Fake News: A Definition

Axel Gelfert

En dépit de son utilisation récente, l'expression «fausses nouvelles» a évolué rapidement. Cet article soutient qu'elle devrait être réservée aux présentations intentionnelles d'allégations (typiquement) fausses ou trompeuses comme si elles étaient des nouvelles véridiques et où elles sont faussées à dessein. L'expression «à dessein» fait ici référence à des caractéristiques systémiques de la conception des sources et des canaux par lesquels les fausses nouvelles se propagent et par conséquent, manipulent les processus cognitifs du public. Cette définition prospective est ensuite mise à l'épreuve: d'abord, en opposant les fausses nouvelles à d'autres formes de désinformation publique; deuxièmement, en examinant si elle aide à cerner les conditions de la prolifération (récente) de fausses nouvelles.
Fake News: A Definition

AXEL GELFERT

Department of Philosophy, Literature, History of Science and Technology
Technical University of Berlin
Berlin, 10623
Germany
axel@gelfert.net

Abstract: Despite being a new term, ‘fake news’ has evolved rapidly. This paper argues that it should be reserved for cases of deliberate presentation of (typically) false or misleading claims as news, where these are misleading by design. The phrase ‘by design’ here refers to systemic features of the design of the sources and channels by which fake news propagates and, thereby, manipulates the audience’s cognitive processes. This prospective definition is then tested: first, by contrasting fake news with other forms of public disinformation; second, by considering whether it helps pinpoint conditions for the (recent) proliferation of fake news.

Résumé: En dépit de son utilisation récente, l’expression «fausses nouvelles» a évolué rapidement. Cet article soutient qu'elle devrait être réservée aux présentations intentionnelles d’allégations (typiquement) fausses ou trompeuses comme si elles étaient des nouvelles véridiques et où elles sont faussées à dessein. L'expression «à dessein» fait ici référence à des caractéristiques systémiques de la conception des sources et des canaux par lesquels les fausses nouvelles se propagent et par conséquent, manipulent les processus cognitifs du public. Cette définition prospective est ensuite mise à l’épreuve: d'abord, en opposant les fausses nouvelles à d'autres formes de désinformation publique; deuxièmement, en examinant si elle aide à cerner les conditions de la prolifération (récente) de fausses nouvelles.

Keywords: fake news, cognitive biases, social epistemology, media literacy, critical reasoning
1. Introduction

Recent political events—notably the controversially close Brexit referendum in the U.K. and the narrow win of Donald Trump in the U.S. presidential election of 2016—have led to a wave of interest in the phenomenon of “fake news”, which is widely believed to have played a significant role in shaping the outcome of both political contests. While much ink has been spilled, by academics and pundits alike, on its disruptive potential and deceptive nature, somewhat less attention has been paid to analyzing and defining the term ‘fake news’. It is, of course, quite natural that a term as recent and controversial as ‘fake news’ should be used in a variety of (sometimes conflicting) ways, thereby making conceptual analysis more difficult. However, if one holds that the term refers to a distinct and identifiable phenomenon, it is all the more important to attempt to come to conceptual grips with it and, if necessary, put forward a definition that can then be refined in the course of future scholarly debate. The present paper contributes to the nascent debate on the concept of fake news, keeping in mind that it derives its significance from the real-world effects of the proliferation of fake news, both as a distinct class of misleading reports and as a rhetorical device for shutting down critical reporting.

The rest of the paper is organized as follows: Section 2 sets the scene by reviewing the role that (well-functioning) news plays in our epistemic lives and by tracing the discussion of “fake news” from its nineteenth-century origins to its most recent resurgence in the run-up to the 2016 U.S. presidential election. Section 3 reviews a number of previous attempts, by journalists, commentators and scholars, to define ‘fake news’, and goes on to identify their shortcomings. Section 4, starting from the broader notion of disinformation (Fallis 2015), gradually develops a workable stipulative definition of ‘fake news’. Fake news, I argue, is best defined as the deliberate presenta-
tion of (typically) false or misleading claims as news, where the claims are misleading by design. The phrase “by design” is then explicated in terms of systemic features of the process of news production and dissemination. Finally, in Section 5, the definition is put to the test in two ways: first, by analyzing whether it can successfully distinguish fake news from other forms of public disinformation; second, by considering whether it is able to highlight conditions for the emergence of fake news, thereby contributing to an understanding of its most recent resurgence. In particular, I argue that online social media have enabled purveyors of fake news to target specific audiences and exploit well-documented cognitive biases and heuristics in an attempt to mislead consumers into propagating fake news claims even further.

2. News, fakes, and “FAKE NEWS”

Before turning to recent attempts to define the compound term ‘fake news’, it is important to reflect briefly on its component parts: ‘news’ and ‘fake’. For, if we are to understand the distinctive aspects of the recent onslaught of fabricated and misleading information that passes itself off as news (i.e., “is faking it”), we must first understand some of the core epistemic functions of the original—what is now, sometimes in a derogatory manner, called “mainstream media”, but which used to be called “news” simpliciter.

To the extent to which news media have received attention in epistemology, the emphasis has been on their reliability as a source of factual information for the consumer. On this account, consumers of news media are the recipients of a specialized form of testimony, and anyone who relies on, say, the New York Times or the BBC World News as authoritative sources of factual reports about the world incurs broadly the same epistemic rights and responsibilities as any recipient of ordinary testimony does. At the same time, it is typically un-
understood that consumers of news also incur epistemic obligations. For example, it is assumed that competent readers can distinguish between factual reports and opinion pieces, and will use the same basic screening methods (e.g., for errors and inconsistencies) they would be using when evaluating the testimony of a trusted and competent interlocutor. Whereas epistemologists typically frame the issue as one of evaluating which news sources (or individual reports) to trust, and which to reject, argumentation theorists focus on the recipient’s ability to critically question appeals to the authority of reports or reporters. From the latter perspective, naively accepting reports without further analysis comes dangerously close to committing the fallacy of the *argumentum ad verecundiam* (i.e., the fallacy of submitting to a potentially irrelevant authority); trust in putative epistemic authorities is by necessity provisional, and basic critical questions—concerning the credibility of the source, its reliability, motives, interests, consistency and track record—should never be entirely suppressed (Walton 1997).

Treating news media as sources of testimonial beliefs emphasizes a central role they play in our epistemic lives: serving as the source of much of what we take ourselves to know. Without reliance on curated news reports we would know precious little about what goes on in other countries, about the world of politics, or even about what the latest sports results are. Of course, communicating reliable information to their readership or viewership is not their only social function. For one, commercial news sources need to return a profit to their proprietors, and even individual consumers do not always turn to the news for knowledge and information only, but may also hope for some entertainment and diversion. Yet, as far as its epistemic significance is concerned, the first—and certainly most widely appreciated—epistemic function of the news is to furnish us with reliable factual information. Put crudely, if a
reputable news source truthfully reports that \( p \), we can come to know that \( p \) simply by taking that report at face value.

Yet, providing its consumers with knowledgeable reports, ready for uptake, is not the only epistemic function of traditional news media. In addition to acquiring specific beliefs by accepting the corresponding news reports, we also rely on the news media for overall coverage. The notion of *epistemic coverage* (Goldberg 2011) refers to those features of an epistemic agent’s epistemic environment that underwrite “the believer's reliance on a source to be both reliably apprised of the facts in a certain domain, and reliably disposed to report on the obtaining of facts in that domain” (Goldberg 2011, p. 93). Whether an epistemic environment reliably provides an agent with coverage is, of course, an entirely contingent matter and depends both on the sources present in the agent’s environment and on the agent’s “epistemic routines” (Gelfert 2014, p. 208)—that is, on whether they regularly consult those sources in the right way. It only takes a moment’s reflection to realize that the particular news media we consume will significantly shape the extent to which we enjoy epistemic coverage. If our environment’s epistemic coverage is deficient, we will no longer be reliably apprised of significant changes in the world around us, and whatever knowledge we may have acquired at the initial point of belief formation will gradually erode. The news, then, serves not only as a direct source of new beliefs, but also (assuming things go well) provides us with epistemic coverage, which in turn safeguards existing knowledge by keeping us abreast of changes in the world.

Finally, the news media also functions as an indirect source of knowledge about one’s epistemic community. For, traditionally at least, the news has been presented in a fixed, aggregate form, e.g., as a news bulletin of a certain duration, or a newspaper edition with a certain number of pages. This necessitates not only tough editorial choices regarding what to in-
clude, and what to leave out, but also presupposes knowledge of the target audience—since, for obvious commercial reasons, the selection of topics must have sufficiently broad appeal. The consumer, in turn, can infer from the selection of articles or news segments what other people in his or her community tend to be interested in. In other words, the selection of topics ‘bundled’ together in a given newspaper edition or news bulletin carries *meta-information* about the reader’s or viewer’s community. As the journalist Stefan Schulz puts it, “Newspapers inform people about what information other people seek out when they wish to learn about the world” (Schulz 2013, p. 38). Such “second-order observation” (*ibid.*) is an important factor in the constitution of what Jürgen Habermas has called the “public sphere”, which in turn provides the basis for the free and frank exchange of ideas and opinions, from which, according to a standard liberal tenet, the truth will emerge.¹

Epistemological analyses of traditional news media have, on the whole, been optimistic, not least by comparison with their electronic counterparts. Here, for example, is Alvin Goldman on the advantages of traditional newspapers over online blogs:

Newspapers employ fact checkers to vet a reporter’s article before it is published. They often require more than a single source before publishing an article, and limit reporters’ reliance on anonymous sources. These practices seem likely to raise the veritistic quality of the reports newspapers publish and hence the veritistic quality of their readers’ resultant beliefs. (Goldman 2008, p. 117)

Yet, lest we succumb to a naive—and by now nostalgic—view of journalists selflessly gathering “all the news that’s fit to print” and publishing it “without fear or favour” it is worth

¹ John Stuart Mill (1859) speaks of emergence of the truth from “collision with error”.

noting that, historically, the newspaper industry has been no stranger to bias, distortion, manipulation, and outright fabrication. As Robert Love puts it, “[i]n the early days of American journalism, newspapers trafficked in intentional, entertaining hoaxes” (Love 2007, p. 33). Among the most influential figures in this regard was William Randolph Hearst, publisher of the New York Journal, whose war-mongering was in no small measure the reason the United States, in 1897, entered into a war with Spain, following the explosion of the USS Maine in the harbour of Havana, which Hearst’s paper resolutely blamed on the Spanish. Competing papers, such as the New York Evening News, lamented the “gross misrepresentation of the facts” and “deliberate invention of tales calculated to excite the public”. More recently, “CNN and the New York Times were used by the U.S. military as unwitting co-conspirators in spreading false information, a tactic known as psychological operations” (Love 2007, p. 34), in the run-up to the invasion of Iraq.

In short, the faking of news stories has been around for a long time, and every iteration of technological advancement, from the telegraph in the 19th century to contemporary social media algorithms, has unleashed new possibilities of deception and fabrication. Writing in the journal Arena in 1898, the critic J. B. Montgomery-McGovern, wrote a scathing piece titled “An important phase of gutter journalism: Faking”, in which he complained about “fake journalism”, which “is resorted to chiefly by news bureaus […] which supply nearly all the metropolitan […] dailies with their most sensational ‘stories’” (1898, p. 240). Among the techniques employed by ‘fakers’ was the use of the “stand-for”, usually a reputable member of the community such as “a doctor, dentist, architect, or other professional or business man” who, for money, would corroborate the story to any reporter that the local paper might send to

---

2 Cited after (McCaffrey 2009, p. 6).
investigate its *bona fides*. “Fake news” then involved deception, not only of the consumer but also of the middle-man: the editor of the local paper who, “now constantly on the alert for ‘fake’ stories, is often deceived, and sends one of the reporters of his own paper to investigate the matter”, only for the local reporter to encounter the planted “stand-for”. By contrast, purveyors of fake news in the twenty-first century often cut out the middle-men and peddle their wares directly to readers, aided by the sharing of sensational stories by trusted acquaintances on social media.

It is worth pausing for a moment to reflect on the connotations of the term ‘fake’. While the term seems straightforward enough, upon closer inspection it reveals a rich variety of meanings. In the artworld, fakes derive their value entirely from the originals they successfully mimic, specifically from the scarcity of the latter. Indeed, the successful forging of an antiquity was sometimes regarded as a “triumph of artistry” (Lowenthal 1990, p. 17). As the art historian Mark Jones puts it, with some hyperbole: “Each society, each generation, fakes the thing it covets most” (1990, p. 13). Contemporary fake news, however, though no doubt parasitic on traditional news media—fake news websites typically mimic the “look and feel” of mainstream sources to garner credibility—does not seem to go hand in hand with an appreciation of the traditional news media. According to a Gallup poll, in September 2016 only 32% of U.S. Americans had “a great deal” or “a fair amount’ of trust” in the mass media when it comes to reporting the news—the lowest figure ever recorded (Swift 2016). Instead, for some consumers, fake news appears to swamp traditional news sources, even subverting the latter’s claim to authoritativeness. This seems to be the overt goal every time Donald Trump uses his Twitter account to denounce a critical news story as, in his trademark all-caps, “FAKE NEWS”, which he did a total of 73 times between 10 December 2016
and 24 July 2017 alone (Rosen 2017). Most of the time, the label is applied not to specific reports, but indiscriminately to news organizations such as CNN, MSNBC, or the New York Times. Even where specific reports are being targeted, these are often factually accurate (as in the true reports of low turnout at Trump’s inauguration rally) or are inaccurate merely as the result of (inevitable) honest mistakes, which are usually swiftly corrected (Stephens 2017).

As recent as the mid-2000s, when it came to be used in relation to satirical TV shows such as The Daily Show (which has been airing since 1996) and The Colbert Report (2005-2014), the term ‘fake news’ even had a mildly progressive ring to it. Mimicking the spin and double-speak of politicians and corporations, which often goes unchallenged on mainstream news media, this TV format has been credited with “conveying real messages”: “As fake news, it satirizes traditional news by reporting in a style similar to network and cable TV news, but it amplifies their biases, mistakes, and deficiencies to ensure that viewers hear them loud and clear” (Gettings 2007, pp. 26-27). Throughout, “[v]arious cues let the audience draw the line pretty clearly between the fake and the real, and the moderately attentive viewer leaves the show better informed about the world, especially the political climate and current state of the media” (ibid.). In satirizing the shortcomings of traditional news media, the “fake news” of The Daily Show “necessitates assumptions about some kind of authentic or legitimate set of news practices” (Baym 2005, p. 262); in other words, it presupposes a certain level of media literacy on the part of its viewers.

Though only a decade old, the view that fake news could be an effective tool for promoting media literacy might now seem quaint, given how the term ‘fake news’ has proliferated in the last few years. No longer is it reserved for sophisticated—and ultimately self-critical media formats that aim to educate their
Fake News: A Definition

viewers. Instead, it has come to be associated with (often anonymous) sources that spread falsehoods by manipulating their consumers’ emotions and tapping into deeply held partisan beliefs. Indeed, legitimate concerns about the trustworthiness of (putative) news sources, which originally motivated the introduction of the term ‘fake news’, are at risk of being drowned out by tactical usage of the phrase in order to cast aspersions on legitimate news organizations. Thus, as Matthew Dentith has recently argued, “the threat that is accusations of ‘That’s just fake news’ comes out of worries that it is […] merely a rhetorical device used by the powerful to crush dissent” (2017, p. 65). Dentith even goes so far as to define fake news as “an allegation that some story is misleading” (p. 66; italics added). On this account, ‘fake news’ would refer not to the misleading claim itself, but to an appeal to the allegedly misleading nature of a claim—that is, it would be a rhetorical device for undermining a given claim’s authoritative status by alleging that it “lacks some context or additional piece of information which, when revealed, undermines either its truth-value, or saliency to some broader claim” (ibid). While I find Dentith’s suggestion intriguing, I do believe it puts the cart before the horse: arguably, both historically and as a socio-cultural phenomenon, the emergence of fake news as a genre of purported factual assertions preceded the emergence of the epithet ‘fake news’ as a tactical way of slandering one’s opponents. I shall briefly return to this point in Section 5. For now, in what follows, the main focus will be on fake news as a class of purportedly factual claims that are epistemically deficient (in a way that needs to be specified), rather than as an accusatory speech act.

3. Surveying recent characterizations of ‘fake news’

Given how recently the term ‘fake news’ has burst onto the political and intellectual scene since 2016, it is perhaps not
surprising that various characterizations of the phenomenon have proliferated. Indeed, the abundance of (tentative) definitions that have been floated has led some to worry that the heterogeneity of the term ‘fake news’ results in it becoming “a catch-all term with multiple definitions” (Lilleker 2017, p. 1). Others have urged journalists, in particular, to “stop calling everything ‘fake news’” (Oremus 2016). Part of the problem, perhaps exacerbated by the heated nature of political discourse following the events of 2016, has been the collapsing of various existing distinctions between different types of public disinformation. Media scholars, folklorists, and even a handful of social epistemologists have long been concerned with demarcating such phenomena as gossip, rumor, hoaxes, and urban legends (see Gelfert 2013). Thus, whereas gossip “possesses relevance only for a specific group” and “is disseminated in a highly selective manner within a fixed social network”, rumours are characterized by the propagation of “unauthorized messages that are always of universal interest and accordingly are disseminated diffusely” (Bergmann 1993, p. 70). Rumors, too, may sometimes become divorced from their specific origins and “solidify” into urban legends (Allport and Postman 1947, p. 162), which in turn may be picked up and reported as factual by mainstream news sources, often acquiring a local flavour through gradual embellishments and the addition of detail. Should urban legends that get reported in otherwise respectable news sources count as fake news? What about hoaxes? After all, hoaxes, in spite of being deliberately fabricated falsehoods that masquerade as the truth—similar, as we shall see, to what a number of definitions of the term ‘fake news’ assert—often serve quite different purposes and, unlike fake news, are typically intended to be found out eventually,
perhaps because they seek to expose the gullibility of a certain segment of the population.\(^3\)

Any putative definition of ‘fake news’ must be situated in relation to these varied forms of public disinformation and distortions of the communicative process. Actual usage, in this context, can hardly be expected to be the only arbiter of the meaning of the term, especially in light of the fact that it has, by now, been co-opted for contradictory political purposes: as a form of media criticism by Stephen Colbert and Jon Stewart, and as a (politically charged) negative epithet by Donald Trump. At the same time, the way the term is being employed—especially by those in academia and the media who have given the underlying phenomenon their serious consideration—provides an important starting point that may bring into clearer focus a multiplicity of intersecting meanings and shared concerns. In what follows, I shall engage closely with a number of proposed definitions, keeping in mind that many of these have been put forward in the context of applied questions, by practitioners from such diverse professions as law and journalism; as such, they do not aspire to the conceptual rigour prized by, say, academic epistemology. Yet, a growing number of professional philosophers, too, have recently entered the fray and have put forward promising definitions of ‘fake news’; while my own definition, to be developed over the next two sections, will differ from these (and, I believe, for good reason), I shall also endeavour to point out continuities and similarities. My overall approach, then, will be in part stipulative, in an attempt to advance the debate by highlighting novel features of the phenomenon of ‘fake news’ that might otherwise be overlooked, and in part precisifying. I take it to

\(^3\) Consider the notorious Sokal hoax, when theoretical physicist Alan Sokal published a fabricated article in *Social Text*, intending to expose the uneven scholarly standards in certain parts of postmodern cultural studies. (See Lingua Franca 2000.)

be a fitting methodological assumption, in particular as concerns a highly controversial phenomenon, that conceptual progress is best achieved through a constructive engagement with a wide range of theoretical approaches and practical concerns.

Before turning to individual definitions in their context, perhaps a quick preview is in order. Surveying extant characterizations of fake news, a number of recurring themes can be readily identified. First, there is the recognition that the medium of the internet (and social media, in particular) has been especially conducive to the creation and proliferation of fake news. Thus, ‘fake news’ is sometimes explicitly defined as “the online publication” of false statements of fact (Klein & Wueller 2017, p. 6), or it is noted that a “core feature of contemporary fake news is that it is widely circulated online” (Bakir & McStay 2017, p. 1). Second, the connection between the content of fake news and the world at large is dubious at best. Fake news, if it is not directly equated with false news (Oremus 2017), is thought to consist of stories “invented entirely from thin air”, to be “completely fabricated”, to transmit “new content [that] is 100% false”, and to have “no factual basis”. Fourth, an element of deliberateness is imputed to the creation and circulation of fake news, which in turn is deemed “deliberately misleading” and involving “intentionally or knowingly false statements of fact”, “deliberate spread of misinformation”, along with the “intention to deceive”. Given that these are widely shared sentiments in relation to fake news, it would certainly be a desideratum of any putative definition of fake news to shed light on their predominance. Why, for example, does fake news seem to have risen to new, or renewed, prominence with the mass adoption of online social media? How does fake news manage to maintain the

---

4 All snippets from the quotations discussed below.
semblance of “news”, even as it becomes detached from reality? And is the emergence of fake news necessarily tied to the malicious intent of individual human agents (a question that may, for example, have important legal ramifications)? While no definition, by itself, can reasonably be expected immediately to provide answers to these, and to a motley of related questions, a good definition should certainly lend itself to their exploration and, in this sense, be fruitful.

One can glean a good sense of the complexity of the term ‘fake news’ in actual usage from this list of (selected) tentative definitions brainstormed by lay participants in a recent media literacy workshop:

“News that contains false or inaccurate information”, “News with an agenda”, “News that is stretched in one way or another/tailored to a certain audience rather than raw facts”, “Clickbait material created without regard for actual true content” (Chapman 2017).

The actual list of participants’ suggestions is much longer, yet instead of listing it in full—or sampling a random number of characterization found in academic, journalistic, and other publications on the topic—it is considerably more insightful to group extant definitions into thematically related clusters.

The first such cluster concerns the medium in which fake news appears and is circulated. Unsurprisingly, the internet features prominently in such discussions, and some authors take its online format to be constitutive of fake news. Thus, lawyers David O. Klein and Joshua R. Wueller, in the Journal of Internet Law, define ‘fake news’ as

the online publication of intentionally or knowingly false statements of fact. (2017, p. 6)

They are, however, willing to consider exemptions for “well-known satirical Web sites such as the Onion, which uses humor and exaggeration to criticize social and political issues”
(2017, p. 6). To be sure, the internet plays an important role in the recent spate of fake news, both as a tool for its dissemination and as a convenient way of dismissing legitimate news as fake (as demonstrated by Donald Trump’s use of Twitter to denounce established news sources). Yet, if fake news is predominantly associated with online social media, then this is itself an *explanandum* and should not be built into the definition of what constitutes fake news. After all, a piece of online fake news does not suddenly cease to be fake news, just because it gets picked up and repeated on AM talk radio or makes its way into an op-ed piece by a newspaper columnist.

If the attempt to define fake news in terms of its medium fails, then perhaps focusing on the nature of its content is more promising. Specifically, what has attracted attention is the (sometimes extreme) disconnect between fake news and reality, as illustrated by the so-called “Pizzagate” episode, when multiple right-wing websites and talk-radio stations in the United States peddled a conspiracy theory that alleged collusion between prominent Democratic politicians (notably Hillary Clinton), pedophile criminals, and a number of popular restaurants, which were said to provide cover for the alleged culprits. This led to at least one shooting incident at one of the restaurants, perpetrated by a self-radicalized 28-year old man who claimed to be “self-investigating” the alleged conspiracy. Perhaps with an eye towards such knock-on effects of fake news stories, Roger Plothow, editor and publisher of the Idaho *Post Register*, in a column on media literacy, argues that,

*Fake news should be defined as a story invented entirely from thin air to entertain or mislead on purpose.* (Plothow 2017, p. A5)

This is echoed by economists Hunt Allcott and Matthew Gentzkow (2017, p. 5) who identify as fake news,
news stories that have no factual basis but are presented as news. (Allcott & Gentzkow 2017, p. 5)

Similarly, the organizers of the *Fake News Challenge*, a media literacy project, define fake news as

A completely fabricated claim or story created with an intention to deceive, often for a secondary gain. (Fake News Challenge 2017)

Most recently, Facebook, which at the end of 2016 had vowed to curb the spread of fake news on its social network, retired the use of the term ‘fake news’ in favour of ‘false news’:

The term “fake news” has taken on a life of its own. False news communicates more clearly what we’re describing: information that is designed to be confused with legitimate news, and is intentionally false. (Oremus 2017)

Any such attempt to reduce the phenomenon of fake news to just any sort of disconnect from reality is problematic. Defining fake news as “news that contains false or inaccurate information” is inadequate, given that even high-quality news sources will make the occasional mistake—yet an honest mistake regarding some irrelevant detail does not render the bulk of the reporting fake news. Facebook’s definition of ‘fake news’ as “false news” is hardly any better. To be sure, it does not equate false news with reports that are false *simpliciter*, but only with reports that are intentionally false, but this would include simple one-off lies as well as, perhaps more problematically, minor falsehoods that are the inevitable result of legitimate attempts to simplify complex matters in a way that makes them more accessible.5 Suggesting that fake news consists of completely fabricated claims and has no factual basis

---

5 Consider the example of a science journalist who describes the orbit of the planets around the Sun as elliptical, knowing full well that, strictly speaking, this is false (due to gravitational and relativistic effects).
likewise mischaracterizes the phenomenon in important ways: much of the initial credibility of fake news derives from real-world back-stories, and almost all fake news purports to be about real-world actors and entities. After all, fake news purports to be news, not fiction. Many fake news stories are not wholly false, but mix deliberate falsehoods with well-known truths as a means of obfuscation. The latter—deliberately misleading one’s audience—can be achieved even without reliance on falsehoods, simply through selective presentation of partial truths. This suggests that it is not fruitful to overstate the degree of disconnect that exists between fake news and reality, or in any case, that it should not be built into the definition.

Perhaps in order to pre-empt such criticisms, some contributors to the debate seem to conceive of fake news as a cluster concept that is best captured by a taxonomy of interrelated phenomena. Thus, Claire Wardle, director of research for the digital education non-profit FirstDraftNews suggests that fake news comes in seven distinct types of “mis- and disinformation”, including (from most to least harmful) fabricated content (“New content [that] is 100% false, designed to deceive and do harm”), false context (“when genuine content is shared with false contextual information”), false connection (e.g., headlines that don’t support the content), and satire/parody. Similarly, the media scholars Vian Bakir and Andrew McStay, drawing on Wardle’s typology, leave open just how much of a distortion and/or fabrication is required for a claim to count as fake news. Instead, they propose to define fake news in a disjunctive way, “as either wholly false or containing deliberately misleading elements incorporated within its content or context” (Bakir and McStay 2017, p. 1).

---

6 The classic example is St. Athanasius, who was trying to escape the Romans, and replied to a Roman officer who asked whether anyone had seen Athanasius: “He is not very far off.”

Such a definition, however, does not escape the criticisms above, against overstating the degree of fabrication and disconnect from reality. A similarly disjunctive view regarding the medium of dissemination is endorsed by Darren Lilleker (2017, p. 2), a professor in political communication, who argues that “fake news is the deliberate spread of misinformation, be it via traditional news media or through social media”.

Having eliminated medium of dissemination and mere lack of veracity as sufficient criteria for what constitutes fake news, two core issues remain: the deliberate manner in which false, unreliable, or otherwise deceptive information is being peddled by the purveyors of fake news, and the way in which the latter mimics the appearance and markers of credibility of established news sources. Thus, regarding the latter, the philosopher Neil Levy, emphasizing that his definition is only “intended to fix the reference for discussion, not serve as a set of necessary and sufficient conditions”, writes:

Fake news is the presentation of false claims that purport to be about the world in a format and with a content that resembles the format and content of legitimate media organisations. (Levy 2017, p. 20)

And Regina Rini, likewise one of the few professional philosophers who have written on the topic, offers what is perhaps the most ambitious and promising among the extant definitions:

A fake news story is one that purports to describe events in the real world, typically by mimicking the conventions of traditional media reportage, yet is known by its creators to be significantly false, and is transmitted with the two goals of being widely re-transmitted and of deceiving at least some of its audience. (Rini 2017: E45)

I have considerable sympathy for both of these definitions, not least since they go some way towards addressing the explanatory challenge of why fake news has become such a force in
the online world: never before has it been this easy to “mimic the conventions of traditional media reportage”, given that digital media have lowered the barrier to entering the information marketplace and that, furthermore, the sharing of articles on social media platforms such as Facebook removes many of the traditional markers of traditional journalism (including such visual markers as professional layout, an official masthead, comprehensive coverage, etc.). They also leave open the exact nature of the deception that renders fake news objectionable. For, as we shall see in the next section, purveyors of fake news need not intend to deceive their audience by getting them to believe the claim in question.

Yet, I believe that further improvements in our conceptual understanding of fake news are possible. This is entirely in keeping with the (previously mentioned) spirit of constructively engaging with current usage and extant definitions, with the goal of delineating more precisely why there is a need for the novel concept of fake news in the first place. Similarity in appearance to legitimate news media is certainly an important feature of fake news aimed at disseminating false or unreliable claims, yet it does not alone suffice as a demarcation criterion. Even when combined with the twin goals of achieving wide circulation and deceiving some of its audience, more is required for an item of disinformation to count as fake news; namely, a certain measure of success in realizing the goal of widespread circulation and uptake.7 In particular, I shall argue

---

7 I recall an episode during my time as a physics student, when a mentally disturbed man would hog the photocopier in the library, making access to it conditional on the purchase of a copy of his semi-professionally produced pamphlet denying Einstein’s relativity theory. Even if he had succeeded in mimicking the look and feel of a professional journal article even more convincingly, this would hardly constitute the scientific equivalent of fake news, since the manner of delivery undercut any chance of achieving wider circulation.
that what is distinctive about the wide circulation of fake news and its deceptive character is not that they are the result of individual intent, but that they are due to systemic features inherent in the design of the sources and mechanisms that give rise to them. In other words, for a claim to be considered fake news, it must in fact mislead a relevant audience—though precisely how large an audience may depend on the case at hand—and it must do so in virtue of the way it is designed to pass itself off as news (at least to the relevant target audience). The mere goal of circulation and deception on the part of its originator, even when combined with the mimicking of traditional conventions governing the news media, does not by itself suffice.

4. Fake news: A stipulative definition

Fake news is a species of disinformation. The term ‘disinformation’, more so than the relatively recent expression ‘fake news’, has by now received considerable attention from epistemologists and has been subjected to extensive conceptual analysis. Like ‘fake news’, ‘disinformation’ derives from a prior, philosophically more “respectable” notion: the notion of information, which in recent years has led to a burgeoning literature in the philosophy of information (Floridi 2011). Because of this structural similarity, and because paradigmatic cases of fake news may themselves be understood as instances of disinformation, we can expect the analyses of both concepts to run in parallel, at least up to a point. As we shall see shortly, however, there are also important specific differences, just as one would expect from phenomena that, though partially overlapping, nevertheless are recognizably distinct.

First, it must be granted that, just as disinformation is a species of information, fake news is, in a sense that needs to be spelled out, a form of news. This claim is by no means uncontentious, since it might appear to create a false equivalence
between epistemically “high-grade” and “low-grade” entities, so to speak. The very suggestion that disinformation is a species of information has been met with forceful criticism. As Fred Dretske puts it, “false information, misinformation, and (grimace!) disinformation are not varieties of information—any more than a decoy duck is a kind of duck” (1981, p. 57).

The problem, of course, is that if something looks like a duck, swims like a duck, and quacks like a duck, it is awfully difficult to recognize that it is not, in fact, a duck. Perspective matters. For the recipient of a piece of disinformation, or someone who is confronted with an instance of fake news, it does little good to be told that they should only accept what they are told if, in fact, it meets the requisite criteria of veracity and truthfulness. To be sure, there may be an objective fact of the matter whether the purported information or news items represents reality, but from the perspective of the recipient, this relationship is epistemically inscrutable. Therefore, building truth and veracity into the very definitions of ‘information’ and ‘news’—in other words, making them success terms—does little to address the pressing epistemological problem: how to respond to claims presented to us as true by a putative news source, given that, for all we know, they might (or might not) be fake news.

Once we grant that fake news falls under the same category of claims presented to us as true by a putative news source as real news, we can then turn our attention to the glaring differences between them. For one, fake news is misleading, in much the same way that disinformation is misleading: it is likely to create false beliefs” (Fallis 2015, p. 406). As discussed earlier, fake news may sometimes be fabricated from thin air, in which case it can at best be accidentally true. More often than not, it is built explicitly around falsehoods—especially claims that, if true, would be sensational—which it promotes and perpetuates. For a claim, in general, to count as

misleading in this way—that is, for it to count as likely to create false beliefs—it does not in fact have to mislead anyone. Perhaps a given claim, as a matter of chance, happens to be encountered only by an especially critical subset of reasoners, none of whom is taken in by it, even though the average person would have easily been fooled. Yet, arguably, what matters in the case of fake news, and gives urgency to it as a socio-political phenomenon, is that sufficiently large numbers of people are in fact taken in by it. What matters, then, is that, all else being equal, and taking into account general background conditions such as overall levels of media literacy, a piece of fake news that is released is likely to result in (and often does cause) false beliefs on the part of its target audience.

The two features identified thus far—being presented as news, and being likely to mislead its target audience by bringing about false beliefs in them—are not yet sufficient to demarcate fake news from, say, merely accidentally false reports. Even the most carefully vetted editorial process cannot entirely avoid the occasional mistake, and cases abound where respectable news sources, such as the Chicago Daily Tribune in its false announcement (“Dewey Defeats Truman”) of the winner of the 1948 U.S. presidential election, have made glaring mistakes. Yet it would not be correct to consider such accidentally false reports fake news. Such reports do mislead their audiences by instilling false beliefs in them, but they do so as the result of an unforeseen defect in the usually reliable process of news production. Fake news, by contrast, is misleading its target audience in a non-accidental way.

What about cases such as The Onion, the (now defunct) Colbert Report, or the British satirical current affairs magazine The Private Eye, the purveyors of the original “fake news” (in the now dated sense of media formats that aim to educate consumers by satirizing the shortcomings of mainstream news sources)? Great care goes into creating spoof reports that bear
many of the hallmarks of traditional reporting, yet have little or no relation to any actual events. Most consumers will, of course, treat them for what they are: as instances of entertainment with a sharp edge of social criticism. But if someone with no awareness of the satirical context were to take any of the fake reports at face value, as Iran’s Fars news agency did in 2012, when it relayed with all seriousness a report by The Onion that rural white Americans would rather vote for Iranian President Mahmoud Ahmadinejad than Barack Obama (BBC News 2012), they would certainly be misled, and not by accident. Yet this does not render The Onion’s original report an instance of fake news (in the revised contemporary sense at issue in this paper). Instead, the latter remains an instance of political satire—not the most subtle one, but apparently one that is still too subtle for some people who are misled by it.

For a putative report to count as fake news, it must be likely to mislead not only in a non-accidental way, but deliberately. As we shall see in a moment, however, this is still not the final word on fake news, considered as a species of disinformation, and it still does not shed much light on what, if anything, is novel about the recent phenomenon of fake news that has prompted our conceptual inquiry in the first place. Why do we need an even more fine-grained characterization of fake news than the one provided thus far? For one thing, it is important to get clearer about what we mean when we say that fake news misleads deliberately. Clearly, there is the (plausible) background assumption that, at some point in the process of creating fake news, human intention is involved: Someone, somewhere, decided to manufacture the claim in question and circulate it. This need not be the proximate source from which we received the claim in question—fake news items do sometimes get picked up by unsuspecting intermediaries, who then pass them on in good faith—but will likely be someone who seeks
to achieve a secondary goal by misleading others into believing the claim in question. What kind of goal might this be?

One might expect the ulterior motive of those who intentionally create fake news to be closely connected to the content of the claims they are manufacturing. Indeed, this has been a common refrain in the wake of the 2016 U.S. presidential election: Fake news, much of it baseless allegations of various transgressions attributed to the Democratic candidate, Hillary Clinton, was thought, at the time, to aim at discrediting her as a political figure. Yet, it is not clear that all of those involved in creating malicious fake news concerning Hillary Clinton had a political stake in the outcome of the presidential race, or were trying to sway the minds of potential voters. Many of those who, during the presidential race, had set up websites spreading fake news about Hillary Clinton (e.g., WorldPolitics.com, USConservativeToday.com, or USADailyPolitics.com) were teenagers based in the Republic of Macedonia, whose main goal was to generate revenue from online ads shown next to fake news articles that served merely as clickbait (Silverman and Alexander 2016). To the extent that they wanted their audience to believe the falsehoods they were spreading, they did so only in the hope that this would lead their audience to share the fake news items on social media, increasing the amount of internet traffic for their websites and thereby their revenue.

The example of the role of Macedonian clickbait farms in the propagation of fake news ahead of the 2016 U.S. presidential election suggests that the deliberate nature of fake news does not necessarily consist in the intention to manipulate others by instilling specific false (or malicious) beliefs in them. Rather, what matters is that the purveyors of fake news deliberately engage in practices that they know, or can reasonably foresee, to lead to the likely formation of false beliefs on the part of their audience, irrespective of whether they themselves
have a stake in those beliefs (as a political activist might have), or whether they are just in it for the money (as the Macedonian website operators were). Importantly, the spread of false beliefs is not merely a side effect of fake news, but is a direct result of its function. Mainstream news sources make mistakes, and the less reliable ones among them make more mistakes. The reports that result can mislead audiences into forming false beliefs, but even a relatively high error rate of a news source is compatible with its function of conveying the truth. This is why, after a series of high-profile mistakes, traditional news sources often seek to make changes to their editorial or vetting process. By contrast, fake news is designed to operate in a way that is unconstrained by the truth, either because it aims to instil falsehoods in its target audience (for example, in order to discredit a political opponent), or because the way it is deliberately operated is objectively likely to mislead its target audience, its real goal being (for example) the generation of clickbait through sensational claims that attract an online audience.

Drawing on the distinctions made thus far, we are now in a position to propose a first stipulative definition of fake news which captures most of its distinctive features:

**(FN)** Fake news is the deliberate presentation of (typically) false or misleading claims as news, where the claims are misleading by design.

---

8 Rini (2017: E45) makes a similar observation when she notes that the entrepreneurial Macedonians “did not care whether anyone ended up believing their fake news, so long as the clicks kept coming.” This is why I applauded Rini earlier (Section 3) for leaving open the exact nature of the deception that renders fake news objectionable. In the case at hand, the deception lies not in getting an audience to believe a false claim, but in getting them to believe it is worth sharing.
The phrase ‘by design’ is intended to reflect that what is novel about fake news—not only, but especially on electronic social media—is its systemic dimension. Individual intentions on the part of its originators (e.g., the intention to deceive an audience, to manipulate public opinion, to increase the circulation of certain pieces of disinformation) have a role to play, not least when it comes to settling the legal question of who should be held responsible for, say, slanderous content. But it is systemic features inherent in the design of the sources and channels through which fake news proliferates that imbue it with its novel significance. The next section will unpack this claim, along with the proposed definition, first by considering the various components of (FN) and explaining how each part contributes to the project of delineating fake news from related phenomena of public disinformation; second, by specifying a set of mechanisms that are typically being exploited in the creation and propagation of fake news. This will further shed light on precisely what is meant by the somewhat flexible phrase ‘by design’, and also holds out the promise of explaining why fake news has recently become a serious problem on electronic social media.

5. Fake news in the Trump age: Exploiting cognitive biases

Two questions need to be considered in relation to the proposed definition (FN): (1) To what extent does it succeed in distinguishing fake news from related, but distinct, types of public disinformation? (2) How well does it fare in explaining the recent rise of fake news, both in the run-up to the 2016 U.S. presidential election and, more generally, on electronic social media? I shall discuss each question in turn.

Misleading claims that are likely to give rise to false beliefs find their way into the public sphere in a variety of ways. A poorly-worded headline of an otherwise accurate article, or an
honest factual mistake, can mislead in this way, yet they do not constitute fake news. (FN) reflects this by reserving the label of ‘fake news’ for (certain cases of) deliberate presentations of false or misleading claims. Not all deliberate presentations of false or misleading claims, however, are instances of fake news. Consider false advertising, which may systematically misrepresent the facts by exaggerating the benefits of a particular product. Such claims are misleading and may get wide circulation, but as long as they are not presented in a format that could be easily mistaken for a news report (that is, as long as they are not presented as news), they do not constitute fake news. Even chronically sloppy journalism, which may result in a large number of false or misleading claims being presented as news, does not in and of itself qualify as fake news—unless, that is, the process of news production is so flawed (and is recognized as flawed by those who would be in a position to make adequate improvements, but fail to do so) that it can legitimately be regarded as being misleading by design. Finally, (FN) does not demand that the purveyors of fake news must always have an ideological agenda: fake news is not the same as propaganda. Instead, (FN) insists that the misleading nature of fake news claims is the result of systemic features of the way content is chosen or presented—in other words, that it is due to the way the process of generation and propagation is designed.

This brings me to my final point. To the extent that the very recent upsurge of fake news, especially in the run-up to the U.S. presidential election in 2016, is itself an explanandum and one of the main motivations for inquiring into the concept of

---

9 This is why the trend towards advertorials (advertisements that give pseudo-objective information about a product in a style that mimics the editorial content of the newspaper or magazine that carries the advertisement) is especially pernicious, since it chips away at the perceived authority of legitimate news sources.
fake news, any prospective definition should be compatible with, and ideally contribute to, an account of what causes fake news to become more prevalent. We should not, of course, expect a *definition* to be able to do the job of explaining the contingent ebb and flow of fake news in the public sphere, but it will certainly count in favour of a proposed definition if it has sufficient structure to allow us to pinpoint conditions for the emergence of fake news in a given set of circumstances. (FN), I argue, is apt in precisely this way. The key to understanding why, lies in the phrase ‘by design’. Recall that (FN) states that, for a claim presented as news to count as an instance of fake news, it must not only be misleading *simpliciter*, but misleading *by design*. That is to say, the originator of an instance of fake news either intends a specific claim to be misleading in virtue of its specific content, or deliberately deploys a process of news production and presentation that is designed to result in false or misleading claims. What makes contemporary fake news a novel phenomenon, and gives it its significance, is the extent to which systemic features inherent in the design of the sources and channels through which fake news is disseminated ensure its proliferation.

The latter, I wish to suggest, adds a sense of urgency to the recent spate of fake news, as propagated especially on electronic social media. In particular, purveyors of fake news have begun to employ strategies of bringing about belief and ensuring continued propagation of their stories (e.g., via “sharing” on social media), by manipulating their consumers’ pre-existing cognitive biases and heuristics. Perhaps the most prominent example is *confirmation bias*, the near-universal tendency to favour new evidence that confirms our existing beliefs or theories. Whereas confirmation bias has received some attention—both in relation to the phenomenon of fake news (Munchau 2017) and to the equally important problem of polarization (e.g., through “filter bubbles”, Pariser 2011), other
Biases are no less significant. Repetition effects entail that repeated exposure to the same information—even when the repeat message comes from the same source and so, from a rational point of view, is redundant—renders that information more persuasive; priming (“Lying Ted”, “Crooked Hillary”) can influence the interpretation of subsequent information to the point of inverting its meaning; affective arousal intensifies partisan bias and evaluative judgments, thereby reducing the overall willingness to negotiate or compromise, and shutting down pathways for rebutting false claims (see, e.g., Berger 2011, Brown and Curhan 2013, Petersen et al. 2015). Even the tactical co-opting of the phrase ‘fake news’ by those in power who wish to dismiss legitimate, albeit negative news coverage (cf. Dentith 2017), has its place in this taxonomy. By whipping up partisanship through the rhetorical pairing of fake news allegations with emotive language (“Fake News CNN made a vicious and purposeful mistake yesterday”, Donald Trump tweeted on 9 December 2017), it encourages the fallacy of poisoning the well—all the more when repetition makes such a linkage readily available to informal reasoning.\(^\text{10}\) Levy notes that “even when we succeed in consuming fake news without coming to acquire beliefs that directly reflect its content”, the representations it draws on “will play a content-reflecting role in our further cognition, in ways that may be pernicious” (2017: 32). What I have tried to argue in this section is that we need to go even further. Far from being a side effect of fake news, the systemic ways in which fake news mobilizes our cognitive biases and heuristics, thereby modulating our reasoning processes, is part and parcel of the function of (contemporary, esp. online) fake news. Confirmation bias, repetition

\(^{10}\) Walton aptly characterizes “poisoning the well” as “a tactic to silence an opponent violating her right to put forward arguments on an issue both parties have agreed to discuss at the confrontation stage of a critical discussion” (2006, p. 273).
effects, affective arousal, and related cognitive biases thus feed into a range of informal fallacies that inhibit critical reasoning, and the systemic features inherent in the design of the sources and channels through which fake news proliferates create additional obstacles to critical reasoning and active inquiry. If this list of factors, which is far from complete, sounds apt in relation to fake news, it is because much of fake news deploys representational means that actively modulate consumers’ cognitive biases, in an attempt to manipulate consumers into propagating (e.g., through online “sharing”) the very claims that misled them in the first place.

Fake news is not itself a new phenomenon. Yet, when combined with online social media that enable the targeted, audience-specific manipulation of cognitive biases and heuristics, it forms a potent—and, as the events of 2016 show, politically explosive—mix. In terms of our definition (FN), online social media, which, as a *Psychology Today* article puts it, work on cognitive biases “like steroids” (Braucher 2016), has opened up new systemic ways of presenting consumers with news-like claims that are misleading by design. As a result, given the increasing permeability between online and offline news sources, and with traditional news media often reporting on fake news in order to debunk it (a worthy goal that is rendered ineffective by further cognitive biases such as source confusion, belief perseverance, and the backfire effect), we find ourselves increasingly confronted with publicly disseminated disinformation that masquerades as news, yet whose main purpose it is to feed off our cognitive biases in order to ensure its own continued production and reproduction.
References


Goldberg, Sanford C. 2011. “If that were true I would have heard about it by now.” In Social epistemology: Essential readings, eds. Alvin I. Goldman and Dennis Whitcomb, 92-108. New York. Oxford University Press.


Rosen, Christopher. 2017. All the times Donald Trump has called the media ‘fake news’ on Twitter. URL accessed 30 August 2017: <http://ew.com/tv/2017/06/27/donald-trump-fake-news-twitter/>


