The Concept of Fake News

Introduction

In 2017, terms such as "post-truth", "fake news" or "alternative facts" suddenly became part of public discourse, and almost immediately scholars began to argue about their meaning. In particular, various definitions of the concept of fake news have been put forward and critically discussed in the literature (Rini 2017, Dentith 2017, Gelfert 2018, Jaster & Lanius 2018, Mukerji 2018, Zimmermann & Kohring 2019, Fallis & Mathiesen 2020). Yet, so far, there has been little explicit reflection on the methodological underpinnings of this conceptual work. Most proponents of definitional proposals appear to operate on a traditional conceptual analysis approach: The assumption seems to be that we have a firm intuitive grasp on the concept of fake news and that this puts us in a position to directly evaluate the merits of a proposed definition by evaluating its suitability in capturing this very concept. There is a striking lack of any previous problematization of methodology, as most proposals are based on this assumption by starting out to capture a set of allegedly uncontroversial cases of fake news.

From the beginning, this strand of the debate was accompanied by criticism. As Habgood-Coote (2019, 2022) and others have pointed out, the task of defining "fake news" may not be that simple: Our ordinary, every-day use of the term "fake news" is gerrymandered, politicized, inflationary, and fuzzy, and so we may not get very far just by looking at linguistic intuitions. Habgood-Coote draws the somewhat radical conclusion that it would be best to "stop talking about fake news" altogether.[1] In our view, however, this throws the baby out with the bath water. As we have argued elsewhere, there are more shared commitments between the seemingly gerrymandered definitions on the market than one would initially think (Jaster & Lanius 2021). It is not at all inconceivable that the proponents in the definitional debate are circling around a real phenomenon that the term "fake news" could be used to refer to. And while there may still be quite some heterogeneity in the understanding of "fake news", this is no reason to refrain from offering a definition. On the contrary, it could arguably help to reduce the heterogeneity in the use of the term or at least facilitate a better understanding of the concept.

We agree, however, that the methodology employed so far may not be entirely suitable for the task at hand. In particular, it may not deliver the means to systematically evaluate the various definitions against one another. The question that has so far received too little attention is: What do we want from a definition of "fake news"? Or, differently put: What are the standards by which to measure a definition? In this chapter, we take steps toward answering this question. Thereby, we hope to provide more solid methodological underpinnings for the discussion about the proper definition of "fake news".

In the first part, we come up with a set of criteria for the evaluation of definitions of "fake news". As we will see, in defining the concept of fake news, various theoretical, epistemic, and practical criteria can be taken into account. In the second part, we use our own definition to

---

[1] For criticisms of this line of reasoning, see e.g. Brown (2019) or Pepp, Michealson and Sterken (2022).

illustrate how the application of the proposed criteria plays out in evaluating a specific proposal. As we go along, we selectively compare the merits and demerits of our proposal to those of some other extant definitions. Finally, in a brief outlook, we sketch how the suggested methodological approach may bring up new research questions in the experimental philosophy of fake news.

Defining the Concept of Fake News

Instead of defining "fake news" exclusively in the spirit of conceptual analysis, we suggest that the first question in the endeavor of defining "fake news" should be: What do we want from such a definition? Or, in other words: What are the goals we should pursue in defining "fake news"?

Some of the goals worth pursuing will clearly be the ones also pursued by traditional conceptual analysis: A definition should not be at odds with paradigm cases of the phenomenon it sets out to capture, it should be precise and informative, and it should locate the concept of fake news in the landscape of related concepts.

But next to these, there are other important goals as well. One such goal, famously emphasized by Carnap (1950) in connection to his methodological paradigm of explication, is for a definition to advance the explanatory power of our theories about a subject matter. More generally, we may pursue a variety of epistemic goals in defining a concept: We will want the definition to be easily graspable, explanatorily relevant, and, in many cases, operationalizable in empirical research. These goals may well interfere with the goals pursued by conceptual analysis.

Epistemic goals are partly practical in nature: In contrast to the purely theoretical goals of conceptual analysis, they are formulated with an eye toward the way people operate with the definition cognitively. To these goals, we may further add purely practical, i.e. moral or political goals: In our case, for instance, the goals of counteracting and regulating fake news come to mind. A definition may be more or less suitable for reaching these aims, and its use for our moral-political goals may cut across theoretical or epistemic aims quite drastically in some cases. Haslanger (2000), for instance, suggests a definition of "woman" that has little to do with the way the term is used – both in ordinary parlance and scientific theories. The goal she pursues in her "ameliorative" project is first and foremost a moral-political one, in this case: to help ending the subordination of women.

Of course, one need not prioritize practical goals as radically as Haslanger does, and it may not always be entirely clear which definition would serve the practical goals best. At the same time it would arguably be unfortunate to disregard practical goals altogether when setting out to define concepts that play a crucial role in societal debates such as the concepts of woman or fake news.

In what follows, we will elaborate in more detail on the goals so far mentioned and on the criteria for a definition of "fake news" that can be derived from them. By drawing attention to the goals and their relevance for setting and evaluating the various criteria, we move away from the paradigm of traditional conceptual analysis and into the methodological realm of

conceptual engineering and re-engineering (Cappelen 2018, Chalmers 2020): Our aim is to provide guide rails to assess the proposed definitions of "fake news" along the theoretical, epistemic and practical criteria thus derived.

Theoretical Requirements

The most obvious theoretical requirement for any definition is its extensional adequacy. Traditionally, this means that the definition fixes the extension in such a way that it clearly and precisely includes all the objects it is supposed to include and excludes all the objects it is supposed to exclude. In the case of "fake news", the extension of the term is somewhat controversial. Yet, there is widespread agreement on a large number of cases and there is an understanding that those engaged in defining the concept of fake news are after the same phenomenon. In what follows, we will focus on three theoretical requirements: the concept's similarity to any pre-existing concept, its exactness, and its connection to related concepts.

Similarity

When engaging in conceptual analysis, we want both extension and intension to be as close as possible to the concept already in use. When re-engineering a concept, a weaker requirement is in place: We want extension and intension to be at least sufficiently similar to the concept already in use. We do not want to end up with an entirely different concept afterwards.[2] In the case of "fake news," this means that any definition "must preserve sufficient similarities to our imprecise and vague ordinary usage of the term" (Grundmann 2020: 6).

On the one hand, a definition should thus be extensionally similar: It should include all or at least most clear cases of fake news and exclude all or at least most clear cases of things other than fake news. There is widespread agreement in the literature on this requirement. Mukerji, for instance, submits that a definition of fake news ought to "fit the [paradigmatic] test cases" (Mukerji 2018, 932) and must be neither overinclusive nor overexclusive (Mukerji 2018: 938ff.).

Now, it is questionable whether there is any distinctive kind in reality that corresponds to the concept of fake news and to which the term "fake news" definitely refers. What we have, however, are intuitions people have when using the concept. Fallis and Mathiesen (2019: 3), for instance, point out that "we would like a definition of fake news that agrees with our intuitions about whether particular cases are or are not instances of fake news".

On the other hand, then, we want the definition to be intensionally similar: It is a merit of a definition to be in line with the intuitions people have about the concept of fake news. Looking at various existing definitions and characterizations of fake news, it becomes apparent that fake news is taken to relate to falsity and misleadingness, to bullshit, lies and deception, fraud and fakery, and to virality and the effects of digitalization (Jaster & Lanius 2021: 25ff.). Ideally, a definition should reflect – or at least make sense of – these intuitions people have about the concept of fake news.

---

[2] While there are conceptual engineering-projects that introduce completely new concepts, the problem of topic change has been recently debated in the literature on conceptual re-engineering quite extensively (Cappelen 2020, Sundell 2020, Belleri 2020, Koch 2021).

Exactness

While taking into account the extension and intension of the concept already in use, a definition would also be considered better if it avoids problematic imprecisions of the ordinary language term. The definition should not yield unnecessary borderline cases, which cannot be determinately classified as fake news or determinately classified as not fake news. It is often pointed out that the term "fake news" is particularly unclear and that it is used in very different senses by different people (Coady 2019; Habgood-Coote 2019). Ideally, we want to preserve the clear cases of the original ordinary language term and determine unclear cases by reference to the definition. In this way, the definition of "fake news" should be informative: It should resolve (as much as possible) borderline and other indeterminate cases in the ordinary usage of the term.

To achieve this end, the definition should avoid unnecessary vagueness (Mukerji 2018: 937) and other unnecessary indeterminacies in the terms used in the definiens: Most ordinary language expressions are not only vague, but also indeterminate in other ways; they are lexically ambiguous, polysemous or standard-relative. These indeterminacies are to be reduced and the definiens should be as linguistically determinate as possible.

Conceptual Interrelatedness and Distinctiveness

The theoretical adequacy of a concept is also reflected in its relation to other concepts. Fake news is plausibly an instance of untruthful or insincere speech. It is a merit of a definition of "fake news" to elucidate the relationship between fake news and other forms of such speech already established in epistemology and philosophy of language, such as "lies, misleading, bullshitting, false assertion, false implicature, being unreliable, distorting the facts" (Habgood-Coote 2019: 1047).

One way the exactness requirement plays out in connection with the concept of fake news is that a definition should counteract the inflationary use of "fake news". There is wide agreement that we should "stop calling everything 'fake news'" (Oremus 2016), and more specifically, that a definition should refrain from "collapsing [...] various existing distinctions between different types of public disinformation" (Gelfert 2018: 94) into a single concept. Instead, "any putative definition of 'fake news' must be situated in relation to these varied forms of public disinformation" (Gelfert 2018: 95), such as biased or selective reporting, satire and parody, propaganda, and conspiracy theories (Jaster & Lanius 2018, 2019, 2021) as well as gossip, rumor, hoaxes, and urban legends (Gelfert 2018: 94).[3]

It is furthermore agreed that fake news needs to be distinguished from inadvertent errors in reporting: The definition is too broad if it "incorrectly counts honest mistakes (...) as fake news" (Fallis & Mathiesen 2019: 4) and collapses the "fundamental difference between fake news and journalistic errors" (Jaster & Lanius 2018: 4; see also Gelfert 2018 and Rini 2020). Conceptual distinctiveness in this sense would be, *ceteris paribus*, a benefit of any definition of "fake news."

---

[3] That is not to say that there may not be an extensional overlap between fake news and some of these other phenomena (Jaster & Lanius 2019, 35ff; 2021, 23ff.).

Epistemic Fruitfulness

By fulfilling these requirements, a definition will ideally contribute to the explanatory power of the theories that science is developing about the phenomenon of fake news. At best, it helps to understand the phenomenon of fake news and perhaps even predict novel phenomena related to it. However, whether the definition of "fake news" is epistemically fruitful in the sense of furthering our understanding of a phenomenon depends not only on the discussed theoretical requirements but also on further aspects such as the definition's transparency and the defined concept's integrability into our theories, its explanatory relevance, and its empirical operationalizability.

Epistemic Transparency

A definition is epistemically transparent if it is easy to understand under which circumstances the defined term applies – both abstractly and concretely. Epistemic transparency is a property of definitions that primarily helps to promote cognitive goals: It facilitates the use and understanding of the defined term and the underlying concept. The more epistemically transparent a concept is, the easier it is, *ceteris paribus*, to recognize the phenomenon in the real world.

Epistemic transparency thus reduces the cognitive load when mentally processing the concept, which may allow for a broader and more uniform reception of the term and reduce the likelihood of misunderstandings. This depends on transparency on other levels: We can distinguish structural, conceptual, semantic and pragmatic transparency with respect to the definition of "fake news."

A definition is structurally transparent if the definiens is logically and syntactically simple, i.e. if it does not contain unnecessary logical operators or grammatical constructions such as disjunctions or "if, then"-clauses. Too many instances of, for example, "and" and "or" in a definition of "fake news" will reduce a definition's structural transparency.

A definition is conceptually transparent if the concepts used in the definiens are easily comprehensible to anyone who is going to use the definiendum. That is, the concepts used in the definiens of "fake news" should themselves be as epistemically transparent as possible.

A definition of a compositum like "fake news" should ideally also be semantically transparent: The choice of the ordinary term in use should be elucidated. More specifically, the "fake" and the "news" in "fake news" should be made sense of: In which sense is fake news *news* and in which sense is it *fake*? Ideally, an answer to this question is implied by the definition.

A definition of a highly politicized and strategically used term should moreover be pragmatically transparent: The specifics of the term's pragmatic usage in actual conversations should be clarified by the definition. How does, for instance, the use of "fake news" in Trump's accusation "You are fake news!" against the New York Times and CNN relate to the concept of fake news as defined? Does Trump employ the same or a different concept? And if it is the same concept, how come it is so prone to being used derogatorily? It would be, *ceteris paribus*,

a benefit of the definition if it helped clarify this usage instead of conceptualizing the phenomenon in a way that would necessitate ad-hoc explanations of fake news accusations.

Explanatory Relevance

Most researchers want a definition of "fake news" to be useful to the empirical study of the phenomenon. Many empirical studies on fake news aim to find out what factors are relevant to its emergence, distribution and influence in societies. A definition might thus be judged on its ability of "addressing the explanatory challenge of why fake news has become such a force in the online world" (Gelfert 2018: 101f.) and which distinctive role it plays in the distribution of conspiracy theories, propaganda, and informational warfare. Certainly, it will "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" (Gelfert 2018: 111). Lacking a definition of "fake news" is a theoretical deficit that can threaten the practical evaluation of the dangers of fake news to society (Zimmermann & Kohring 2018: 530).

Integrability and Operationalizability

There is a lot of research in media studies, communication science, computer and information science as well as political science and political philosophy (including critical theory, critical race theory, and feminist theory) related to fake news. Preferably, our use of "fake news" should not "[cut] us off from these theoretical resources" (Habgood-Coote 2022, 501). Differently put, the definition of "fake news" should be theoretically useful in being integrable into our various theories. At best, it might even help to consolidate the various theories themselves.

To this end, the concept of fake news should be operationalizable in order to facilitate measuring the causes, effects, and the propagation of fake news. If elements of the definiens are impossible to observe, intransparent or too imprecise, empirical researchers will have a hard time to put the definition into use. As a consequence, they may stipulatively define concepts such that they be easily operationalizable, but fail to meet theoretical requirements such as similarity or exactness: Different research projects employ different concepts that are designed exclusively with an eye toward the specific empirical endeavor at stake, but lack similarity both with the common use of the term and with one another. What is called "fake news" in one study may then be altogether different from what is called "fake news" in another. A suitable definition will respect operationalizability, facilitate communication between different researchers and disciplines, and help to reduce misinterpretations. This will increase the explanatory power of our theories in the different sciences that are concerned with the phenomenon of fake news in one way or another.

Practical Goals

When engineering a concept, practical goals are sometimes most decisive. Broadly speaking, we eventually want our concepts to help solve problems and sometimes even further moral-political progress. Reflecting on his own proposed definition of "fake news," Mukerji holds that the "most important and interesting question is [...] a practical one: What, if anything, should we do about fake news?" (Mukerji 2018: 942). His assessment is shared by many politicians,

social media platforms, NGOs, and journalists: Fake news is standardly taken to contribute to false beliefs, polarization and radicalization. A suitable definition should thus "help us deal with the threat to knowledge and democracy posed by fake news" (Fallis & Mathiesen 2019: 3; see also Zimmermann & Kohring 538).

A definition may serve this overarching practical goal in a number of ways. First, by enabling constructive scientific and political discourse about the threats fake news poses to democracy. Secondly, by its suitability to figure in legislation and thereby contribute to the regulation of fake news: A proper definition of "fake news" may be useful for enacting legal norms and enforcing them to effectively prevent or weaken the threat of fake news. Thirdly, by enabling the recognition of and the proper reaction to fake news: At best, a definition will serve the education of epistemic competences needed to effectively deal with fake news (Lanius 2021).

Putting the Methodology to Work

In the last section, we have identified a number of criteria by which to judge the overall quality of a definition of "fake news." Ideally, a definition would satisfy all of the criteria as fully as possible. In practice, conceptual work will usually consist in trade-offs between the various requirements. For reasons of space, we cannot systematically evaluate all definitions on the market. We will, instead, present our own definition, evaluate it against each of the criteria introduced in the last section, and selectively compare its merits and demerits to other definitions as we go along. In doing so, we demonstrate the need for trade-offs between the various goals at stake. We do not want to argue for our own definition here; its analysis is intended to merely serve as illustrative material for the methodology we are proposing and to deliver a blueprint for a more systematic and comprehensive assessment of the extant definitions.

A Definition of "Fake News"

In our view (Jaster & Lanius 2018, 2019, 2021), a news report is fake news iff it *lacks truth* in the sense of being false or misleading and *lacks truthfulness* in the sense of being distributed with an intention to deceive or a disregard for the truth.[4]

According to the first conjunct of the definition, fake news is false or misleading. The point is well-taken in the philosophical literature: While typically false, a fake news report need not literally state a falsehood (Fallis 2016: 338f.; Gelfert 2018, 100; Mukerji 2018: 935f.). Instead, it may be misleading. We understand misleadingness in terms of falsity in pragmatic content: A misleading statement may well state something that is literally true, while communicating something false (Jaster & Lanius 2018, 209f.; 2021: 21f.). This can be done by implicating, impliciting or presupposing something false, while literally saying something true.

A good example for a fake news report that is misleading in this sense is a Breitbart report (Hale 2017), according to which a Muslim mob had been calling "Allahu akbar" and set fire to a church in Germany on a new year's eve. The report does not make a false statement: There

---

[4] Once a piece of fake news is in circulation, it may be further distributed in good faith. As in the case of lies, the origin of fake news reports requires a certain mind-set of their creators, while the reports may then be spread by people who genuinely believe them (Jaster & Lanius 2021, 23).

had been a gathering of Muslim men, they had been calling "Allahu akbar" to celebrate the new year, and fireworks had accidentally set a safety net on fire that was attached to a scaffolding at the church's outer walls. While not literally false, however, the report is clearly misleading: The wording suggests that the crowd was aggressive, pursued ideological motives, and purposefully set the church itself (and not just the safety net) on fire.

The second conjunct of the definition states that fake news is distributed with an intention to deceive or a disregard for the truth. Cases in which fake news is distributed with an intention to deceive easily come to mind: Consider Russian state media reports about the war in Ukraine or the Breitbart report mentioned above. Fake news is often distributed for political or ideological purposes and in those cases, the originators of the report typically exhibit deceptive intentions.

At the same time, there is widespread agreement in the philosophical literature that "[n]ot all purveyors of fake news intend to deceive people into believing the stories that they post" (Fallis & Mathiesen 2019: 5). Instead, some originators of fake news show a disregard for the truth of their reports. The classic example for this is the infamous case of a group of Macedonian teenagers who invented large amounts of politically tendentious news reports for click-baiting purposes (Subramanian 2017). In such cases, fake news is what is commonly called "bullshit" (Frankfurt 1986).

Testing for Theoretical Requirements

Testing for Similarity

How does this definition fare with respect to the similarity dimension? Presumably, it accounts for a number of intuitions people have about the concept of fake news: The definition picks up on the notions of falsity and misleadingness, relates fake news to lies and bullshit, and integrates ideas about fraud and fakery by emphasizing the lack of truthfulness on the originator's part.

Also, the definition seems to capture the paradigmatic cases of fake news. As we have argued elsewhere (Jaster & Lanius 2018, 2019, 2021), this is because all of them fall into one of the following four categories:

(1) False news propagated with an intention to deceive
(2) False news propagated with a disregard for the truth
(3) Misleading news propagated with an intention to deceive
(4) Misleading news propagated with a disregard for the truth

To our knowledge, there are only two kinds of cases that are recognized as fake news by other accounts in the literature, while not, or not obviously, being recognized as such by our definition. But neither of the two seem to be paradigmatic instances of fake news.

First, there are instances of straightforwardly true (i.e., neither false nor misleading) bullshit reports (Mukerji 2018: 938). However, Mukerji discusses such cases (which clearly do not fall under our definition) only in anticipation of the objection of his account being overinclusive. The majority of proposed definitions exclude cases of true news from fake news.

Second, there is the case of algorithmically distributed falsities:

> Consider a case in which, as an unintended and unexpected by-product of the technical architecture or its algorithms, our internet browsers filter stories in such a way that their users receive mostly false news. The production of this news does not need to involve any epistemically bad intentions. [Instead] the stories that are disseminated were originally just meant as fiction or satire (Grundmann 2020: 5)

Whether or not such cases fall under our definition will depend on whether or not we are prepared to ascribe an intention or other mental attitude (such as bullshit) to algorithms. Be that as it may, the reports thus generated, while clearly a potential source for massive misinformation, are not paradigmatic instances of fake news either: Whether or not the case should be covered under a definition is an open question that can and presumably will be controversially assessed.

Testing for Exactness

While the definition determines most cases quite clearly, there remain some borderline cases. Sometimes, it will be unclear whether the threshold for indifference to the truth has been surpassed to count as lack of truthfulness, since there is no clear boundary between extreme negligence and an utter disregard for the truth (for instance, when ruthlessly monetizing attention). In the propagation of some yellow press reports truth is clearly not a priority, but does this mean that they are bullshit? Here, the definition apparently allows for a gray area on a scale of stronger or weaker forms of disregard for the truth.

Does this amount to a demerit of the definition? Arguably, the remaining vagueness corresponds to a continuity in the underlying phenomenon: That there is a spectrum of degrees of disregard for the truth that people in fact exhibit. If that is the case, the remaining vagueness may be theoretically eliminable, but its elimination would result in reducing the fit between the concept and its underlying phenomenon. Transitions between negligent reporting and fake news are continuous – and that is to be expected because disregard for truth comes in degrees.

While not being maximally precise, then, our definition is more precise than the ordinary language use of "fake news" and arguably more precise than other definitions on the market: Gelfert, for instance, defines fake news as the "deliberate presentation of (typically) false or misleading claims as news, where these are misleading by design" (Gelfert 2018: 84). Not only does the "typically"-clause render the definition inexact, there is also no exact resolution to the requirement that a claim is misleading "by design". Note, too, that neither instance of inexactness in Gelfert's definition is of scalar kind. The formulations leave it open whether particular cases are *potentially* covered by the definition at all.

In contrast, what counts as being false, being misleading, intended to deceive or being bullshit is sufficiently clear to decide most cases – due to the use of well-established technical vocabulary from the philosophy of language. However, the definition is probably less precise than other definitions such as Mukerji's or Fallis and Mathiesen's who also rely on well-established vocabulary, but whose definitions are less complex (see below).

Testing for Conceptual Interrelatedness and Distinctiveness

As already noted, the definition is formulated in terms of concepts familiar from the literature: Fake news is *false* or *misleading* and it is distributed with an *intention to deceive* or a *disregard for the truth*, i.e. with a *bullshit* attitude. We spell out misleadingness in terms of *falsity in pragmatic content*. Thus, our definition establishes a clear connection between the concept of fake news and these other concepts well established in the philosophy of language.

The relation to lying is less clear. Depending on one's definition of lying, some fake news will indeed be lies. According to the traditional account, a "lie is a statement made by one who does not believe it with the intention that someone else shall be led to believe it" (Isenberg 1973, 248). Fake news reports that are propagated with an intention to deceive will usually be lies on this account. Other fake news reports will not be lies, however, because they are propagated only with a disregard for truth, but not with an intention to deceive. This is also the case according to most modern accounts (Carson 2006; Fallis 2012; Saul 2012; Faulkner 2013; Lackey 2013; Stokke 2013). While only some of them explicitly require the liar to have an intention to deceive, most of them require her to believe what she says to be false.

Habgood-Coote (2018) is right, then, that "fake news" can be defined in terms of the established vocabulary from the field of untruthful or insincere speech. Importantly, however, that does not mean that the concept of fake news collapses into any of these established concepts. As we have argued elsewhere (Jaster & Lanius 2018, 2019, 2021), the definition – as do most other definitions on the market – successfully helps to set fake news apart from other forms of problematic reporting, such as journalistic errors and selective reporting, as well as other phenomena related to disinformation, such as conspiracy theories, satirical and parodistic "news," and propaganda.

In most of these cases, there is an extensional overlap between the phenomena. Some fake news is also propaganda and *vice versa*, for instance. Propaganda can be any communication "designed to manipulate a target population by affecting beliefs, attitudes, or preferences in order to obtain behavior compliant with the political goals of the propagandist" (Benkler et al. 2018: 29). A fake news report distributed with an intention to deceive in order to further some political goal will clearly fall into this category. Thus, some fake news is also propaganda. But there are also instances of fake news that are not propaganda: Think of the reports distributed without any political goals by the Macedonians. And there is propaganda that does not qualify as fake news: Think of propagandistic imagery showing a resolute soldier or a happy family.

In other cases, our definition of fake news excludes – as do virtually all other definitions – such an overlap. Satirical "news", for example, is not fake news because it is not set in a truth warranting context and therefore does not qualify as news reports in the first place (Jaster & Lanius 2021: 23). This is something most definitional proposals aim to set apart – presumably also for practical goals (see below).

Testing for Epistemic Fruitfulness

Testing for Epistemic Transparency

Above we said that a definition is epistemically transparent if it is easy to understand under which circumstances it applies. How well-suited is the proposed definition with respect to this goal? Structurally, the definition is not as simple as one would hope. There are two necessary conditions (lack of truth and lack of truthfulness) each of which can be satisfied in two different ways. Grundmann therefore considers our definition "hopelessly disjunctive" (Grundmann 2020: 3) – and he is right.

Now, many extant definitions are somewhat unwieldy. Rini (2017: E45), for instance, defines a "fake news report" as "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." Even the seemingly simple definition by Gelfert (2017: 108) is more complex than it looks at first sight. Gelfert defines "fake news" as "the deliberate presentation of (typically) false or misleading claims as news, where the claims are misleading by design." Like our definition, Gelfert's contains a disjunction ("false or misleading"), but it also adds a "typically"-qualifier and specifies deliberateness, misleadingness and the "by design"-clause as further distinct constraints.

Yet, there are simpler definitions as well. Fallis and Mathiesen (2019), for instance, define "fake news" as "counterfeit news", which means that "a story (...) is not genuine news, but is presented as genuine news, with the intention and propensity to deceive." The definition is quite simple in that it treats fake news as the deceitful presentation of not-news as news, which does not impose a particularly complex structure on the definition. Likewise, Mukerji (2018, 929) defines "fake news" as "bullshit asserted in the form of a news publication." The superordinate concept is bullshit, which then is specified as one that is asserted in a particular form. Focusing on structural transparency, these two definitions are evidently better than ours, Gelfert's, or Rini's.

Turning to conceptual transparency, our definition contains two rather subtle notions: the concept of misleadingness we use is technical and differs from the ordinary concept of misleadingness. Moreover, as the literature on bullshit reveals, "indifference to the truth" has proven difficult to cash out.[5] Thus, the definition turns out to be not quite as conceptually transparent as it may appear at first sight. Similar problems beset Gelfert's "by design"-phrase, Mukerji's appeal to bullshit, and, to some extent, Fallis and Mathiesen's notion of counterfeit news. Rini, in contrast, formulates her definition almost exclusively in ordinary parlance.

Concerning semantic transparency, our definition is clear about the "news" in the compositum "fake news": We take fake news to be a species of news, where "news" comprises any report about typically recent events which is directed to a public (Jaster & Lanius 2018: 208). In which sense is fake news *fake*? Here, our definition points to the lack of truthfulness exhibited by

---

[5] One challenge is to spell out what exactly the bullshitter is indifferent to (see, for instance, Fallis & Stokke 2017). Virtually unacknowledged in the literature, but even harder to answer is the challenge of spelling out what the bullshitter's attitude of indifference amounts to in the first place.

fake news originators: Fake news originators actively try to deceive the public or do not care about the truth of their reports. How this turns their news reports into "fake" news reports is not entirely transparent, though. Other definitions account for the "fake" aspect more straightforwardly: On Fallis and Mathiesen's understanding, for instance, fake news is not news of sorts, but mere fakes of news in deceitfully presenting as such.

How does the definition fare with respect to its pragmatic transparency, i.e. its capability to explain the politicized use of "fake news"? On the one hand, it is descriptive in the sense that it provides truth conditions for statements about fake news and renders them principally verifiable. At the same time, however, it is part of the truth conditions of fake news statements that the originator of the report is untruthful and thus exhibits a highly problematic, and indeed objectionable, attitude. "Fake news" turns out to be a thick concept: Accusing someone of distributing fake news entails a negative evaluation of that agent's behavior (Jaster & Lanius 2018: 8). This explains the derogatory use of the term "fake news" and its proneness to instrumentalization in political discourse. When Trump accuses the New York Times of distributing fake news, he accuses them of untruthfulness. Most other extant definitions allow for the same point: On Mukerji's account, the accusation will be bullshitting, for Rini it will be deceitfulness, for Gelfert deliberate misleadingness, and for Fallis and Mathiesen fakery. Any definition with an inbuilt negative valence will account for the politicized use of "fake news."[6]

How epistemically transparent is the definition overall? Of course, the mentioned conceptual subtleties affect its epistemic transparency. How easily fake news is identified depends, in large part, on the ease with which the concepts in the definition can be applied to real-life cases. At the same time, this need not require the full grasp of each concept's subtleties. Often, real-life cases will be clear instances of misleadingness or indifference to the truth, despite subtleties on the conceptual level. How well various extant definitions serve the project of identifying fake news seems to be a question for empirical research.

There is a real drawback to the definition that concerns epistemic transparency: In requiring a certain attitude on the originator's part, it will often be hard to establish whether a report is fake news or not. Clearly, this epistemic intransparency has serious ramifications for the concept's integrability and operationalizability, as we will discuss below. As Grundmann points out, this difficulty lies with many extant definitions:

> either (i) fake news is based on an intention to deceive about the content (Rini 2017), or (ii) it is based on an attitude of indifference to truth (Mukerji 2018), or (iii) it is based on an intention to deceive about its status as news (Fallis and Mathiesen 2019), or (iv) it is based on either an intention to deceive consumers about the content of the story or an attitude of indifference to its truth (Jaster and Lanius 2018).

Grundmann's own definition appears to be the only one on the market that attempts to get by without any reference to the propagator's attitude. On his account, fake news is "news that is produced or selected in general ways such that it has the robust disposition to lead, at the

---

[6] Other definitions such as Dentith's (2017, 66) treat fake news directly as "an allegation that some story is misleading – it contains significant omissions – or even false – it is a lie – designed to deceive its intended audience." While being pragmatically transparent, it seems to aim at a quite different phenomenon, and was thus dropped as a definition of "fake news" in later works by Dentith.

time of publication and under normal conditions, to a significant amount of false beliefs in a significant number of the addressed consumers" (Grundmann 2020: 8). While Grundmann's definition scores highly on epistemic transparency, it has deficits with respect to some crucial theoretical requirements such as similarity and exactness. For instance, when is a report that is disposed to generate false beliefs produced or selected i*n general ways*? Consider the case of a journalistic error due to false testimony by sources that the journalist rightfully took to be independent and reliable. Intuitively, the resulting report should not come out as fake news. At the same time, the way in which the report is produced – relying on false testimony by sources one rightfully takes to be independent and reliable – is clearly disposed to generate false beliefs in the audience.

Testing for Explanatory Relevance

When it comes to "addressing the explanatory challenge of why fake news has become such a force in the online world" (Gelfert 2018: 101f.) and the evaluation of the dangers of fake news to society, our definition has a mixed balance. There are various problems attributed to fake news that are not directly reflected in our definition. For instance, it does not become clear from our definition why fake news is a more serious threat now than twenty years ago. Related to this, our definition may even obscure the societal challenge posed by fake news.

Based on these grounds some have cautioned against the use of the term "fake news" if conceived in such a manner. Bennett and Livingston (2018: 124) find that "it tends to frame the problem as isolated incidents of falsehood and confusion", which obscures the real threat from disinformation and propaganda as a systematic attempt to undermine democracy. In the same vein, Habgood-Coote (2022: 501) claims that speaking of "fake news" "offers an oversimplified picture of the problems we face". And, indeed, the definition seems to locate the problem of fake news entirely on the individual level and neglect the systematic features of what may be called "post-truth era" – in contrast to definitions such as Gelfert's and Grundmann's respectively, which focus on the effects of fake news, or Rini's, according to which fake news is necessarily transmitted with the goal of being widely re-transmitted (Rini 2017: E45).

At second glance, the shortcoming of our definition may be less severe, however. For while it does not explicitly address the novelty of "post-truth" dynamics in the internet age, it emphasizes that its propagators consciously violate two norms that are considered essential to deliberative democracy, namely, the norm of truth and the norm of truthfulness. Starting off from this insight, one might work easily enough toward a theory of fake news that addresses also the structural issues related to "post-truth" (Jaster & Lanius 2021: 38f.). Apart from this, it is anything but certain that many of these issues are as novel as they are made out: In our view, the phenomenon of fake news is not new at all – what's new is merely the scale of the problem it generates (Jaster & Lanius 2019: 46f.). Nevertheless, the definition is obviously not as directly relatable to the perceived problem of fake news as other definitions.

Testing for Integrability and Operationalizability

Due to its less than optimal epistemic transparency, our definition has also deficits on the level of operationalizability and integrability. The need to assess the mental attitudes of fake news

propagators is a hindrance to finding out whether some particular news report is fake news. This in turn makes it difficult to use the definition for empirical research and legislative endeavors. It is almost impossible to systematically measure the intentions of fake news propagators, and even the falsity or misleadingness of fake news is often determined only by reference to fact checking websites due to the difficulties of automatically measuring such a context-dependent variable.[7]

As a result, what is often studied in the empirical research on fake news is not fake news according to most definitions in the literature (including ours) or what is commonly understood by the term. Instead, what media or communication scientists actually study is false news, misleading news, misinformation, untruthful speech, information or news that is labeled "fake news" by fact checking organizations or news with certain, for instance, stylistic characteristics. Most probably, our definition is not particularly suitable for the kind of empirical research currently done under the label "fake news."

At least from a conceptual point of view, we take it, however, that there is no real option of getting rid of some requirement like lack of truthfulness. Defining "fake news" with the sole goal of easily measuring something more or less similar to the phenomenon called "fake news" all too often results in a substantial change of topic. And, as a consequence, we do not only speak past each other, but also lose track of the problems that we set out to solve in the first place.

Testing for Practical Goals

When talking about the practical goals tied to a definition of fake news, we mentioned the goal of enabling constructive political and societal discourse about the phenomenon and its dangers as well as the legal regulation of fake news. We also mentioned the goal of strengthening citizens' capabilities for recognizing and effectively debunking fake news. As far as a definition can further these goals at all, the success of this endeavor will arguably depend on the definition's epistemic fruitfulness, most prominently its epistemic transparency and explanatory relevance. The concept should meaningfully relate to legal concepts already in place and allow for easy use so that politicians, NGOs, journalists, educators, and citizens are able to understand and address the dangers of fake news to individuals and societies.

We have already pointed out that, due to their lack of transparency, definitions which stress the propagator's attitude may not be easily put to work in legislation. When it comes to the political or societal discourse about the threats and dangers of fake news as well as educative projects strengthening citizens' capabilities in debunking fake news, however, it seems that our definition has some merits: The definition draws attention to the importance of the norms of truth and truthfulness – norms that appear to be essential to deliberative democracy (Jaster & Lanius 2021: 38f.). By emphasizing the importance of these norms and the ways in which they may be undermined to further political and ideological goals, our definition can be expected to steer public discourse about fake news to a core problem of our time and direct citizens' attention to the crucial difference between good and bad reporting.

---

[7] This might change in the near future, however, due to the steady progress in the training of large language models.

How well our definition – or any definition – scores overall on the level of practical goals is hard to say, however. The practical goals are general in nature and most definitions will arguably contribute in *some* ways to these goals. At the same time, the contribution of a particular term, concept, or definition to each of the practical goals will at best be indirect and therefore notoriously hard to evaluate. We take it that much more philosophical and empirical research would be required to even estimate the overall suitability of any definition of "fake news" with respect to practical goals in a reasonable way.

Wrap-up and Outlook

Our aim in this paper was to come up with a more refined set of criteria for the evaluation of competing definitions of "fake news." As we have argued, the existing proposals can be assessed with respect to various theoretical, epistemic, and practical goals. We also saw that there are considerable trade-offs between the various goals one may pursue with a definition of "fake news." Because of this, different definitions may be more or less suitable. Competing definitions will fare varyingly well, depending on the weight one puts on one requirement or the other. Consequently, different definitions may be better or worse suited in different contexts of inquiry: A context in which practical regulatory questions are at stake may favor a different definition of "fake news" than a context which focuses on a purely theoretical taxonomy of fake news and other forms of insincere speech.

What are we to make of that? Are definitions of "fake news" only better or worse in relation to a specific epistemic inquiry or particular practical goal? One may think that this leads to a relativistic approach to defining "fake news": Various competing definitions of "fake news" will be equally legitimate relative to different theoretical, epistemic, or practical goals. Indeed, for some research projects or debates a specialized definition of "fake news" will be most appropriate. Given the interdisciplinary relevance of fake news, however, an important desideratum is also to establish one concept that is suitable for interdisciplinary exchange. Thus, conceptual work on fake news should be guided by the attempt to integrate the various goals as best as possible. For reasons of space, doing so is beyond the scope of this chapter. Integrating the various goals would require the systematic comparative evaluation of extant definitions on the market. All we are able to provide here are the guide-rails for a more thorough methodological discussion.

A result from this chapter is that once we move away from traditional conceptual analysis, a variety of new research questions in experimental philosophy emerge. Within a conceptual analysis approach, the role of empirical research is somewhat limited, as "most research in experimental philosophy makes use of a collection of closely connected methods that in some way involve the study of intuitions" (Knobe & Nichols 2017). More specifically, x-phi has been primarily concerned with the empirical basis for claims about linguistic intuitions and the similarity of the defined concept to the original one: Does a particular definition really conform to our linguistic intuitions better than the others? Are the paradigmatic cases really as clear as assumed? A natural task for standard approaches in x-phi would be to investigate, for example, whether competent speakers are inclined to count Grundmann's filter algorithms as instances of fake news or not.

Focusing on different sets of criteria for a suitable definition and the potential interplay of these criteria opens up an even broader spectrum of empirical questions. For instance, x-phi may investigate the epistemic transparency of the various definitions of "fake news". How easily do people understand and apply various definitions? How difficult do people find it to decide whether an individual instance of reporting falls under the definition or not? Moreover, there could be fruitful empirical research about the operationalizability of the various definitions of "fake news" and their integrability into extant theories. How well do various definitions serve the project of measuring the effects of fake news? How smoothly can they be integrated into the research programs in various disciplines? Finally, it is a task for empirical research to determine how well-suited the proposed definitions are for achieving various practical goals: What kind of definition contributes most effectively to the legislation of fake news? What definition is best suited to help people debunk fake news?

As we hope to have shown, there is a wide array of empirical questions arising within a conceptual engineering-paradigm that go beyond the more traditional research program of x-phi. There are theoretical, epistemic, and practical goals that can be pursued when defining the concept of fake news, which requires trade-offs in evaluating the overall merit of any definition. Some decisions when weighing the various requirements will have to rely on facts that may only be answered empirically. Here, x-phi can step in and enhance the debate about the definition of the concept of fake news.

References

- Belleri, D. (2021): Downplaying the change of subject objection to conceptual engineering, Inquiry: 1–24.
- Benkler, Y., Faris, R. & Roberts, H. (2018): Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics. New York, NY: Oxford University Press.
- 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.
- Brown, E. (2019): "Fake news" and conceptual ethics, Journal of Ethics and Social Philosophy, 16(2): 144–154.
- Cappelen, H. (2018): Fixing language: An essay on conceptual engineering. Oxford: Oxford University Press.
- Carnap, R. (1950): Logical Foundations of Probability. Chicago: University of Chicago Press.
- Carson, T.L. (2006): The Definition of Lying, Noûs, 40(2): 284–306.
- Chalmers, D. (2020): What is conceptual engineering and what should it be?, Inquiry.
- Coady, D. (2019): The Trouble With "Fake News", Social Epistemology Review and Reply Collective 8(10): 40-52.
- Dentith, M.R.X. (2017): The Problem of Fake News, Public Reason, 8(1-2): 65–79.
- Fallis, D. (2012): Lying as a Violation of Grice's First Maxim of Quality, Dialectica, 66: 563–581.
- Fallis, D. and Mathiesen, K. (2019): Fake news is counterfeit news, Inquiry: 1–20.

- Fallis, D. and Stokke, A. (2017): Bullshitting, Lying, and Indifference toward Truth, Ergo, 4(10): 277–309.
- Faulkner, P. (2013): Lying and Deceit, in H. Lafollette (ed.) International Encyclopedia of Ethics, Hoboken: Wiley-Blackwell, 3101-3109.
- Frankfurt, H.G. (1986): On Bullshit, Raritan Quarterly Review, 6(2): 81–100.
- Gelfert, A. (2018): Fake News: A Definition, Informal Logic, 38(1): 84–117.
- Grundmann. T. (2020): Fake news: the case for a purely consumer-oriented explication, Inquiry, DOI: 10.1080/0020174X.2020.1813195.
- Habgood-Coote, J. (2018): Stop Talking about Fake News!, Inquiry: An Interdisciplinary Journal of Philosophy: 1–33.
- Habgood-Coote, J. (2022): Fake news, conceptual engineering, and linguistic resistance: reply to Pepp, Michaelson and Sterken, and Brown, Inquiry, 65(4): 488–516.
- Hale, V. (2017): Revealed: 1,000-Man Mob Attack Police, Set Germany's Oldest Church Alight on New Year's Eve. Breitbart:: https://www.breitbart.com/europe/2017/01/03/dortmund-mob-attack-police-church-alight/.
- Haslanger, S. (2000): Gender and Race: (What) Are They? (What) Do We Want Them To Be?, Noûs, 34(1): 31–55.
- Isenberg, A. (1973): Deontology and the Ethics of Lying, in A. Berleant and A. Isenberg (eds.) Aesthetics and the Theory of Criticism. Selected Essays of Arnold Isenberg, 245–64. Chicago: University of Chicago Press.
- Jaster, R. and Lanius, D. (2018): What Is Fake News?, Versus, 127(2): 207–227.
- Jaster, R. and Lanius, D. (2019): Die Wahrheit schafft sich ab: Wie Fake News Politik machen. Stuttgart: Reclam.
- Jaster, R. and Lanius, D. (2021): Speaking of Fake News. Definitions and Dimensions, in S. Bernecker, A. Flowerree, T. Grundmann (eds.) The Epistemology of Fake News, Oxford: Oxford University Press: 19–45.
- Knobe, J. and Shaun N. (2017), Experimental Philosophy, in E. N. Zalta (ed.) The Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/archives/win2017/entries/experimental-philosophy.
- Koch, S. (2021): Engineering what? On concepts in conceptual engineering. *Synthese*, 1991(1): 1955-75.
- Lackey, J. (2013): Lies and deception: an unhappy divorce, Analysis, 73: 236–248.
- Lanius, D. (2021): Wie sollten Lehrende mit Fake News und Verschwörungstheorien im Unterricht umgehen?, in J. Drerup, M. Zulaica y Mugica & D. Yacek (eds.) Dürfen Lehrer ihre Meinung sagen? Demokratische Bildung und die Kontroverse über Kontroversitätsgebote. Stuttgart: Kohlhammer Verlag: 188–208.
- Mukerji, N.S. (2018): What is Fake News?, Ergo, 5: 923–946.
- Oremus, W. (2016): Stop Calling Everything "Fake News." Slate: https://slate.com/technology/2016/12/stop-calling-everything-fake-news.html.
- Pepp, J., Michaelson, E. and Sterken, R. (2022): Why we should keep talking about fake news, Inquiry, 65(4): 471–487.
- Rini, R. (2017): Fake News and Partisan Epistemology, Kennedy Institute of Ethics Journal, 27: 43–64.
- Saul, J. (2012): Lying, Misleading, and What Is Said, Oxford: Oxford University Press.

- Stokke, A. (2013): Lying, Deceiving, and Misleading, Philosophy Compass, 8: 348–359.
- Subramanian, S. (2017): The Macedonian Teens Who Mastered Fake News. Wired: https://www.wired.com/2017/02/veles-macedonia-fake-news/.
- Sundell, T. (2020): Changing the Subject, Canadian Journal of Philosophy, 50(5): 580–593.
- Zimmermann, F. and Kohring, M. (2018): „Fake News" als aktuelle Desinformation. Systematische Bestimmung eines heterogenen Begriffs, M&K Medien & Kommunikationswissenschaft, 66(4): 526–541.