detection, as it should be combined with an analysis of other sorts of data such as information relating to author and user characteristics (see also Bondielli

and Marcelloni 2019). Finally, the focus on textual characteristics also implies that we do not consider other sorts of online deceptive messages, such as deepfakes. Given their potential impact on trust in news (e.g., Vaccari and Chadwick 2020), it is important that future research takes up this challenge.

We are fully aware of the fact that any review or line of research enhancing the identification of intentionally deceptive information does not change the increasingly thorny relationship between news and truth. Nor do fact checking activities or measures against misleading content on social media platforms. Graves and Wells convincingly (2019) argue that scholars tend to perceive the fake news phenomenon mainly as a matter of information availability, which has led to a focus on remedying falsehoods by containing the spread of bad information and replacing it with facts. In a sense, our paper fits in this tradition as it also builds on the assumption that when falsehoods are detected, we have moved closer towards a cure. It needs to be noted though that such an informational approach does not have much leverage if or when truth is not embedded in and supported by institutions that enable and require citizens as well as institutions to respond in good faith to public truth claims (Graves and Wells 2019, 39–40). In other words: When intentionally deceptive information is successfully detected and reported, and yet it does not seem to matter, there is not much to be won. If this is the case, more fundamental system change might be necessary to restore the power of truth in modern democracy.

## Notes

1. <https://www.bbc.com/news/technology-52281315>.
2. See: [https://www.meteo.expert/wp-content/uploads/2018/03/The-Global-Risks-Report-2018\\_light\\_MMCLogo.pdf](https://www.meteo.expert/wp-content/uploads/2018/03/The-Global-Risks-Report-2018_light_MMCLogo.pdf).
3. More specifically, when articles were published in journals not published by established publishers, we checked the journals using [www.predatoryjournals.com](http://www.predatoryjournals.com) and [www.journalguide.com](http://www.journalguide.com).

## Disclosure Statement

No potential conflict of interest was reported by the author(s).

## Funding

This work is part of the research project “Knowledge Resistance: Causes, Consequences, Cures,” funded by Riksbankens Jubileumsfond for the Advancement of the Humanities and Social Sciences [grant number M18-0310:1].

## ORCID

Alyt Damstra  <http://orcid.org/0000-0001-7753-018X>  
Hajo G. Boomgaarden  <http://orcid.org/0000-0002-5260-1284>  
Elina Lindgren  <http://orcid.org/0000-0002-9032-1347>  
Jesper Strömbäck  <http://orcid.org/0000-0001-7392-9791>  
Yariv Tsfati  <http://orcid.org/0000-0001-7395-3257>  
Rens Vliegenthart  <http://orcid.org/0000-0003-2401-2914>

## References

- Abonizio, H. Q., J. I. de Moraes, G. M. Tavares, and S. Barbon Junior. 2020. "Language-Independent Fake News Detection: English, Portuguese, and Spanish Mutual Features." *Future Internet* 12 (5), 87.
- Ahmed, H., I. Traore, and S. Saad. 2018. "Detecting Opinion Spams and Fake News Using Text Classification." *Security and Privacy* 1 (1): e9.
- Ahmed, W., J. Vidal-Alaball, J. Downing, and F. L. Seguí. 2020. "Dangerous Messages or Satire? Analysing the Conspiracy Theory Linking 5G to COVID-19 Through Social Network Analysis." *Journal of Medical Internet Research*. Advance online publication. doi:10.2196/preprints.19458.
- Allcott, H., and M. Gentzkow. 2017. "Social Media and Fake News in the 2016 Election." *Journal of Economic Perspectives* 31 (2): 211–236.
- Asubiaro, T. V., and V. L. Rubin. 2018. "Comparing Features of Fabricated and Legitimate Political News in Digital Environments (2016–2017)." *Proceedings of the Association for Information Science and Technology* 55 (1): 747–750.
- Bakir, V., and A. McStay. 2018. "Fake News and the Economy of Emotions: Problems, Causes, Solutions." *Digital Journalism* 6 (2): 154–175.
- Benkler, Y., R. Faris, and H. Roberts. 2018. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. New York: Oxford University Press.
- Bondielli, A., and F. Marcelloni. 2019. "A Survey on Fake News and Rumour Detection Techniques." *Information Sciences* 497: 38–55.
- Bonet-Jover, A., A. Piad-Morffis, E. Saquete, P. Martínez-Barco, and MÁ García-Cumbreras. 2021. "Exploiting Discourse Structure of Traditional Digital Media to Enhance Automati Fake News Detection." *Expert Systems with Applications* 169 (2021): 114340.
- Bourgonje, P., J. M. Schneider, and G. Rehm. 2017. "From Clickbait to Fake News Detection: An Approach Based on Detecting the Stance of Headlines to Articles." In *Proceedings of the 2017 EMNLP Workshop: Natural Language Processing Meets Journalism*, 84–89.
- Bradshaw, S., P. N. Howard, B. Kollanyi, and L. M. Neudert. 2020. "Sourcing and Automation of Political News and Information Over Social Media in the United States, 2016–2018." *Political Communication* 37 (2): 173–193.
- Chen, Y., N. J. Conroy, and V. L. Rubin. 2015. "Misleading Online Content: Recognizing Clickbait as" False News"." In *Proceedings of the 2015 ACM on Workshop o Multimodal Deception Detection*, 15–19, November.
- Chu, S. K. W., R. Xie, and Y. Wang. 2021. "Cross-Language Fake News Detection." *Data and Information Management* 5 (1): 100–109.
- Clayton, K., S. Blair, J. A. Busam, S. Forstner, J. Glance, G. Green, A. Kawata, et al. 2020. "Real Solutions for Fake News? Measuring the Effectiveness of General Warnings and Fact-Check Tags in Reducing Belief in False Stories on Social Media." *Political Behavior* 42 (4): 1073–1095.
- Conroy, N. J., V. L. Rubin, and Y. Chen. 2015. "Automatic Deception Detection: Methods for Finding Fake News." *Proceedings of the Association for Information Science and Technology* 52 (1): 1–4.
- Depoux, A., S. Martin, E. Karafillakis, R. Preet, A. Wilder-Smith, and H. Larson. 2020. "The Pandemic of Social Media Panic Travels Faster Than the COVID-19 Outbreak." *Journal of Travel Medicine* 27 (3): 1–2.
- Egelhofer, J. L., and S. Lecheler. 2019. "Fake News as a Two-Dimensional Phenomenon: a Framework and Research Agenda." *Annals of the International Communication Association* 43 (2): 97–116.
- Faris, R., H. Roberts, B. Etling, N. Bourassa, E. Zuckerman, and Y. Benkler. 2017. *Partisanship, Propaganda, and Disinformation: Online Media and the 2016 US Presidential Election*. *Berkman Klein Center Research Publication*, 6.
- Faustini, P. H. A., and T. F. Covões. 2020. "Fake News Detection in Multiple Platforms and Languages." *Expert Systems with Applications* 158: 113503.
- Freelon, D., and T. Lokot. 2020. "Russian Twitter Disinformation Campaigns Reach Across the American Political Spectrum." *Harvard Kennedy School Misinformation Review* 1 (1): 1–9.



- Freelon, D., and C. Wells. 2020. "Disinformation as Political Communication." *Political Communication* 37 (2): 145–156.
- Fuller, C. M., D. P. Biros, and D. Delen. 2011. "An Investigation of Data and Text Mining Methods for Real World Deception Detection." *Expert Systems with Applications* 38 (7): 8392–8398.
- Graves, L. 2017. "Anatomy of a Fact Check: Objective Practice and the Contested Epistemology of Fact Checking." *Communication, Culture & Critique* 10 (3): 518–537.
- Graves, L., and F. Cherubini. 2016. *The Rise of Fact-checking Sites in Europe*.
- Graves, L., B. Nyhan, and J. Reifler. 2016. "Understanding Innovations in Journalistic Practice: A Field Experiment Examining Motivations for Fact-Checking." *Journal of Communication* 66 (1): 102–138.
- Graves, L., and C. Wells. 2019. "From Information Availability to Factual Accountability." In *Journalism and Truth in an Age of Social Media*, edited by J. E. Katz, and K. K. Mays, 39–57. New York: Oxford University Press.
- Grindberg, N., K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer. 2019. "Fake News on Twitter During the 2016 U.S. Presidential Election." *Science* 363 (6425): 374–378.
- Guess, A., B. Nyhan, and J. Reifler. 2018. "Selective Exposure to Misinformation: Evidence from the Consumption of Fake News during the 2016 US Presidential Campaign." European Research Council, 9. <http://www.dartmouth.edu/~nyhan/fake-news2016.pdf>.
- Gupta, A., P. Kumaraguru, C. Castillo, and P. Meier. 2014. "Tweetcred: Real-time Credibility Assessment of Content on Twitter." In *International Conference on Social Informatics*, 228–243. Cham: Springer, November.
- Hameleers, M. 2020. "Separating Truth from Lies: Comparing the Effects of News Media Literacy Interventions and Fact-checkers in Response to Political Misinformation in the US and Netherlands." *Information, Communication & Society*: 1–17. Advance online publication. doi: 10.1080/1369118X.2020.1764603
- Hameleers, M., T. E. Powell, T. G. L. A. Van Der Meer, and L. Bos. 2020. "A Picture Paints a Thousand Lies? The Effects and Mechanisms of Multimodal Disinformation and Rebuttals Disseminated via Social Media." *Political Communication* 37 (2): 281–301.
- Hameleers, M., and G. L. A. Van der Meer. 2019. "Misinformation and Polarization in High Choice Media Environment: How Effective are Political Fact-Checkers?" *Communication Research*. doi:10.1177/0093650218819671.
- Hameleers, M., T. van der Meer., and R. Vliegenthart. 2021. "Civilized Truths, Hateful Lies? Incivility and Hate Speech in False Information—evidence from Fact-checked Statements in the US." *Information, Communication & Society*: 1–18. Advance online publication. doi:10.1080/1369118X.2021.1874038.
- Horne, B. D., and S. Adali. 2017. "This Just in: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire Than Real News." Eleventh International AAAI Conference on Web and Social Media, May.
- Howard, P. N., B. Ganesh, D. Liotsiou, J. Kelly, and C. François. 2019. *The IRA, Social Media and Political Polarization in the United States, 2012–2018*. Computational Propaganda Research Project (17 December 2018). <https://comprop.oii.ox.ac.uk/wpcontent/uploads/sites/93/2018/12/IRA-Report-2018.pdf>.
- Humpherys, S. L., K. C. Moffitt, M. B. Burns, J. K. Burgoon, and W. F. Felix. 2011. "Identification of Fraudulent Financial Statements Using Linguistic Credibility Analysis." *Decision Support Systems* 50 (3): 585–594.
- Humprecht, E. 2019. "Where 'Fake News' Flourishes: a Comparison Across Four Western Democracies." *Information, Communication & Society* 22 (13): 1973–1988.
- Kalogeropoulos, A., R. Fletcher, and R. K. Nielsen. 2019. "News Brand Attribution in Distributed Environments: Do People Know Where They Get Their News?" *New Media & Society* 21 (3): 583–601.
- Krafft, P. M., and J. Donovan. 2020. "Disinformation by Design: The Use of Evidence Collages and Platform Filtering in a Media Manipulation Campaign." *Political Communication* 37 (2): 194–214.
- Lasorsa, D. L., and J. Dai. 2007. "Newsroom's Normal Accident? An Exploratory Study of 10 Cases of Journalistic Deception." *Journalism Practice* 1 (2): 159–174.

- Lazer, D., M. Baum, N. Grinberg, L. Friedland, K. Joseph, W. Hobbs, and C. Mattson. 2017. Combating Fake News: An Agenda for Research and Action. <https://shorensteincenter.org/combating-fake-news-agenda-for-research/>.
- Liu, Y., K. Yu, X. Wu, L. Qing, and Y. Peng. 2019. "Analysis and Detection of Health-Related Misinformation on Chinese Social Media." *IEEE Access* 7 (1): 154480–154489.
- Marchal, N., B. Kollanyi, L. M. Neudert, and P. N. Howard. 2019. *Junk News During the EU Parliamentary Elections: Lessons from a Seven-Language Study of Twitter and Facebook*. Oxford: Oxford Internet Institute.
- Marwick, A., and R. Lewis. 2017. *Media Manipulation and Disinformation Online*. New York: Data & Society Research Institute.
- Neudert, L. M., P. Howard, and B. Kollanyi. 2019. "Sourcing and Automation of Political News and Information During Three European Elections." *Social Media+ Society* 5 (3): 1–13.
- Neudert, L., B. Kollanyi, and P. N. Howard. 2017. *Junk News and Bots During the German Parliamentary Election: What are German Voters Sharing Over Twitter?*
- Newman, M. L., J. W. Pennebaker, D. S. Berry, and J. M. Richards. 2003. "Lying Words: Predicting Deception from Linguistic Styles." *Personality and Social Psychology Bulletin* 29 (5): 665–675.
- Nielsen, R. K., and L. Graves. 2017. "News You Don't Believe": Audience Perspectives on Fake News.
- Nyhan, B., E. Porter, J. Reifler, and T. J. Wood. 2020. "Taking Fact-checks Literally but not Seriously? The Effects of Journalistic Fact-checking on Factual Beliefs and Candidate Favorability." *Political Behavior* 42 (3): 939–960.
- Nyhan, B., and J. Reifler. 2015. "The Effect of Fact-Checking on Elites: A Field Experiment on U.S. State Legislators." *American Journal of Political Science* 59 (3): 628–640.
- Omezi, N., and H. Jahankhani. 2020. "Proposed Forensic Guidelines for the Investigation of Fake News." In *Policing in the Era of AI and Smart Societies*, edited by Hamid Jahankhani, Babak Akhgar, Peter Cochrane, and Mohammad Dastbaz, 231–265. Cham: Springer.
- Ortoleva, P. 2019. "Canards, Fausses Nouvelles, Paranoid Style." In *Journalism and Truth in an Age of Social Media*, edited by J. E. Katz, and K. K. Mays, 119–132. New York: Oxford University Press.
- Pierri, F., A. Artoni, S. Ceri, and F. Zollo. 2020. "Investigating Italian Disinformation Spreading on Twitter in the Context of 2019 European Elections." *PloS one* 15 (1): e0227821.
- Potthast, M., S. Köpsel, B. Stein, and M. Hagen. 2016. "Clickbait Detection." In *Advances in Information Retrieval*. ECIR 2016. Lecture Notes in Computer Science, Vol. 9626, edited by N. Ferro, et al., 1–8. Cham: Springer.
- Rashkin, H., E. Choi, J. Y. Jang, S. Volkova, and Y. Choi. 2017. "Truth of Varying Shades: Analyzing Language in Fake News and Political Fact-checking." In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, 2931–2937.
- Reddy, H., N. Raj, M. Gala, and A. Basava. 2020. "Text-mining-based Fake News Detection Using Ensemble Methods." *International Journal of Automation and Computing* 17 (2): 210–221..
- Rehm, G. 2017. "An Infrastructure for Empowering Internet Users to Handle Fake News and Other Online Media Phenomena." In *International Conference of the German Society for Computational Linguistics and Language Technology*, edited by Georg Rehm and Thierry Declerck, 216–231. Cham: Springer.
- Ruchansky, N., S. Seo, and Y. Liu. 2017. "CSI: A Hybrid Deep Model for Fake News Detection." In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 797–806.
- Sahoo, S. R., and B. B. Gupta. 2021. "Multiple Features Based Approach for Automatic Fake News Detection on Social Networks Using Deep Learning." *Applied Soft Computing* 100: 106983.
- Schudson, M. 2019. "Belgium Invades Germany: Can Facts Survive Politics?" In *Mays Journalism and Truth in an Age of Social Media*, edited by J. E. Katz and K. K. Mays, 21–35. doi:10.1093/oso/9780190900250.001.0001.
- Shu, K., A. Sliva, S. Wang, J. Tang, and H. Liu. 2017. "Fake News Detection on Social Media: A Data Mining Perspective." *ACM SIGKDD Explorations Newsletter* 19 (1): 22–36.
- Silverman, C. 2016a. "Viral Fake Election News Outperformed Real News on Facebook." BuzzFeed News. See: <https://www.buzzfeednews.com/article/craigsilverman/viralfake-election-news-outperformed-real-news-on-facebook>.

- Silverman, C. 2016b. "Here are 50 of the Biggest Fake News Hits on Facebook from 2016." Buzzfeed News. See: <https://www.buzzfeednews.com/article/craigsilverman/top-fakenews-of-2016>.
- Søe, S. O. 2019. "A Unified Account of Information, Misinformation, and Disinformation." *Synthese* 198 (6): 5929–5949.
- Srivastava, A., G. Rehm, and J. M. Schneider. 2017. "DFKI-DKT at SemEval-2017Task 8: Rumour Detection and Classification Using Cascading Heuristics." In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval 2017)*, 486–490.
- Tandoc, E. C., Jr., Z. W. Lim, and R. Ling. 2018. "Defining "Fake News" A Typology of Scholarly Definitions." *Digital Journalism* 6 (2): 137–153.
- Tfsati, Y., J. Strömbäck, H. Boomgaarden, R. Vliegenthart, E. Lindgren, and A. Damstra. 2020. "Causes and Consequences of Mainstream Media Dissemination of Fake News: Literature Review and Synthesis." *Annals of the International Communication Association*. Advance online publication, doi:10.1080/23808985.2020.1759443.
- Tucker, J. A., A. Guess, P. Barberá, C. Vaccari, A. Siegel, S. Sanovich, D. Stukal, and B. Nyhan. 2018. "Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature." Political Polarization, and Political Disinformation: A Review of the Scientific Literature.
- Vaccari, C., and A. Chadwick. 2020. "Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News." *Social Media+ Society* 6 (1): 2056305120903408.
- Van Der Zee, S., R. Poppe, A. Havrileck, and A. Baillon. 2018. A Personal Model of Trumpery: Deception Detection in a Real-world High-stakes Setting. arXiv preprint arXiv:1811.01938.
- Volkova, S., K. Shaffer, J. Y. Jang, and N. Hodas. 2017. "Separating Facts from Fiction: Linguistic Models to Classify Suspicious and Trusted News Posts on Twitter." In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics* (Volume 2: Short Papers), 647–653, July.
- Wahl-Jorgensen, K. 2013. "Subjectivity and Story-Telling in Journalism: Examining Expressions of Affect, Judgement and Appreciation in Pulitzer Prize-Winning Stories." *Journalism Studies* 14 (3): 305–320.
- Wahl-Jorgensen, K. 2019. *Emotions, Media and Politics*. Cambridge: John Wiley & Sons.
- Wahl-Jorgensen, K. 2020. "An Emotional Turn in Journalism Studies?" *Digital Journalism* 8 (2): 175–194.
- Waisbord, S. 2018. "Truth is What Happens to News: On Journalism, Fake News, and Post Truth." *Journalism Studies* 19 (13): 1866–1878.
- Walter, N., J. Cohen, R. L. Holbert, and Y. Morag. 2020. "Fact-checking: A Meta-Analysis of What Works and for Whom." *Political Communication* 37 (3): 350–375.
- Wang, W. Y. 2017. "Liar, Liar Pants on Fire": A New Benchmark Dataset for Fake News.
- Wingfield, N., M. Isaac, and K. Benner. 2016. "Google and Facebook Take Aim at Fake News Sites." *The New York Times*, 11, 12. Accessed November 1, 2019. [https://www.mediapicking.com/medias/files\\_medias/nytimes—google-and-facebooktake-aim-at-fake-news-sites-0237488001479491012.pdf](https://www.mediapicking.com/medias/files_medias/nytimes—google-and-facebooktake-aim-at-fake-news-sites-0237488001479491012.pdf).
- Zhang, X., and A. A. Ghorbani. 2020. "An Overview of Online Fake News: Characterization, Detection, and Discussion." *Information Processing & Management* 57 (2): 102025.
- Zhou, L., J. K. Burgoon, J. F. Nunamaker, and D. Twitchell. 2004. "Automating Linguistics-Based Cues for Detecting Deception in Text-Based Asynchronous Computer-Mediated Communications." *Group Decision and Negotiation* 13 (1): 81–106.
- Zhou, Z., H. Guan, M. M. Bhat, and J. Hsu. 2019. Fake News Detection via NLP is Vulnerable to Adversarial Attacks. arXiv preprint arXiv:1901.09657.
- Zhou, X., and R. Zafarani. 2019. "Network-based Fake News Detection: A Pattern-Driven Approach." *ACM SIGKDD Explorations Newsletter* 21 (2): 48–60.
- Zhou, X., and R. Zafarani. 2020. "A Survey of Fake News: Fundamental Theories, Detections Methods & Opportunities." *ACM Computing Surveys* 53 (5): 1–40.
- Zimmermann, F., and M. Kohring. 2020. "Mistrust, Disinforming News, and Vote Choice: A Panel Survey on the Origins and Consequences of Believing Disinformation in the 2017 German Parliamentary Election." *Political Communication* 37 (2): 215–237.