SCM
Studies in Communication and Media

FULL PAPER

Stop studying "fake news" (we can still fight against disinformation in the media)

Schluss mit der Forschung zu „Fake News" (wir können trotzdem weiter gegen Desinformation in den Medien kämpfen)

Benjamin Krämer

 Studies in Communication and Media, 10. Jg., 1/2021, S. 6–30, DOI: 10.5771/2192-4007-2021-1-6

Benjamin Krämer (PD Dr.), Ludwig-Maximilians-Universität München, Institut für Kommunikationswissenschaft und Medienforschung, Oettingenstr. 67, 80538 München, Germany. Contact: kraemer(at)ifkw.lmu.de. ORCID: https://orcid.org/0000-0002-1351-2414

FULL PAPER

Stop studying "fake news" (we can still fight against disinformation in the media)

Schluss mit der Forschung zu „Fake News" (wir können trotzdem weiter gegen Desinformation in den Medien kämpfen)

Benjamin Krämer

Abstract: The problem of "fake news" has received considerable attention both in public discourse and in scholarship. However, many have argued that the term should be avoided for ideological reasons or because it lacks clarity. At the same time, a growing body of literature investigates "fake news" empirically. We complement this discussion by reflecting on epistemological and methodological problems with the term "fake news" and the implications for possible solutions to the problem of disinformation such as automatic detection and increased media literacy. Based on the principle of symmetry established in the sociology of scientific knowledge, we show that a classification of messages according to the researcher's assessment of their truthfulness can lead to biased or tautological explanations. We argue that many researchers commit themselves to the truth or falsehood of messages in cases where they should not and avoid such a commitment when it is necessary.

Keywords: Fake News, principle of symmetry, automatic recognition of disinformation, media literacy.

Zusammenfassung: Das Problem der „Fake News" hat in der öffentlichen Diskussion und in der Forschung viel Beachtung gefunden. Es wurde jedoch vielfach argumentiert, dass der Begriff aus ideologischen Gründen oder mangels einer klaren Bedeutung vermieden werden sollte. Zugleich untersucht eine wachsende Zahl von Publikationen „Fake News" empirisch. Wir wollen zu dieser Diskussion beitragen, indem wir die epistemologischen und methodischen Probleme mit dem Begriff „Fake News" sowie ihre Implikationen für mögliche Lösungen wie automatische Erkennung und verbesserte Medienkompetenz reflektieren. Auf der Grundlage des Symmetrieprinzips aus der Wissenschaftssoziologie zeigen wir, dass es zu verzerrten oder tautologischen Erklärungen führen kann, wenn Forschende Botschaften nach dem Wahrheitsgehalt einteilen, den sie ihnen zuschreiben. Wir argumentieren, dass sich viele Forschende auf die Wahrheit oder Falschheit von Botschaften festlegen, wenn sie es nicht sollten, und eine solche Festlegung vermeiden, wenn sie notwendig wäre.

Schlagwörter: Fake News, Symmetrieprinzip, automatische Erkennung von Desinformation, Medienkompetenz.

Stop studying "fake news" (We can still fight against disinformation in the media)

We all know the media narratives about heroic researchers who, due to a coincidence of luck and genius, discover some important truth that they successfully defend against their opponents and against authorities who wish to suppress such enlightenment. This is not too different from certain narratives that are still being related in the history of science. However, scholars in the history and sociology of science have warned against such perspectives. If we exclusively focus on scientific theories that have come to be accepted as true, our notion of how science produces knowledge will be biased. We may even fail to explain anything at all if we only tell the story of how the truth ultimately had to reveal itself or if we think that only false beliefs require explanation.

And we all know the reports about some sinister propagandist who, due to the coincidence of a heated political climate and manipulative skills, is able to create exceedingly deceptive messages disguised as news items which they successfully spread among gullible social media users, defying all attempts by the established media to set the record straight and to get through to citizens with fact checks and truthful news. Would it be as problematic for researchers to base their work on such an anti-heroic narrative as it would be to study the history of science based on heroic narratives? We would argue that, in a certain way and apart from many important differences, the underlying problems are similar. In this article, we would therefore like to critically reconsider research on "fake news" – at least as it follows a certain logic – in the light of a principle that has been established in science and technology studies in response to the flaws of some narratives of scientific progress: the principle of symmetry. We will use it to reflect on epistemological and methodological problems with the term "fake news."

We start by shortly considering in what way many researchers have approached "fake news" so far, why, according to some, the term should (not) be used, and how our own analysis relates to these existing lines of research and positions. We then present a categorization of messages with regard to how different actors (researchers, communicators and recipients) judge their truth value and situate "fake news" in this typology. From the discussion of this categorization, we conclude that it is the perspective of the researcher that often dominates the classification of messages. In the next chapter, we argue why it is problematic if researchers define their object of study as messages that are, in their own eyes, false, or if they compare such messages with those we consider to be true. Such a perspective violates the principle of symmetry which we will define in that section. We illustrate the problems arising from research that violates that principle and provide examples of "symmetric" approaches. As we will show, the logic behind the criticism of asymmetry in empirical research of "fake news" can also lead us to take a critical view on two main solutions that have been proposed to fight against fake news: the automatic detection of disinformation and increased media literacy (and calls to re-establish epistemic authorities).

1. The interest in "fake news"

In the recent scholarly, scientific, and public discussions of "fake news," we can identify a number of main interests: 1) the amount of "fake news" that people are exposed to (e.g. Alcott & Gentzkow, 2017; Guess, Nyhan, & Reifler, 2018), 2) to understand how people process and why they believe and share certain questionable messages and the resulting patterns of diffusion (e.g. Allcott, Gentzkow, & Yu, 2019; Del Vicario et al., 2016; Levy, 2017; Lewandowsky & Ecker, 2017; Pennycook & Rand, 2017; Shin, Jian, Driscoll, & Bar, 2018; Swire, Berinsky, Lewandowsky, & Ecker, 2017; Tandoc, Lim, & Ling, 2020; Vosoughi, Roy, & Aral, 2018; Zollo & Quattrociocchi, 2018), 3) how people react to fact checking (see Chan, Jones, Hall Jamieson, & Albarracín, 2017, and Walter, Cohen, Holbert, & Morag, 2019, for meta-analyses on debunking), 4) how to contain the perceived flood of fake news by, among other strategies, detecting them automatically (e.g. Altunbey Özbay & Alatas, 2020; Saquete, Tomás, Moreda, Martínez-Barco & Palomar, 2020; Zhang & Ghorbani, 2020, for systematic overviews) and 5) the social causes and consequences as well as the normative implications (Ball, 2020; Lewandowsky, Ecker, & Cook, 2017; Quinn, 2017; Rini, 2017. For an overview on these and other topics, see also Lazer et al., 2018; Müller & Denner, 2018; Tucker et al., 2018). A relatively large amount of literature has been produced in a short period of time in order to address these important concerns revolving "around what is or is not considered true and is able to serve as the basis for public discussion" (Jankowski, 2018, p. 249). The aim of this article is to take a step back and to reflect on the assumptions underlying these approaches. We argue that, somewhat paradoxically, many researchers commit themselves to the truth or falsehood of messages in cases where they should not and avoid the complexities of such a commitment when it is necessary.

There may be reasons not to use the term "fake news" that one could consider to be legitimate, such as its ambiguity (it has been used to describe parodies of news, satire, deliberately deceptive messages and simple errors, advertising and public relations messages disguised as news, etc.; Perez-Rosas, Kleinberg, Lefevre, & Mihalcea, 2017; Tandoc, Lim, & Ling, 2018; Wardle, 2018) or its use as a political invective against journalism (e.g., Dentith, 2017; Zimmermann & Kohring, 2018). Habgood-Coote (2019) also dismisses the term on the grounds that it does not have a stable descriptive but only an evaluative meaning. While one might object that political concepts are usually contested and that vague terms can be defined (as the author also acknowledges), we may agree that the term is ideologically problematic and rather redundant with regard to earlier concepts. The new term suggests that we have suddenly entered an era of post-truth (for different critical perspectives on this claim, see, e.g., Farkas & Shou, 2018; Neuberger et al., 2019; Sismondo, 2017) and that we can return to a golden age of truth if we appeal to individual responsibility and good leadership, learn to trust established institutions again, and develop new technical solutions – instead of more radical reforms and instead of acknowledging social diversity (Habgood-Coote, 2019).

In the following, we will stay with the eye-catching label "fake news" because even if the phenomenon is labeled differently, the problems with the underlying concept that we would like to address remain the same: epistemological or methodological problems that arise from an approach that begins with "fake news" as a category, defined as news or news-like messages that are (knowingly) false (or what Zimmermann & Kohring, 2018 call "aktuelle Disinformation [topical disinformation]", i.e., "wissentlich und empirisch falsche Informationen zu neuen und relevanten Sachverhalten mit dem Anspruch auf Wahrheit [untruthful and empirically false information regarding novel and relevant issues that are claimed to be true, translations by Z & K]" which, however, they do not limit to messages that closely mimic journalistic formats)[1] and intentionally and systematically misleading (in particular by imitating real news) but that may be consumed and proliferated by people who believe them to be true (Gelfert, 2018).

2. The concept of "fake news" and perspectives on the truth value of messages

We can create a systematic typology of messages based on the relationship between the beliefs of different actors, including researchers (see Table 1). We can start with the judgment of communicators: Do they believe what they claim or not? We may further distinguish between deceptive messages in the proper sense and those that are intended to be recognized as false or fictional. The audience can be competent or incompetent to recognize such works of fiction or satire. This can sometimes pose a problem, but we certainly worry more about "fake news" in the narrow sense – deceptive messages – in particular if it is being believed by parts of the audience. There are two types of "fake news" in this sense: messages claiming that something one considers true is false, and those claiming that something one considers false is true. (This is perhaps the more prototypical case of "fake news": the fabricated report that invents its own truth.)

Two other types of problems can be identified: legitimate news that is not accepted by some recipients (including the debunking of false information), and er-

1 The distinction between "true" and "false" is of course a simplification. In reality, communicators, researchers and/or recipients may hold that some information is "probably true", "mostly true", "unconfirmed rumor" etc., and a news item itself may also mention such qualifications. Furthermore, messages cannot only be simply false but also otherwise misleading, e.g., by taking information out of context (Wardle, 2018; Zimmermann & Kohring, 2018). However, first, the present argument is about "fake news" and the field is often defined by these news items actually being untrue, and they are considered relevant just because they are false. Therefore, at least the category of "(certainly) false" is of particular interest here.
Second, more nuanced ascriptions of truth and falsehood do not substantially change our argument. Its main point will be that we should not depart from our judgment of truth, whatever it may be, when explaining others' judgments and behaviors. In the case of "fake news", this judgment happens to be that some information is considered false *by the researcher*. However, the argument would also apply to other cases, for example to studies of rumor or of different types of misleading messages. In such cases, we should not base our analysis on our own decision on whether some information is unconfirmed or misleading. If we commit ourselves to the judgment that one cannot know whether something is true (or that something is presented in a misleading manner), this is only our own attitude which cannot play a role when it comes to explaining others' judgments or the spread of what we classify this way. For the sake of simplicity, we will therefore use the simple distinction between "true" and "false" in the following.

rors by communicators, in particular, if they are accepted by the audience. Communicators may either mistakenly believe something to be true which is not, or *vice versa*.

The remaining cells in Table 1 represent the rare cases of ill-informed communicators trying to deceive recipients but actually describing reality in a way we as observers would accept. Their lies are "accidental truths" (Gelfert, 2018; Reed, 2000) that lead to recipients either holding accidentally true beliefs or ultimately wrong beliefs if they reject the message.

Looking at Table 1, we may make two observations. First, the researcher and their judgment sit at the very top of the table. It is no coincidence that we put them there: Research on "fake news" starts with a distinction made by researchers, dividing messages into those that they consider true and those they consider false. It is their perspective that informs the whole analysis. Even if we use different terms according to the subjective attitudes and intentions of communicators or the form of the message instead of categorizing everything as "fake news," this general perspective remains the same. For example, Shin, Jian, Driscoll, and Bar (2018) distinguish between disinformation and misinformation, trolling and fake news in the narrow sense depending on whether communicators believe the message to be true or false and whether they aim, for example, at others' reactions or financial gains or other advantages (Wardle's, 2018, definition of disinformation similarly locates it at the intersection of what is known to be false by communicators and what is communicated due to certain undesirable intentions, however describing the communicators' aims somewhat differently as "intent to harm"). However, all these concepts are then used to designate false information *according to the researcher* or, in the case of rumor information that is unconfirmed – but still from the perspective of the observer.

Second, we may identify a kind of symmetry in the table. For example, "fake news" and legitimate forms of news appear twice in diagonally opposed cells. It may even be concluded that the judgment of the researcher is irrelevant when it comes to explaining the interplay between communicators and their audience. Communicators decide whether to be truthful, and recipients decide whether to accept messages. And when recipients judge messages, the actual beliefs and intention of the communicators does not matter either and cannot explain that judgement. An external perspective can only identify *normative* problems if we would like to commit ourselves to the truth or falsehood of certain information, but the processes that require explanation instead of normative judgments by the researcher are represented in the third to eighth line of the table: how communicators decide what to believe and communicate and, independently from these decisions, how recipient evaluate messages. However, the typology itself does not provide such explanations (which are beyond the scope of this article).

In the sociology of science and (scientific) knowledge, symmetry is also an important principle when explaining the production or acquisition of knowledge and the establishment or acceptance of something as a fact (see, e.g., Sismondo, 2010, ch. 5, for an introduction). Below, we will define this principle and point to some pitfalls when studying "fake news" without following the principle of symmetry.

Table 1. Types of messages according to the truthfulness ascribed by different categories of actors

	Judgment of researcher							
	True				*False*			
	Judgment of communicator				**Judgment of communicator**			
	True		*False*		*True*		*False*	
	Claim by communicator				**Claim by communicator**			
Judgment of recipients	*True*	*False*	*True*	*False*	*True*	*False*	*True*	*False*
True	Legitimate news accepted by audience	"Fake news" rejected by audience	Accidentally true lie accepted by audience	Errors of communicators accepted by audience		Accidentally true lie rejected by audience	Successful deception by "fake news"	Unsuccessful debunking of "fake news"/ recipient unable to recognize satire
False	Legitimate news rejected by audience	Successful deception by "fake news"	Accidentally true lie rejected by audience	Errors of communicators rejected by audience		Accidentally true lie accepted by audience	"Fake news" rejected by audience	Successful debunking of "fake news"/ recipient able to recognize satire

3. The lack of symmetry

As mentioned above, "fake news" is also used as a political invective. One person's "fake news" can be another's truth, and *vice versa*. Researchers may also weigh in on such a controversy. But what can their judgment mean for (empirical) research? We would argue that even if we are concerned that parts of the population use and accept media messages that we consider false or misleading, it is not a good strategy to select case studies or samples based on our own beliefs on the correctness or falseness of these messages or to treat both types of messages differently in studies.

The principle of symmetry has been established in the sociology of scientific knowledge as part of the so-called Strong Program: Our belief in the correctness of a statement cannot explain others' belief – or disbelief (see Bloor, 1999, also on the following). Therefore, our belief is completely irrelevant when it comes to explaining why others do or do not believe something. The principle of symmetry requires using the same kinds of explanations for why people accept statements that we consider correct and for why they accept statements we believe to be false. "The Strong Program [discourages] the analyst from dividing agents into two evaluative categories, namely those who subscribe to what we take to be true beliefs on some subject matter, and those who don't" (Bloor, 1991, p. 105). This principle demands that when developing explanations, we bracket our beliefs about what is true, but, of course, it does not require us to abandon them altogether.

The principle of symmetry as such only applies to cognitive beliefs and we certainly do not have to give up our normative convictions that misinformation is a problem and should be fought against, and that, consequently, it is important to understand why different types of messages (including those that communicators or we believe to be false) are spread and believed. We only argue that, in doing so, we should follow the principle of symmetry which we will now explicate in more detail.

In the history and sociology of scientific knowledge, we are tempted to think that truth reveals itself, and therefore we look at an error as something that requires particular explanations. However, we must refer to the same sets of factors such as structural conditions, preexisting theories, practices of research, etc. in order to explain why both conclusions that are established as true today were reached *and* why others were not introduced by researchers or not accepted as true. We may find differences between accepted and rejected theories, but we cannot simply rely on our beliefs when developing our explanation or simply assume that true findings can be explained rationally and false ones by some "irrational" factors"the truth" as an explanation. Instead, we must refer, for example, to the fact that practices of knowledge production and standards of evaluation have evolved. We can then expect to find our own judgments of a theory to be correlated with the scientific practices used to establish and test it: We are generally more inclined to accept findings that are based on methodologies that we consider up-to-date. However, our own evaluation of an assumption cannot play a role in the explanation of how it came to be accepted (some may also have accepted the

theory for other reasons than those that convinced us). Conversely, it is not a contradiction to commit oneself to a paradigm and a set of theories and methods in order to conduct research on the production of knowledge and to use the same approach to explain why our approach is accepted (at least among some researchers) (an approach can be a topic and a resource at the same time; Bloor, 1999).

For example, the theory of relativity is generally accepted as valid in the scientific field, whereas older theories of the aether have been discarded. If we seek to explain these developments, we cannot simply assume that truth ultimately reveals itself and that only false beliefs require explanations. This reasoning is tautological (things are considered true because they are true) or, more precisely, does not really explain anything. It runs as follows: "A theory that we consider to be true will ultimately be accepted as true (because, after all, it is true) – unless it is not (yet) accepted as true for some reason that we have to find out." Instead, we must investigate the social backgrounds, institutional positions, interests, pre-existing worldviews and epistemological standards, research practices, arguments, rhetorical strategies, etc. of the advocates of both theories and try to explain why they and others came to be convinced and to find out what were the most important factors in this decision. For example, different critics of the theory of relativity were either dissatisfied with its mathematical form or metaphysical implications, interpreted certain experiments differently, or rejected it because it was "Jewish" (there is a vast literature on the reception of the theory; see, e.g., Brush, 1999; Glick, 1987; Wazeck, 2014). Similarly, the advocates of the theory had to make sense of its formulas, accept experimental practices and results as evidence based on their practical knowledge on how to conduct "good" experiments, their epistemological or ideological beliefs, etc. The principle of symmetry requires the same type of explanation for all kinds of assumptions (whether we consider them true or false), although, of course, different factors may be decisive in different cases. In contrast, our own belief or disbelief in a theory or the attitude of an epistemic community we belong to or trust cannot be an explanation for others' beliefs.

The case of "fake news" is somewhat different, but the principle of symmetry has important implications here as well. We are certainly not tempted so much to assume that truth will prevail by itself. However, we also tend to introduce our own beliefs as a factor in explanations or as a criterion of sampling in our research.

One may intuitively grasp what the lack of symmetry means for research on "fake news" if one imagines that whenever someone uses the term "fake news" affirmatively, they actually mean "news I believe to be false/misleading" (as good as their reasons for this belief may be). If our object of investigation is simply "messages we consider incorrect" or if we compare (in our view) "true" and "false" messages, we risk missing important commonalities and differences between types of media content and different users' judgments of the truth value of messages.

Another way to illustrate the problem is to imagine a research project that investigated some factors that were assumed to explain why certain "fake news" were shared and believed. Then imagine that the researchers later discover that those messages were in fact true. Would they have to change their explanations or

withdraw their publication, or wouldn't we still accept their findings as valid explanations of why these messages were spread and considered true?

Shin et al. (2018) refer to an instance where mainstream news outlets had to apologize for spreading false claims. This illustrates the pitfalls of explaining diffusion by the ascribed truth value of claims. Our explanation of what happened at a point in time in a presidential campaign several years ago cannot depend on whether claims have since been confirmed or debunked.

If we compare the propagation and reception of "fake news" to legitimate news (and also draw our sample accordingly), we may risk misattributing the reasons for recipients' beliefs and disbeliefs and for why ordinary users or different types of media outlets share a story. Or we would even commit a category mistake if we explained others' beliefs and behavior mostly with our own beliefs. Certainly, our judgment may correlate with others' judgment and actions. However, the reason for this is that we evaluate news in certain ways and others may use the same or different standards, referring to the same or different properties of sources and messages and the same or different bodies of knowledge to compare a news item with. Even if we include such factors, we may estimate their influence incorrectly if we continue to (explicitly or implicitly) use our own evaluations as explanations as well. And, in the most extreme case, if we do not include any other factors beyond our own beliefs, we cannot provide a real explanation at all.

In the following, we will offer a rather uncharitable reading of some studies (hoping that the authors will forgive us for using their work in such a way) and point to the problem that research on fake news may invite such interpretations. Afterwards, we will suggest a more charitable and, at the same time, more symmetric and thus epistemologically justifiable interpretation.

For example, Shin et al. (2018) found that the number of tweets regarding rumors flagged as false by fact-checking organizations peaked several times, whereas rumors confirmed by those institutions peaked only once. However, this result as such only describes a highly relevant *problem* (from the perspective of those agreeing with the fact checkers, falsehoods are shared too often) or a mere correlation (between temporal patterns of sharing and the results of fact checking or the belief of the authors concerning the truth of the rumors). However, such a finding as such does not provide an *explanation*. An explanation would have to refer to, for example, the criteria of newsworthiness and the practices of selection in different parts of the public sphere: As the authors themselves note, each peak is usually associated with another, mostly non-mainstream, outlet or prominent Twitter account, and they explain this, for example, in terms of campaign tactics.

Pennycook and Rand (2019) find that subjects who are more prone to analytical reasoning believe "fake news" less often. They explain this result with the observation that "fake news" often makes implausible claims that contradict everyday knowledge and can therefore be identified by analytic thinking. However, the assertion that the messages are "implausible" is ambiguous: Are messages objectively implausible or implausible in the context of certain previous beliefs? The first interpretation may lead to a rather tautological explanation: Subjects identify "fake news" as false because it is indeed false (or at least implausible, i.e., easy to categorize as probably false).

Van Duyn and Collier (2019) analyze the effects of elite discourses on "fake news." Their experiment shows that participants who have read tweets by elite persons addressing "fake news" were able to identify real news with less accuracy and did not recognize "fake news" easily. However, if the authors find an effect for one category of news but not the other, how can this be explained? After all, these categories reflect the researchers' assessment of the veracity of the messages, while the reader's judgment is the dependent variable. Therefore, the researchers' categories do not have equivalents on the part of the participants and their reactions must be explained otherwise.

Finally, studies on "debunking" actually only investigate the effects of contradictory information based on a sample of initial messages that researchers consider false. Here, findings risk being biased because researchers' judgments may be contingent on the typical content and formal features of the messages. If the messages used in such studies exhibit typical properties, those properties may then determine the typical outcome of such analyses. Recipients simply must process the first message and the message debunking it and make a judgment. Their evaluation is logically independent of which message the researchers consider true. Instead, participants' judgments will be based on the concrete properties and content of the messages interacting with the participants' prior attitudes.

The findings of such studies may be prone to a highly asymmetrical reading that essentializes the concept of "fake news" and leads to biased explanations that commit category mistakes. This interpretation goes as follows: There is this thing called "fake news." They are spread and believed more easily than real news – because that is the way it is with fake news. Or one might add that this is the case due to *the* properties of fake news, or that they are disbelieved because they are "implausible" (because, after all, they are false!).

A more charitable and symmetric reading (that some authors also offer, even if it is not always clearly separated from the essentializing reading) would be: Some messages with certain properties are being spread and believed more easily due to their characteristics and other factors; they are plausible to some because they fit their prior beliefs and worldviews, and implausible to others because… etc. And, what makes this particularly problematic and relevant, many of the messages with those properties happen to be false because in the present historical, cultural and political context, "fake news" tend to exhibit those features.

However, instead of classifying messages according to our own beliefs, we should categorize them with regard to other criteria that could explain people's judgments in general, such as theories on trust, worldviews and belief systems, folk epistemology, etc. (the truth and falsehood of certain types of messages may then add to the relevance of such research). A symmetrical approach also leads us to reflect on our own trust in certain media outlets and journalistic institutions and practices. What makes us believe that events we do not experience directly are covered correctly?

Some authors have already explored other lines of research bracketing judgments on the truthfulness of messages, instead focusing on discourses on "fake news" (or the use of "fake news" as a label as opposed to fake news as a type of

messages, Egelhofer et al., 2020; Egelhofer & Lecheler, 2019) and analyzing them at a meta-level.

Carlson (2018) describes "fake news" as a symbol in a moral panic in which the established media, protecting their authority and relevance, present themselves as trustworthy as opposed to deviant others. To Farkas and Shou (2018), "fake news" is not only a question of truth or falsehood, but a symbol in political struggles, a "floating signifier" that is being used to gain discursive hegemony. They identify three "moments" in this struggle: "(1) a critique of digital capitalism, (2) a critique of right-wing politics and media and (3) a critique of liberal and mainstream journalism. A fourth moment, which is not included in the analysis, includes mobilization of the 'fake news'-signifier as part of techno-deterministic critiques of digital media technologies (e.g. Facebook is bad for democracy)" (p. 303). The authors come to an interesting conclusion concerning the present diagnosis that we live in an era of "post-truth": Truth has not so much lost importance as we live in times in which everyone tries to obsessively define what truth is. Or, to put it otherwise: While there are certainly actors for whom truth does not matter in their strategies, the diagnosis of a "post-truth" era is tied to the egocentric position of one party in the struggle for truth which observes that its own vision of reality does not necessarily prevail.

While she also draws normative conclusions, Marres (2018) analyzes the conception of truth behind fact-checking functionalities on social media platforms and the social relations they imply without committing herself to the truth or falsehood of certain specific claims. According to her interpretation, this conception seems to imply a logical-positivist correspondence theory and a "demarcationist" understanding of truth: By segmenting reality into testable statements and elementary facts, a boundary between legitimate and illegitimate claims can be drawn. This distinction corresponds with a social one: although everyone would in principle be responsible for the correctness of their claims and beliefs, fact checkers must help those that are unable to fulfill this responsibility on their own. And those who are ultimately able to recognize the truth are set apart from those who cannot, which presently means "educated progressives [from] less educated supporters of populist and nationalist causes" (Marres, 2018, p. 430). While the sources of claims and individual users who share them are held responsible for what they spread, the design of social media platforms, their algorithmic selection, the tracking, targeting, and collective filtering are not related to ascriptions of truth value (Marres, 2018).

Although they depart from a strongly normative perspective that fake news, Iosifidis & Nicoli (2020) descriptively analyze Facebook's announcements of how it intends to combat disinformation. The authors commit themselves to the existence of a certain type of false messages but do not explain anything based on their falsehood.

4. The problem with (certain) countermeasures

The studies reviewed in the second part of the previous section demonstrate the analytical potential of a symmetrical approach towards "fake news." Does this agnostic approach inevitably lead to relativism and into the post-truth culture that they seem to critically analyze? It does not, because we can and should hold convictions about the truth and falsehood of important news items that we may only bracket in order to analyze discourses at another level. As noted above, the principle of symmetry does no longer apply when it comes to our normative convictions and to defining the goals and success criteria of measures against misinformation. However, developing ideas on how to fight against falsehoods holds its own, but related challenges as we will argue in the following. We discuss two possible solutions: automatic recognition of false messages and improving media literacy. At first sight, the following argument may seem unrelated to the above discussion based on the principle of symmetry. However, the logic behind our criticism of the two types of countermeasures is as follows.

Both when explaining others' judgements asymmetrically and when engaging in automatic recognition or media literacy education, we mostly rely on pre-established judgments on the truth and falsehood of messages: There are "fake news" which we know to be false. Thus, we immediately move to explain and combat them and fail to reflect on the basis and role of our judgements: In one case, that of explanations of the receptions and spread of misinformation, we should not commit ourselves too strongly to our own judgements but bracket them during the analysis. In the other case, that of countermeasures, we would have to commit ourselves even more strongly to our judgements and properly explicate our criteria for truth and falsehood or trust. How else can we be sure that the functioning of a technical system reflects our actual understanding of what makes a news item false and applies it correctly to new messages? And how can we be sure that others really learn how to assess media content if we are not sure what makes it true or false or when to trust it?

The argument from the principle of symmetry has shown that we may have our reasons for judging something to be true or false but that others have their own reasons which may or may not be the same as ours. It is this distinction between our reasons and others' reasons that is both at the basis of the principle of symmetry and of the following argument. If we want to implement countermeasures against disinformation which we consider to be normatively justifiable, they ultimately need to implement *our* criteria of truth as faithfully as possible – they have to be based on *our* actual reasons for believing something. Countermeasures we can endorse and implement cannot define truth based on criteria that are not ours, that are only sometimes congruent to ours by chance or that are only contingently related to ours. We will argue below that it is impossible for the criteria implied in the two types of countermeasures to be completely the same as ours but that they are mostly heuristic and thus only contingently related to ours, or at least incomplete: What automatic systems do to "recognize" false messages cannot match usual standards of truth we would commit ourselves to, and the knowledge and competences usually considered as part of media literacy are not suffici-

ent to allow recipients to judge news items in way that would satisfy usual standards of truth.

However, if we aim to communicate what we hold to be true, we have to develop strategies that start from people's own beliefs and criteria of truth, potentially challenging them. This can only be achieved if we investigate them symmetrically, by bracketing our own criteria.[2] Therefore, symmetry cannot play a role in the definition of aims of countermeasures and the way the countermeasures discussed below actually work is neither based on real symmetry nor on clear normative "asymmetrical" criteria, as we will show. But symmetry can play a role when it comes to the development of strategies to argue against falsehoods.

We cannot systematically discuss theories of truth in relation to potentially false news items, nor can we suggest better intervention against misinformation. This has to be left to other publications and authors. We will only make a number of rather commonsense arguments for why the criteria used in countermeasures cannot be satisfactory. However, the following argument does not imply the media literacy and automatic recognition are unable to provide important cues for when to be particularly skeptical. They simply cannot help us to *definitely* ascertain the truthfulness of messages.

4.1 Automatic recognition?

We have argued that some current approaches commit themselves to the truth or falsehood of information and use this categorization as an explanans instead of addressing the empirical reasons for these effects. Somewhat paradoxically, the current attempts to staunch the perceived flood of fake news by means of automatic detection use some of the criteria that might be part of the symmetrical approaches we call for: features of messages that may contribute to shaping users' judgments of their truth value. Many recipients do not have direct access to certain events and sources but must either trust media institutions or their peers to inform them correctly, or they have to rely on heuristic assessments of the messages based on certain cues. This is also the case for most approaches to the detection of "fake news."

At the same time, research on technical solutions usually does not provide a normative framework to justify commitments to the veracity of news items. These approaches usually do not commit themselves to a sufficiently clear standard of truth but use messages that have already been classified as true or false (Badia, 2020). Are the existing systems or will future systems ultimately be able to "fact check" messages and in what sense, e.g. in the same way human reporters and researchers do?

While technical systems may be successful in sorting out a considerable number of textbook cases of what is currently considered "fake news" (Zhou, Guan, Bhat, & Hsu, 2019), we should be extremely careful not to claim that those systems come anywhere close to what we might consider real fact checking. This

2 The author would like to thank an anonymous reviewer for suggesting this differentiation.

argument is not merely based on the rather low rates of success.[3] A universal fact-checking tool would be equivalent to general artificial intelligence, a system that is not restricted to processing formal properties in a restricted field. Instead, information on any social or natural phenomenon might be relevant when verifying a given statement. And no news report or similar message employs concepts that are independent of cultural meaning. To check a news item in the proper sense, agents need to be socialized into an overall culture (for the following, see the literature on abilities of technical systems and "artificial intelligence" in particular, e.g., Collins & Kusch, 1998; Dreyfus, 1992). They would not only need knowledge about the world such as that water is wet and that South Africa is in Southern Africa – this type of knowledge, trivial as it might seem, really matters when we interpret descriptions, because we need to fill in the blanks of what has been left out merely because it is common knowledge.

Agents that would be able to fact check in the narrow sense would also have to be familiar with semantic rules as part of their socialization in order to reconstruct the meaning of possible "fake news." However, meaning does not only imply semantic rules and formal knowledge, but also relevance, a sense for what is important that is based on embodied competences and motivations. This sense helps us to decide what to take into account, where to start and stop when interpreting and judging messages, and to which assumedly true proposition we should compare what we think is the meaning of a part of the report – or whether we should modify our knowledge because we are convinced by something in the report.

Machine learning is not the same as socialization and the acquisition of actual knowledge of the world or an apprenticeship in fact-checking and much less the study of philosophical theories of truth. As mentioned above, researchers themselves often simply define a set of messages as true or false for the purpose of training by shifting the problem of what is true and false to journalists or crowdsourcing. This approach avoids the complex normative (or "asymmetrical") discussion of when we are really justified in considering a report truthful – or at least in generally trusting a source even if we cannot completely meet this standard of verification every time.

Even if the messages used for training have been checked by humans, current systems dealing with new information can only use peripheral properties, such as linguistic features of the messages, user reactions and patterns of diffusion, or properties of the source (e.g. Conroy, Rubin, & Chen, 2016; Saquete et al., 2020)

3 See Altunbey Özbay & Alatas (2020) and Saquete et al. (2020) for a systematic overview of different measures of success: The success rates are obviously lower under more realistic conditions and using standard datasets, and even if the authors do not discuss what would be satisfactory rates and which tradeoffs are to be made, the values seem to indicate that the systems are at best usable as heuristic prefilters for manual fact-checking. However, we cannot simply assume that much higher rates can be attained by simply refining the technology in use because what the systems is doing is something very different from actual factchecking (which is based on an "asymmetrical" attitude with clear criteria of truth). And we would not consider it a step towards actual factchecking but a heuristic alternative that does not turn into real factchecking by incremental improvement. Thus, even if the rates were higher, this would not necessarily mean that what the systems do is to actually fact-check the messages.

that are only contingently related to the veracity of the news items according to professional journalistic or scientific standards.

Asymmetric thinking that essentializes "fake news" may imply that they have certain more or less stable properties that are intrinsically linked to their being false and we only have to learn to detect more successfully. However, "fake news" that no longer exhibits these formal features may already exist somewhere or at least appear in the near future. Zhou, Guan, Bhat, and Hsu (2019) indeed demonstrate that detection systems are vulnerable to attacks using slightly modified versions of existing news items and are biased against accurate messages that are written in a non-journalistic style. Conversely, technical systems can create false positive results (and outsourcing not only the heuristic search for false information but also its control to technical systems can lead to dangerous unintended effects and the same infrastructure can be used for more sinister purposes).

And even if we are aware that we use heuristic criteria for detection, this is not somehow more "symmetrical" and thus somehow epistemologically appropriate because this is not the place for symmetry (and due to their typical intransparency, the systems are not very helpful in developing and testing symmetrical theories related to fake news). Nor is this approach "asymmetrical" in the right way because it does not match (or even really try to approximate) normative criteria of truth.

We would not deny that technical systems can give us hints as to which reports should be met with skepticism and that in a given context, such systems may be able to flag many false messages correctly based on learned heuristic criteria. Still, the technical solutions should not distract us from the necessity to support investigative journalism and to study how it can be strengthened and made more effective.

4.2 Media literacy and authority?

Improving media literacy is often suggested as the best response to the problem of "fake news," and one would expect that a communication researcher will recommend this solution. However, we would see it as half of the solution at best (see also Waisbord, 2018). Mihailidis and Viotty (2017, p. 449) consider this approach as insufficient because it places the burden on the (hopefully well-educated) individual and considers the problem solved. Kelkar (2019) similarly criticizes research on motivated reasoning in the context of post-truth as narrowly focused on (assumed) universal tendencies of the human mind instead of political and media-related contexts that explain why some controversies are politicized and others are not, and why and how alternative information ecosystems are being established.

More importantly in the present context, media literacy also appears as only half of the solution even if we decide to focus on individual knowledge, competences, and judgements of truth. Without a clear normative commitment defining what one needs to know to judge the truthfulness of messages, we risk focusing on the wrong types of knowledge and competences we seek to convey pedagogically or measure empirically. And if we define certain competences (such as media literacy) as necessary and thus worthy of investigation instead of symmetrically

analyzing which ones explain certain judgements, we may end up with spurious findings and problematic recommendations for interventions.

Imagine, for example, someone sharing a link to an article that claims that refugees coming to Sweden are offered free cellphones by state authorities. We may advise them to check the website for signs of its journalistic quality, for indications of political bias, etc. However, this consideration might be inconclusive, and the recipient may have already reached a verdict based on their knowledge about the social world: Regardless of the country, refugees usually do not receive more benefits than natives. Based on this information, readers can judge the story at least as relatively implausible without relying on media literacy in the proper sense, but on knowledge from earlier media use or general background knowledge.

For example, while Kahne and Bowyer (2017) report that media literacy education, not political knowledge, is the relevant factor when adolescents and young adults identify misinformation, this relationship is not entirely plausible: If, in an experimental study, all other properties of messages and communicators are held constant, how can media literacy explain how participants recognize that claims about exaggeratedly high or low tax loads are implausible and that other claims from the same source are plausible? Users will probably employ knowledge about tax policy that is correlated to the researchers' measurement of media literacy but not of political knowledge.

In order to debunk "fake news," the inconsistency with what one already knows about the world will probably often be no less important than media literacy. Media literacy education can only teach the actual or desirable norms and routines of journalistic work in detail and how journalists and other communicators follow them or fail to do so. However, even the highest degree of media literacy has its limits if general education, world knowledge and institutional structures that would allow people to judge the plausibility of reports are lacking (and thus also the basis for the evaluation of new sources). Thus, to rely on media literacy often means to rely on properties of sources and messages that are only accidentally related to what we would consider actual criteria of truth. Again, this does not mean that media literacy programs are useless – they can teach important heuristic rules and criteria for trust. But only if they were also to convey an understanding of valid journalistic research and reporting and if journalism were sufficiently transparent to check whether these standards are met, this would enable recipients to verify reports according to criteria of truth we would commit ourselves to. Without explicit normative standards of truth and without truly symmetrical analyses of judgments, media literacy is neither a sufficiently well-defined aim of education nor a useful analytical concept.

If we compare the discussion about "fake news" to reflections on scientific knowledge one last time, we notice a similarity between the paradoxes inherent in media literacy education and those in science education or science communication that point to more fundamental problems. The main goal of these endeavors should not be to teach people that what scientists have found out is "the facts," that we should trust people wearing lab coats or sitting in front of a large bookshelf, or certain trustworthy, "quality" news outlets, i.e., simply try to reestablish epistemic authorities. Instead, it is more important to understand methods (inclu-

ding the ways of thinking in the social sciences, humanities, etc. and methods of reporting), the types of knowledge they do and do not produce, and the degree of certainty of different findings (and we must discuss "knowledge democracy" in an even more encompassing way – see below). Conversely, if we seek to understand why people do or do not accept certain findings, we need to symmetrically reconstruct the worldviews in which certain scientific findings and methods or certain criticisms thereof make sense to people (for example, the ideologies that suggest that climate change is a hoax spread by left-wing, pseudo-scientific ideologues).

Researchers from the tradition of science and technology studies (STS) have joined the current discussion on "fake news" and addressed the thesis of a "breakdown of authority" (Bennett & Livingston, 2018, p. 128). We already referred to Marres' (2018) analysis of "demarcationism" behind fact-checking functionalities. The author stresses the importance of a public discussion of facts that goes beyond a mere top-down communication from experts or authorities to citizens, in particular a discussion of statements whose ascribed truth value changes over time. As noted above, this is not the end of a normative attitude towards the epistemic value of statements. It is at the core of a "knowledge democracy." Hoffman (2018) reminds us that it was authors from science and technology studies who pioneered the discussion of this concept. Taking up Marres' (2018) thesis that "we can't have our facts back," Hoffman (2018), however, warns that it would be dangerously idealist to insist on abstractions like "demaracationism," "knowledge democracy," and "symmetry" while others are "burning down the house" (Hoffman, 2018, p. 447) and working to systematically discredit science and journalism and to cultivate a public that is susceptible to "alternative facts." And isn't it strange to see "[s]ocial constructionists and postmodernists march in support of scientific facts against the conservative onslaught on science" (Waisbord, 2018, p. 1868)?

We would insist that research based on the idea of symmetry can best contribute to our understanding of how certain claims are being established as facts and what explains belief and disbelief among different social groups. Having analyzed this with an agnostic attitude, we can then insist on the difference between, for example, how science and producers of "fake news" establish facts and on the value of scientific methodology (as Hoffman, 2018, also demands), or reflect on the naive realism that many in journalism express in reaction to current criticism and attacks (Waisbord, 2018). Symmetry in the analysis of beliefs (including scientific knowledge) and knowledge democracy are not the same as false equivalence between scientific research and politically motivated claims imitating scientific findings in staged controversies, and the concept of symmetry did not really foreshadow or inspire some post-truth and anti-science era (Lynch, 2017, 2020).

"Embracing epistemic democratization does not mean a wholesale cheapening of technoscientific knowledge in the process. STS's detailed accounts of the construction of knowledge show that it requires infrastructure, effort, ingenuity and validation structures. Our arguments that 'it could be otherwise' (e.g. Woolgar and Lezaun, 2013) are very rarely that 'it could easily be otherwise'; instead, they point to other possible infrastructures, efforts, ingenuity and validation structures. That doesn't look at all like post-truth. […] Epistemic democratization has to

involve more equitable political economies of knowledge – and so critique does not 'run out of steam' with symmetry" (Sismondo, 2017, p. 3)

In sum, researchers can and should commit themselves to the idea that to collect and analyze data systematically, to reason and interpret rigorously, and to be as unprejudiced but also critical as possible is to be preferred to guesswork, hearsay, cherry picking, arguments from authority, sophistry etc. – even if their methods and conclusions may always be criticized not only by selected peers but by everyone, and even if research in STS paints a more nuanced picture of knowledge production than the opposition between "scientific facts" and ideologies, feelings, irrational personal beliefs, lies etc. suggests. Researchers' commitment to normative principles of research are fully compatible with their (temporary) bracketing of commitments to the truth or falsehood of certain news items (or scientific findings). In fact, the latter is the precondition to analyze certain processes surrounding such messages in a systematic and unbiased way, as we hope to have shown. And researchers should critically reflect whether the proposed solutions to the problem of "fake news" are really based on criteria of truth they would commit themselves to (even if the standards of good journalism cannot exactly match strict theories of truth and scientific methods) or only on heuristics and accidental properties of sources and messages – as useful as these solutions may otherwise be.

5. Conclusion

Terms, such as "fake news," that suddenly rise to public attention may only be buzzwords, but they may also find themselves at the center of an important discourse. We as researchers are tempted to think that a discourse based on these terms might become more rational and informed if we use the opportunity to weigh in – or at least citations and funding may beckon if we take up this much-debated issue. In any case, it is difficult to resist a terminology that is well-established in public discourse. We may well adopt such terms in order to link to the public conversation, but we should avoid certain problems often associated with everyday language and thinking. It can be a treasure trove of sedimented experience, but it may also follow an asymmetric logic. Social-scientific thinking can then break up essentialized social categories and the unquestioned egocentricity, ethnocentrism, and similar asymmetries.

For example, if in everyday conversations and political debates, only the others are ideologues, social theory and research teaches us to think of ourselves and everyone as someone who looks at the world from a particular and value-laden perspective. Historical and cultural comparison as well as critical theory then teach us not to essentialize our views. And if we shake our heads at others' acceptance of "fake news" and condemn the recklessness of its creators, we can still hope that communication research will not only help us reconstruct others' judgments about news and the media, but also teach us to reflect on the bases of our own convictions and our own trust in certain media institutions and bodies of knowledge. Instead of leading into relativism, such a symmetrical perspective can

enable us to understand others' worldviews and to critically engage with other perspectives in a true knowledge democracy.

Another "symmetrical" way in which researchers can relate to public discourses on "fake news" is to make them the object of study, as some authors have already done. The analysis of discourses and the analysis of how recipients judge and spread certain information are not mutually exclusive options. On the contrary, we can inform public discourse all the better if we can, on the one hand, present arguments for a theoretically informed and symmetric view on the diffusion and appropriation of messages, and findings from the symmetric analysis of how people deal with messages, and, on the other hand, if we know how discourses on disinformation are structured. Furthermore, such discourses are a research topic in their own right. It is important to understand who contributes to them (and with what ideological and strategical background), what interpretations are circulated and articulated with other discourses, and what the impact of such discourses may be.

Although politically and in terms of critical analysis, it may be highly relevant to deal with "fake news" (or "disinformation"), we have argued that it is epistemologically problematic to define one's object of analysis as news-like messages that are considered wrong by oneself. However, given the social relevance of the problem, we have critically discussed a number of solutions to the problem of messages we consider false being spread. While we concluded that technical solutions will be difficult to find and come with serious unintended (or, in the hands of more malicious actors, intended and dangerous) consequences, technology can support the struggle against misinformation by identifying the most unambiguous cases and by warning users about possibly problematic sources and messages while avoiding the repressive and chilling effects of the technological regulation of speech. And while we concluded that the call for "media competence" risks missing the point because the problem is more often about world knowledge and worldviews, media competence is certainly relevant when users try to identify reliable sources and to research and communicate competently. However, new competences can only really be conveyed if we understand peoples' previous knowledge. Our argument based on epistemological symmetry addresses this more fundamental level: Its function is to guide us towards a clearer understanding of the worldviews (and contexts) that lead people to accept claims that we consider false and even absurd (and those we consider to be true and even self-evident!). It is meant to give the otherwise perplexed and helpless question of "How can anyone believe *that*?" a constructive turn.

References

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, *31*(2), 211–36. https://doi.org/10.1257/jep.31.2.211

Allcott, H., Gentzkow, M., & Yu, C. (2019). Trends in the diffusion of misinformation on social media. *Research & Politics, 6*(2), 1–8. https://doi.org/10.1177/2053168019848554

Altunbey Özbay, F., & Alatas, B. (2020). Fake news detection within online social media using supervised artificial intelligence algorithms. *Physica A: Statistical Mechanics and its Applications, 540*, 123174. https://doi.org/10.1016/j.physa.2019.123174

Badia, A. (2020). An overview (and criticism) of methods to detect fake content online. In M. Khosrow-Pour (Ed.), *Encyclopedia of criminal activities and the deep web* (pp. 1053–1061). Hershey: IGI Global.

Ball, B. (2020). Defeating fake news. On journalism, knowledge, and democracy. *Moral Philosophy and Politics*. Advance online publication. https://doi.org/10.1515/mopp-2019-0033

Bennett, W. L., & Livingston, S. (2018). The disinformation order: Disruptive communication and the decline of democratic institutions. *European Journal of Communication*, *33*(2), 122–139. https://doi.org/10.1177/0267323118760317

Bloor, D. (1991). *Knowledge and social imagery*. Chicago: Chicago University Press.

Bloor, D. (1999). Anti-Latour. *Studies in History and Philosophy of Science Part A, 30*, 81–112. https://doi.org/10.1016/S0039-3681(98)00038-7

Brush, S. G. (1999). Why was relativity accepted? *Physics in Perspective, 1*, 184–214. https://doi.org/10.1007/s000160050015

Carlson, M. (2018). Fake news as an informational moral panic: the symbolic deviancy of social media during the 2016 US presidential election. *Information, Communication & Society*, 1–15. https://doi.org/10.1080/1369118X.2018.1505934

Chan, M. P. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: a meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, *28*(11), 1531-1546. https://doi.org/10.1177/0956797617714579

Collins, H., & Kusch, M. (1998). *The shape of actions. What humans and machines can do*. Cambridge: MIT Press.

Conroy, N., Rubin, V., & Chen, Y. (2016). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology, 52*, 1–14. https://doi.org/10.1002/pra2.2015.145052010082

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., … Quattrociocchi, W. (2016). The spreading of misinformation online. *PNAS, 113*, 554–559. https://doi.org/10.1073/pnas.1517441113

Dentith, M. R. X. (2017). The problem of fake news. *Public Reason, 8*(1–2), 65–79.

Dreyfus, H. (1992). *What computers still can't do. A critique of artificial reason*. Cambridge: MIT Press.

Egelhofer, J. L., Aaldering, L., Eberl, J. M., Galyga, S., & Lecheler, S. (2020). From novelty to normalization? How journalists use the term "fake news" in their reporting. *Journalism Studies, 21*, 1323–1343. https://doi.org/10.1080/1461670X.2020.1745667

Egelhofer, J. L., & Lecheler, S. (2019). Fake news as a two-dimensional phenomenon: A framework and research agenda. *Annals of the International Communication Association, 43*, 97–116. https://doi.org/10.1080/23808985.2019.1602782

Farkas, J., & Schou, J. (2018). Fake news as a floating signifier: hegemony, antagonism and the politics of falsehood. *Javnost – The Public*, *25*(3), 298–314. https://doi.org/10.1080/13183222.2018.1463047

Gelfert, A. (2018). Fake news: A definition. *Informal Logic, 38*, 84–117. https://doi.org/10.22329/il.v38i1.5068

Glick, T. F. (Ed.) (1987), *The comparative reception of relativity*. Dordrecht: Reidel.

Guess, A., Nyhan, B., & Reifler, J. (2018). *Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign*. European Research Council.

Habgood-Coote, J. (2019). Stop talking about fake news! *Inquiry, 62*, 1033–1065. https://doi.org/10.1080/0020174X.2018.1508363

Hoffman, S. G. (2018). The responsibilities and obligations of STS in a moment of post-truth demagoguery. *Engaging Science, Technology, and Society*, *4*, 444–452. https://doi.org/10.17351/ests2018.259

Iosifidis, P., & Nicoli, N. (2020). The battle to end fake news: A qualitative content analysis of Facebook announcements on how it combats disinformation. *International Communication Gazette, 82*, 60–81. https://doi.org/10.1177/1748048519880729

Jankowski, N. W. (2018). Researching fake news: A selective examination of empirical studies. *Javnost – The Public*, *25*, 248–255. https://doi.org/10.1080/13183222.2018.1418964

Kahne, J., & Bowyer, B. (2017). Educating for democracy in a partisan age: confronting the challenges of motivated reasoning and misinformation. *American Educational Research Journal, 54*, 3–34. https://doi.org/10.3102/0002831216679817

Kelkar, S. (2019). Post-truth and the search for objectivity: political polarization and the remaking of knowledge production. *Engaging Science, Technology, and Society, 5*, 86–106. https://doi.org/10.17351/ests2019.268

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., & Schudson, M. (2018). The Science of fake news. *Science*, *359*, 1094–1096. https://doi.org/10.1126/science.aao2998

Levy, N. (2017). The bad news about fake news. *Social Epistemology Review and Reply Collective, 6*(8), 20–36.

Lewandowsky, S., Ecker, U., & Cook, J. (2017). Beyond misinformation: understanding and coping with the 'post-truth' era. *Journal of Applied Research in Memory and Cognition, 6*, 353–369. https://doi.org/10.1016/j.jarmac.2017.07.008

Lynch, M. (2017). STS, symmetry and post-truth. *Social Studies of Science, 47*, 593-599. https://doi.org/10.1177/0306312717720308

Lynch, M. (2020). We have never been anti-science: reflections on science wars and post-truth. *Engaging Science, Technology, and Society, 6*, 49–57. https://doi.org/10.17351/ests2020.309

Marres, N. (2018). Why we can't have our facts back. *Engaging Science, Technology, and Society*, *4*, 423–443. https://doi.org/10.17351/ests2018.188

Mihailidis, P., & Viotty, S. (2017). Spreadable spectacle in digital culture: civic expression, fake news, and the role of media literacies in "post-fact" society. *American Behavioral Scientist, 61*, 441–454. https://doi.org/10.1177/0002764217701217

Müller, P., & Denner, N. (2018). *Was tun gegen „Fake News"? Eine Analyse anhand der Entstehungsbedingungen und Wirkweisen gezielter Falschmeldungen im Internet* [What to do against "fake news." An analysis based on the conditions of emergence and the effects of deliberate disinformation on the internet] [2nd edition]. Potsdam: Friedrich-Naumann-Stiftung für die Freiheit.

Neuberger, C., Bartsch, A., Reinemann, C., Fröhlich, R., Hanitzsch, T., & Schindler, J. (2019). Der digitale Wandel der Wissensordnung. Theorierahmen für die Analyse von Wahrheit, Wissen und Rationalität in der öffentlichen Kommunikation [The digital change of the knowledge order. A theoretical framework for the analysis of truth, knowledge and ration-

ality in public communication]. *Medien & Kommunikationswissenschaft, 67*, 167–186. https://doi.org/10.5771/1615-634X-2019-2-167

Pennycook, G., & Rand, D. (2017). *Who falls for fake news? The roles of analytic thinking, motivated reasoning, political ideology, and bullshit receptivity*. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3023545

Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition, 188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

Perez-Rosas, V., Kleinberg, B., Lefevre, A., & Mihalcea, R. (2017). *Automatic detection of fake news*. Available online: https://arxiv.org/abs/1708.07104v1

Quinn, A. (2017). Fake news, false beliefs, and the need for truth in journalism. *International Journal of Applied Philosophy, 31*, 21–29. https://doi.org/10.5840/ijap201771884

Reed, B. (2000). Accidental truth and accidental justification. *Philosophical Quarterly, 50*, 57–67. https://doi.org/10.1111/1467-9213.00167

Rini, R. (2017). Fake news and partisan epistemology. *Kennedy Institute of Ethics Journal, 27* (S2), 43–64. https://doi.org/10.1353/ken.2017.0025

Saquete, E., Tomás, D., Moreda, P., Martínez-Barco, P., & Palomar, M. (2020). Fighting post-truth using natural language processing: A review and open challenges. *Expert Systems with Applications, 141*, 112943. https://doi.org/10.1016/j.eswa.2019.112943

Shin, J., Jian, L., Driscoll, K., & Bar, F. (2018). The diffusion of misinformation on social media: Temporal pattern, message, and source. *Computers in Human Behavior*, *83*, 278–287. https://doi.org/10.1016/j.chb.2018.02.008

Sismondo, S. (2010). *An introduction to science and technology studies* [2nd edition]. Chichester: Wiley-Blackwell.

Sismondo, S. (2017). Post-truth? *Social Studies of Science, 47*, 3–6. https://doi.org/10.1177/0306312717692076

Swire, B., Berinsky, A., Lewandowsky, S., & Ecker, U. (2017). Processing political misinformation: Comprehending the Trump phenomenon. *Royal Society Open Science, 4*(3), 1–21. https://doi.org/10.1098/rsos.160802

Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining "fake news" A typology of scholarly definitions. *Digital Journalism, 6*, 137–153. https://doi.org/10.1080/21670811.2017.1360143

Tandoc, E. C., Lim, D., & Ling, R. (2020). Diffusion of disinformation: How social media users respond to fake news and why. *Journalism, 21*, 381–398. https://doi.org/10.1177/1464884919868325

Tucker, J., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S. … Nyhan, B. (2018). *Social media, political polarization, and political disinformation: A review of the scientific literature*. Menlo Park: William and Flora Hewlett Foundation. Available online: https://www.hewlett.org/wp-content/uploads/2018/03/Social-Media-Political-Polarization-and-Political-Disinformation-Literature-Review.pdf

Van Duyn, E., & Collier, J. (2019). Priming and fake news: The effects of elite discourse on evaluations of news media. *Mass Communication and Society, 22*, 29–48. https://doi.org/10.1080/15205436.2018.1511807

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

Waisbord, S. (2018). Truth is what happens to news. *Journalism Studies, 19*, 1866–1878. https://doi.org/10.1080/1461670X.2018.1492881

Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2019). Fact-checking: A meta-analysis of what works and for whom. *Political Communication, 37*, 350–375. https://doi.org/10.1080/10584609.2019.1668894

Wardle, C. (2018) The need for smarter definitions and practical, timely empirical research on information disorder. *Digital Journalism, 6*, 951–963. https://doi.org/10.1080/21670811.2018.1502047

Wazeck, M. (2014). *Einstein's opponents: The public controversy about the Theory of Relativity in the 1920s.* Cambridge: Cambridge University Press.

Woolgar, S. & Lezaun, J. (2013). The wrong bin bag: A turn to ontology in science and technology studies? *Social Studies of Science 43*, 321–340. https://doi.org/10.1177/0306312713488820

Zhang, X., & Ghorbani, A. A. (2020). An overview of online fake news: Characterization, detection, and discussion. *Information Processing & Management, 57*(2), 102025. https://doi.org/10.1016/j.ipm.2019.03.004

Zhou, Z., Guan, H., Bhat, M. M., & Hsu, J. (2019). *Fake news detection via NLP is vulnerable to adversarial attacks*. Available online: https://arxiv.org/abs/1901.09657

Zimmermann F., & Kohring, M. (2018) „Fake News" als aktuelle Desinformation. Systematische Bestimmung eines heterogenen Begriffs ["Fake news" as topical disinformation. A systematic definition of a heterogenous concept]. *Medien & Kommunikationswissenschaft, 66*, 526–541. https://doi.org/10.5771/1615-634X-2018-4-526

Zollo, F., & Quattrociocchi, W. (2018). Misinformation spreading on Facebook. In S. Lehmann & Y. Ahn (Eds.), *Complex spreading phenomena in social systems. Influence and contagion in real-world social networks* (pp. 177–196). Cham: Springer.