

Checking PolitiFact's Fact-Checks

Sakari Nieminen & Valtteri Sankari

To cite this article: Sakari Nieminen & Valtteri Sankari (2021) Checking PolitiFact's Fact-Checks, *Journalism Studies*, 22:3, 358-378, DOI: [10.1080/1461670X.2021.1873818](https://doi.org/10.1080/1461670X.2021.1873818)

To link to this article: <https://doi.org/10.1080/1461670X.2021.1873818>



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



Published online: 19 Jan 2021.



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



Article views: 771



View related articles [↗](#)



View Crossmark data [↗](#)

Checking PolitiFact's Fact-Checks

Sakari Nieminen and Valtteri Sankari

Department of Philosophy, Contemporary History and Political Science, University of Turku, Turku, Finland

ABSTRACT

In this article we examine PolitiFact's fact-checking process. We collect a random sample of 858 fact-checks and evaluate them in the light of criteria based on or inspired by fact-checking literature and the International Fact-checking Network's code of principles. Our analysis reveals the following: in general, PolitiFact fared well. However, from the point of view of the criteria, its practices leave room for improvement. The biggest issue is complex propositions. These are statements containing multiple claims, i.e., more than one proposition. In 279 cases (33% of our sample), PolitiFact checks a complex proposition and assigns one truth rating to it. This is problematic as the reader might misinterpret the truthfulness of an individual claim. PolitiFact also checks claims that we considered uncheckable. These are statements whose truthfulness cannot be defined in practice, e.g., claims about the future and vague claims. In 92 cases (11% of our sample), PolitiFact checked a claim like this. The article ends with a discussion about the limitations of the criteria used here.

KEYWORDS

Fact-checking; journalistic practice; political journalism; PolitiFact; truthfulness; truth-value

Introduction

The fact-checking industry has grown remarkably in recent years and has become firmly established in the U.S. media field (Graves 2016). Many organizations conduct fact-checking routinely, not only during elections. The number of fact-checking projects has also increased: in 2014, there were 44 active fact-checking projects (Stencel 2017), while the number in November 2019 was 226 (Stencel and Luther 2019). The international fact-checking movement is also emerging. The International Fact-Checking Network (hereafter IFCN) was launched at the Poynter Institute in 2015 (Mantzaris 2015), and in 2016, this network published fact-checking's code of principles (Mantzaris 2016).

Fact-checking means evaluating the truthfulness of claims presented in public. The purpose is to discover and publish whether a claim is accurate or not. This is based on various information sources, such as scientific studies, experts and official statistics. Usually, fact-checkers are interested in claims made by politicians and other influential actors. In the American context, fact-checking is mainly a journalistic practice, but globally there is more diversity (Graves 2018).

CONTACT Sakari Nieminen  saanni@utu.fi

© 2021 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One of the most famous fact-checking institutions is American PolitiFact. It started its work back in 2007 (Drobnic Holan 2018a), and is nowadays a well-established actor in U.S. political discourse (Graves 2016, 8). PolitiFact won a Pulitzer Prize in 2009 (Ibid), and its operations are widespread: besides the national institution, PolitiFact also has 23 state editions, 14 of which are currently active (The number and the activity status are from Duke Reporter's Lab database: <https://reporterslab.org/fact-checking/#>, cited December 9, 2020). PolitiFact has also served as a model for many later fact-checking outlets and it has directly advised organizations in Africa, Australia, Europe and South America (Graves 2018).

Fact-checking has great potential in the field of communications. It has been shown, for example, that legislators make fewer false statements when they are reminded of the potential damage that negative fact-check ratings could cause to their reputation (Nyhan and Reifler 2015). Nevertheless, perhaps the most significant thing that fact-checking could offer is its capability to convey information about statements appearing in public. This is what many fact-checkers themselves mention as their main goal: to inform their audience (Amazeen 2013, 5).

If done poorly, however, this goal could suffer. Critics have pointed out that some of the fact-checkers' methods might create misleading impressions about the truthfulness of the claims checked (Uscinski and Butler 2013, 167). There has also been concern that if different fact-checkers give dissimilar ratings, the audience could be left confused and uncertain about what to believe (Marietta, Barker, and Bowser 2015).

In this study, we focus on the processes and practices of PolitiFact. We compile a set of 24 criteria inspired by and based on fact-checking literature and the IFCN's code of principles, and examine how PolitiFact's fact-checks compare with these criteria. PolitiFact is a justifiable object of study as it is one of fact-checking's flagships and a model for many other fact-checkers.

The structure of this article is as follows: first, we discuss previous literature relating to our study. Then we introduce the criteria that we are going to use and present the research question in its final form. After that, we outline the method of our study and present the results of our analysis. The article ends with a general discussion.

Previous Research

Fact-Checking Literature

So far, most of the fact-checking research has focused on either the effects of fact-checking or fact-checking as a profession (Nieminen and Rapeli 2019). For our study, this last branch of research is more relevant, so we will focus solely on it. Graves (2016) outlines (among other things) the history and development of fact-checking. Even though the recent emergence of fact-checking outlets is considered as a new phenomenon, politicians' statements have long been refuted, e.g., on op-ed pages. He depicts that one of the motivations behind this current fact-checking movement was due to disappointment in conventional journalism's ability to dispute politicians' false assertions (Graves 2016, 36–63). Lowrey (2017) has also analyzed the development of fact-checking as well as its legitimacy. Since 2010 the fact-checking world has become more fragmented, and the legitimacy of the practice has increased over time.

Graves (2017) examines the methods of fact-checking. Based on his fieldwork among three U.S. fact-checkers, Graves identifies five elements that a typical fact-check contains: (1) Firstly, fact-checkers choose which claims to check for, considering e.g., the news-worthiness and political significance of the claim. (2) Secondly, fact-checkers often give the author of the claim a chance to present their account. (3) Thirdly, fact-checkers try to trace the original source and spreading of the claim. (4) Fact-checkers also repeatedly consult experts for interpretation of the data, which is sourced from nonpartisan sources, preferably government agencies (5) Finally, fact-checkers make the verification process public (steps, sources, etc.) for the sake of transparency. However, critics have questioned the reliability of the fact-checking method, claiming that it leaves too much room for subjectivity when rating claims (Uscinski and Butler 2013).

Critics have also rebuked fact-checkers for judging claims that cannot be fact-checked (e.g., claims about the future) or are difficult to fact-check (e.g., causal claims). The tendency of fact-checkers to pick and choose the claims has also raised criticism, because without systematic case selection, fact-checkers' ratings might create false impressions about the overall honesty of politicians. (Uscinski and Butler 2013; see also Uscinski 2015). This last criticism has been responded to by noting that the main purpose of fact-checking is to provide information about public claims, not to tell, "who lies most" (Amazeen 2015).

Graves (2018) examines fact-checking as a global phenomenon. Variation exists, for example, in rating systems, lengths of fact-checks, and in whether to contact the speaker. Many fact-checkers place the claims they check in pre-existing categories, like "true", "half-true" and "false". However, some reject these rating systems, considering them reductive and unscientific. Length is another divider; some fact-checkers prefer brief pieces while others favor lengthy explanatory articles. Approaches also vary on whether to contact the author of a claim and let them present their version. Even though Graves (2017) identifies this as an element of a typical fact-check, the issue seems to be somewhat debated. The practice is justified by thoroughness and fairness. However, opponents state that contacting might open the door for subjectivity and political inference (Graves 2018, 624–626).

There are also studies that have scrutinized the fact-checks of PolitiFact. Uscinski and Butler (2013) critically analyzed PolitiFact's work to highlight flaws in the methods and practices associated with fact-checkers. Marietta, Barker, and Bowser (2015), Amazeen (2016) as well as Lim (2018) have examined whether PolitiFact checks the same claims as other major fact-checkers and whether they reach the same conclusions. Amazeen found consistency between fact-checkers ratings, while Marietta et al. found that the agreement between fact-checkers varies from topic to topic. Lim, on the other hand, found that fact-checkers agree well in obvious truths or falsehoods but that the agreement rate is lower in more ambiguous statements. Bucciol (2018) utilized PolitiFact's fact-checks to analyze what the background variables (such as the speaker's age, gender, party, etc.) of false claims look like.

Our study is positioned on this "fact-checking as a profession" branch of research, as we evaluate how one particular fact-checker, PolitiFact, conducts its fact-checks. Our study can be justified from the point of view of Lowrey's (2017) results. As fact-checking's legitimacy has increased (reflected by more frequent fact-check citations and less often questioned motivations), it becomes equally important to examine how fact-checking and its different elements (e.g., what kind of claims are checked) are carried. Our study also has interfaces to the fact-checking critique, as it has inspired some of the criteria we are going to use.

Studies on Accuracy and Verification

The evaluation and description of journalistic practices and output is not a new thing. For example, academics have been analyzing the accuracy (referring here to a lack of errors) of news articles at least since the 1930s (Charnley 1936, for more recent studies, see e.g., Maier 2005 and Porlezza, Maier, and Russ-Mohl 2012). Journalistic verification is another issue that has attracted scholarly attention. Brandtzaeg et al. (2016), for instance, interviewed journalists to discover how they verify social media content, while Godler and Reich (2017) analyzed journalistic cross-verification practices and how particular factors relate to the utilization of cross-verification. Rauchfleisch et al. (2017) examined journalists' verification practices on Twitter during the Brussels attacks in March 2016. They found that many journalists did not follow proposed best verification practices. A great deal of tweets requesting information or content did not try to verify it.

Both of these aforementioned aspects could be studied, when evaluating the work of fact-checkers. One could examine whether the fact-checks contain any errors or assess whether the truth-value judgments are justified. Alternatively, one could appraise whether the sources cited by fact-checkers are of high-quality and how the sources' information conveys into truth-values. We do not analyze these issues here as our focus is on claim choosing and the checking process in general. However, there is some interface with these kinds of studies and our work, as we are evaluating the actual work of (fact-checking) journalists.

Criteria for Journalistic Quality

Compiling quality criteria for journalism is not a new thing either. Rögner and Wormer (2017), for example, developed criteria for good environmental journalism based on literature reviews, environmental journalists' views and journalism students' opinions. Their criteria highlight things like the transparency of sources, use of evidence and lack of scaremongering and trivialization. In a similar fashion, Ashoorkhani et al. (2017) outline a checklist for health research journalism.

Shapiro, Albanese, and Doyle (2006) survey judges in journalism award programs in order to discover the criteria they have for excellence in journalism. Things like writing quality, reporting rigor, independence and fairness came up as indications of journalistic excellence. Ellis (2012), on the other hand, analyzes various codes of ethics in Anglo-American journalism and thus illustrates what journalists themselves consider good practices.

Our study could be considered as a part of this research tradition, as we compile criteria for the evaluation of political fact-checking. However, our motivations are empirical. We are compiling the criteria because we want a tool for analyzing PolitiFact's work, not for the sake of the criteria itself. In addition, our viewpoint is narrower than the one in those more general checklists. Our interest is in the claim choosing and the checking process, not in the output. This is also one of the reasons why we compile the analysis tool ourselves, instead of adapting the existing ones.

Criteria and Research Question

We have used the criteria we compile here in publications aimed for our national audiences (Nieminen and Wiberg 2018a, 2018b; Sankari and Wiberg 2019). Those works are

more empirically oriented, and do not discuss the criteria in detail. Thus, this is the first time our criteria are justified at this length. The criteria are inspired by and based on fact-checking literature and the IFCN's code of principles.

The IFCN's code of principles is one of the bases, since it is drafted by the fact-checking industry itself (IFCN 2020a), and thus reflects practices that fact-checkers themselves consider legitimate.

A methodological criticism of fact-checking (Uscinski and Butler 2013; Uscinski 2015) is considered, since it introduces a scientific perspective to the issue. This criticism raises important points about what some scholars consider to be problematic in fact-checking methodology and practice in general.

What fact-checkers check is also significant. Statements are only a small fragment of political communication and not all statements are verifiable. Hence, it is important to ponder the cases where a truth-rating is not possible or meaningful. The methodological criticism highlights few problematic aspects (Uscinski and Butler 2013, 170–175), but this viewpoint could be wider. Therefore, the criteria, in this respect, are based on the work of Makkonen and Wiberg (2017) as well as Nieminen, Raiskila, and Wiberg (2017, 60; 186–210), wherein they reflect on the kind of speech that is beyond the reach of fact-checking. Merpert et al. (2018) is also an interesting study related to the issue, in which they study how capable citizens are in differentiating checkable claims from uncheckable claims. However, they are not listing uncheckable statement types, in particular, so their pondering is not the basis of the criteria.

The criteria compiled here are only designed for an analysis of the claim choosing and checking process and ignore certain important aspects of fact-checking. However, the ignored aspects are not necessary for our analysis. They deal with whether the claims are rated correctly, the quality of the evidence, as well as matters concerning the transparency of funding, organization and the correction policy.

Below are the 24 criteria we use to evaluate fact-checks made by PolitiFact:

- (1) Complex propositions are divided into parts
- (2) Claims are checked in their immediate context
- (3) The fact-checked actor is not accused of lying, unless there is evidence of the intention to deceive
- (4) Each claim is checked using the same criteria
- (5) Fact-checks are not concentrated on any one particular side (e.g., party)
- (6) The fact-checker does not advocate for policy positions
- (7) Fact-checks are based on evidence
- (8) Fact-checking sources are published
- (9) Fact-checking methodology and the definitions of possible truth-value categories are published

The following sentence types are considered uncheckable:

- (10) Claims about the future
- (11) Ambiguous or vague claims
- (12) Claims about non-existent things (e.g., "The current king of France is bald")
- (13) Liar sentences ("I am lying")

- (14) Claims containing aesthetic, moral or ethical values
- (15) Claims concerning the supernatural or matters of faith
- (16) Claims referring to personal experience
- (17) Tautologies (“All bachelors are unmarried”)
- (18) Performatives
- (19) Figures of speech
- (20) Opinions
- (21) Goal setting
- (22) Interrogative sentences
- (23) Non-declarative exclamatory sentences
- (24) Imperative sentences

Joseph Uscinski’s and Ryden Butler’s (2013) critique of fact-checking methodology inspires the first three criteria. They criticize fact-checkers for occasionally compounding multiple statements into one factual claim and then assigning a “meta-rating” to it. According to Uscinski and Butler (2013, 166–168), this practice is sometimes misleading. If a politician makes a statement containing five different propositions and the fact-checker bundles them all together, providing one singular rating (like half-true), an audience may think that all of these propositions are half-true, even though some of them might be false and others true. Hence our first criterion.

The first criterion means that claims containing multiple propositions should be divided into separate fact-checks, each containing exactly one claim. That is, different propositions should not be compounded under one single truth rating. If a politician claims that his opponent A is an ex-convict and an immigrant, instead of one fact-check (“A is an ex-convict and an immigrant”) and one truth-value rating, there should be two separate fact-checks (“A is an ex-convict”; “A is an immigrant”) with two separate ratings.

However, our first criterion is not entirely in line with Uscinski’s and Butler’s critique (ibid.), since they also think that for the sake of contextualization, it might be necessary to aggregate multiple claims into one. Additionally, they argue that since the decision on whether to compound multiple claims depends on context, and because the right context in politics is contestable, fact-checking is highly subjective.

While we are not claiming that Uscinski and Butler are wrong when saying that sometimes compounding is necessary, we still believe that complex propositions should be divided. Although it is occasionally unavoidable to compound elements from separate propositions to be able to check what the speaker is implicating, even then the checked content could (usually) be reformulated into a single proposition. Imagine a situation where a fact-checker examines the following comparison: “*Politicians’ average salary is X, while teachers’ average salary is Y*”. Both numbers could be accurate, but the fact-checker considers them incomparable and thinks that the proper rating is half-true. If the claim is checked in a form, “*Politicians’ average salary is X, while teachers’ average salary is Y: half true*”, this might leave the audience with a wrong impression that X and Y are not the actual average salaries. Reformulating this claim “*Comparison of politicians’ and teachers’ average salaries: half true*”, reduces this risk.

The second criterion relates to the importance of context. According to Uscinski and Butler (2013, 167–168), context matters in fact-checking, but the right context for the

claim is not “the surrounding words of the speaker but the objective reality represented by the factual claim” and this context is usually legitimately contestable in politics. Thus, fact-checkers’ ratings are inevitably based on subjective notions about the right context. While Uscinski and Butler might be right about the subjectivity of fact-checking, we do not think that this necessarily eradicates the legitimacy of fact-checking, as long as the judgments made are justified for the audience, based on evidence and the sources are transparently specified (for a more detailed discussion about the subjectivity, justification and epistemological critique of fact-checking, see Graves 2017, 520–523). Though their argument about the significance of the context is reasonable, we understand the context here in a narrower way. Uscinski and Butler seem to hold a broad notion of context, but our interests are more immediate: who says what, where, how and what else is said. All of these factors are things to consider when judging the truthfulness of a claim, so that humorous flings, for example, are not fact-checked or that fact-checkers do not ignore surrounding words that would change the interpretation. Hence, our second criterion.

The third criterion is based on a small (though significant) side note that Uscinski and Butler (2013, 171–172) make. They mention a root of the tendency (among some fact-checkers) to classify supposedly inaccurate statements as lies. They remind that lies are intentional falsehoods and that there is rarely evidence of such intention. This distinction is important, since it is one thing to say that a politician is not correct, but another to allege that they are intentionally spreading falsehoods. Therefore, the third criterion.

Uscinski and Butler (2013, 168–170; 172–175) also point out the problems of fact-checking causal claims and the overall leeway fact-checkers have when defining truth-ratings. We do not include these remarks to our analysis, because we do not take a stand on whether the truth-value judgements are correct. We are interested in the claim choosing and the checking process, not its product. Their critique about the selection effect (that is, fact-checking might make politicians look more honest or dishonest than they truly are) is neither considered, as the purpose of fact-checking is usually not to establish the biggest liar, but to provide information about claims appearing in public (Amazeen 2015).

Criteria 4–9 are based on the IFCN’s code of principles (IFCN 2020a). These principles require signatories to fact-check every claim using the same standard and not to focus the fact-checks on any one particular side. This is the basis for our criteria four and five. The code also prohibits the signatory organizations to advocate for policy positions on the fact-checked issues. Therefore, criterion six.

The code of principles demands that the evidence should determine the conclusions. To study whether the evidence determines the conclusion comes close to analyzing whether the claim’s rating is correct. Thus, we do not apply this demand as such, but we deduce from it our criterion seven: *Fact-checks are based on evidence*. This means that there must be evidence on which the fact-check is based; fact-checkers should not rate the claims off the top of their heads. The code also requires a disclosure of sources and methodology. Hence, our criteria eight and nine.

Transparency of funding, organization and an open correction policy are also in the code of principles. While these are important tenets for fact-checking, we do not include these to our criteria. This is because transparency of funding and organization

are not part of the checking process as such and any evaluation of correction policies would require us to take positions on whether the truth-value ratings of the fact-checkers are correct. And even though the use trustworthy and ideally nonpartisan sources is an essential component of fact-checking, we do not include this to our criteria either. We believe that to determine whether some particular source is reliable vis-a-vis some claim is more about assessing whether the claim is rated correctly and less about the fact-checking process itself. Additionally, the operationalization of trustworthy sources is difficult.

It should also be noted that the code of principles has been revised since we compiled our criteria. However, the core principles remain the same and thus this does not require changes to the criteria. For the sake of simplicity, we cite the revised version throughout the article.

Criteria 10–24 are based on the work of Makkonen and Wiberg (2017) as well as Nieminen, Raiskila, and Wiberg (2017, 60; 186–210), though Uscinski and Butler (2013, 170–172, 174–175) also discuss the problem of claims about the future and comment on ambiguous sentences. These criteria concern sentence types that cannot be meaningfully fact-checked in practice.

Claims about the future (criterion 10) cannot be fact-checked, since there is no way to know the future beforehand. Even though good predictions may exist, one still cannot say now whether something is true in the future. Claims about the supernatural or matters of faith (criterion 15, e.g., “*Our country loses blessing, if same-sex marriages are permitted*”) and claims referring to personal experience (criterion 16, e.g., “*I am excited about the Super Bowl*”) can neither be checked, since there is no reliable way to verify these. The same is true for claims about non-existent things (criterion 12). Even though information about their existence can be offered, it might be wise for fact-checkers to reformulate these statements only to refer to the actuality of such things (e.g., “*The current king of France is bald*” → “*There is a current king of France*”). Otherwise, they might create an impression that such non-existent things exist.

Ambiguous or vague claims (criterion 11) and figures of speech (criterion 19) can neither be fact-checked, because it is not clear what the claim means, or on what grounds the truthfulness could be defined. Liar sentences, tautologies and performatives (criteria 13 and 17–18) are special cases of language and are not meaningfully checkable. The truth-value of liar sentences is paradoxical from the point of view of fact-checking and tautologies are true by their definition. Performatives are not true or false in the traditional way (see Austin 1962).

Claims containing aesthetic, moral or ethical values (criterion 14, e.g., “*Rembrandt’s paintings are better than Picasso’s.*”) cannot be fact-checked, since these concern personal values. This is also true for opinions (criterion 20). Goal setting (criterion 21) cannot be checked, because no matter how ridiculous or unrealistic one’s goals are, it is still impossible to define them as true or false as such.

While rhetorical questions could be checkable in theory, distinguishing them from interrogative questions might be difficult. This problem is averted by not checking questions. Hence, our criterion 22. Exclamatory and imperative sentences (criteria 23–24) are not checkable, unless some proposition is deducible from them. For example, from the sentence “*Damn the government for cutting pensions!*” one could deduce the proposition that “*The government has cut pensions*”.

For many parts, criteria 10–24 are based on a correspondence-theoretical approach of truth. According to the correspondence theory, a proposition is true if and only if it corresponds to some state of affairs or fact (Goldman 1999, 42–43). The truth is thus a relation between the claim and the world. In order for the claim to be true, it must correspond to the reality, i.e., how things are. Therefore, there must be some way to approach this reality in order to determine whether something is true. That’s why claims about the future, for example, cannot be checked beforehand: there is no reality the claim could yet correspond to. Also consider claims about aesthetic values: the reality does not contain an objective gauge that would make the claim true or false. However, this is not the only way to approach the issue. With a different theoretical approach to truth, our criteria could be different. The discussion about what e.g., a coherence theory-based criteria would look like, is beyond the scope of this study.

We use these 24 criteria to answer the following question:

RQ: Do PolitiFact’s fact-checks comply with the 24 criteria inspired by and based on fact-checking literature and the IFCN’s code of principles?

We are not claiming that this is the only way to evaluate fact-checking. One could, for example, take the efficacy of fact-checking for a starting point and judge whether or not the formats used by fact-checkers are optimal (see e.g., Young et al. 2018 and Ecker et al. (2019) for studies on this subject). Neither do we claim that the criteria necessarily contain all aspects that could be considered in this kind of analysis. However, we feel confident that it can provide meaningful results. We have also had good experiences with it (Nieminen and Wiberg 2018a, 2018b; Sankari and Wiberg 2019).

We are also aware, that there is some circularity in our criteria from the viewpoint of this study. PolitiFact is one of the institutions that helped to establish fact-checking as a practice and a movement (Graves 2016, 8) and it has influenced many later fact-checking outlets (Graves 2018). So, it could be said that PolitiFact has been shaping those practices that many other fact-checkers themselves consider legitimate and fair, and thus those very same practices that the IFCN’s code of principles are based on, which, in turn, is the basis for our criteria 4–9. So, in a way we are analyzing whether PolitiFact still complies with the principles they themselves have been establishing. However, applying these criteria to PolitiFact’s fact-checks is still meaningful, as there could be cases where PolitiFact does not comply with the practices it helped to establish. Additionally, this circularity is not due to the criteria itself, but to our research subject. Using the criteria to analyze the fact-checks of a fledgling outlet, for example, would not result in this kind of circularity.

Method

To answer the research question, we conducted an analysis of a random sample of PolitiFact’s fact-checks.

Data

We collected the data from the PolitiFact website. We chose the option “All” from the website’s Truth-O-Meter menu and then collected links to all fact-checks made by PolitiFact

National between 2 May 2007 and 31 July 2019, adding up to a total of 5,250 fact-checks (including both Truth-O-Meter and Flip-O-Meter). Six web links were broken, so those fact-checks are not included to our analysis or any numbers mentioned here.

Sample

We drew a random sample of 66 fact-checks from each year with SPSS statistical software, resulting in 858 fact-checks that were included to our final analysis, and collected their actual text parts. Then we conducted the analysis.

Coding Procedure

We coded the data ourselves. As both authors have used the criteria in earlier studies, there was no need for pilot testing in this study. The second author went through all 858 fact-checks and the first author independently analyzed 20% of the sample to determine intercoder reliability. We examined the fact-checks one by one and evaluated, criterion by criterion, whether the fact-check in question fulfills the criterion or not. Afterwards, the coders argued for their rulings in the cases where there were coding differences and then decided together whose rulings prevail. Criteria 5 and 9 were exceptions as they are not about individual fact-checks, but PolitiFact in general. In the case of criterion 5, the second author assessed whether or not all the fact-checks were aimed at one particular side. In the case of criterion 9, the authors examined PolitiFact's website to determine whether they could find a description of the fact-checking method.

Our sample contained 12 Flip-O-Meter checks, in which PolitiFact estimates whether a speaker has changed their mind on a particular issue. The intercoder sample contained four of these checks. We decided to include these, because criteria 2 and 4–9 are suitable for analysis of these checks. Both coders agreed unanimously in the four overlapping cases, and none of these checks violated any of our criteria, so the analysis section will concentrate only on the 846 ordinary fact-checks. These Flip-O-Meter checks are neither included in the intercoder reliability calculations discussed below.

No variables were dropped during the analysis, and no reliability problems emerged, apart from those mentioned in the next section.

Measures

Each criterion was treated as a dichotomous variable (1=the fact-check fulfills the criterion; 0=the fact-check does not fulfill the criterion). [Table 1](#) displays the operationalizations as well as the intercoder reliability levels (Krippendorff's alpha) of each variable. To calculate the Krippendorff's alpha, we used the KALPHA macro provided by Hayes and Krippendorff (2007). In the cases of complete agreement between coders, the intercoder reliability was not calculated.

Solving the Reliability Issue

As one can see from the [Table 1](#), there was complete agreement between coders in the case of criteria 2–9, 12–13, 15–19 and 21–24. Intercoder reliability levels for criteria 1 and

Table 1. The operationalizations and the intercoder reliability levels of our criteria.

| Criterion | Operationalization | Intercoder reliability level (Krippendorff's alpha) |
|-----------|--|--|
| 1 | Coders assess whether more than one claim is assigned under one truth-rating. If only one claim is assigned under one truth-rating, then the criterion is fulfilled . If more than one claim is assigned under one truth-rating, or the claim is quoted in a form that contains multiple propositions , even though only one is rated, then the criterion is not fulfilled . However, the criterion is also fulfilled in cases where more than one claim is assigned under a single truth-rating, but (a) coders consider the other claims to be general knowledge, (b) the other claims are not unique propositions but different ways to say the same thing and (c) the rating given to the complex proposition is "True", and in the text part of the fact-check PolitiFact considers each individual claim true. (Note: The condition (c) is added, because even though it might be unclear for the reader regarding which claim(s) the rating refers to (without reading the whole article), at least the fact-check should not cause any serious misinterpretations, as no matter what claims the reader thinks that PolitiFact considers true, they got the right impression.) | .9176 |
| 2 | Coders assess whether the immediate context is accounted for when checking the claim. This means information about who said something, where it was said, how it was said and what else is said. If the immediate context is accounted for , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders, intercoder reliability not calculated |
| 3 | Coders assess whether the speaker is accused of lying. The criterion is fulfilled if the speaker is not accused of lying, or if the speaker is accused of lying and evidence of intent to deceive is presented . Otherwise, it is not. | Complete agreement between coders |
| 4 | Coders assess whether the claim is rated by using the same principles applied to the previous claims of the sample. If the same principles are used, then the criterion is fulfilled . Otherwise, it is not. The first claim to be evaluated under this criterion will always be considered to fulfill the criterion. | Complete agreement between coders |
| 5 | One of the coders assesses whether different entities are targeted for fact-checking or if the fact-checks are concentrated on one entity. If different entities are subjected to fact-checking, then the criterion is fulfilled . Otherwise, it is not. | Not applicable |
| 6 | Coders assess whether a policy position is being advocated in the fact-check. If the fact-checker is not advocating a policy position, then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 7 | Coders assess whether the fact-check is based on evidence, instead of relying solely on the fact-checker's opinions, personal feelings, etc. If the fact-check is based on evidence , then the criterion is fulfilled . Otherwise, it is not. Coders do not assess the quality of the evidence. | Complete agreement between coders |
| 8 | Coders assess whether the sources used in the fact-check are disclosed. If the sources are disclosed , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 9 | Coders assess whether the methodology and the definitions of truth-values are disclosed to the audience. If the methodology and the definitions of truth-values are disclosed , then the criterion is fulfilled . Otherwise, it is not. | Not applicable |
| 10 | Coders assess whether the checked claim is about the future. If the claim is not about the future , then the criterion is fulfilled . The criterion is also fulfilled if the fact-checker checks the claim but does not assign a truth-rating to it and declares the claim uncheckable. Otherwise, it is not fulfilled . Checking estimates does not necessarily violate this criterion if the fact-checker only verifies the existence of an estimate. (Note: Each of our operationalizations for criteria 10–24 contain the | .8337 |

(Continued)

Table 1. Continued.

| Criterion | Operationalization | Intercoder reliability level (Krippendorff's alpha) |
|-----------|--|---|
| | qualification that "if this type of claim is not truth-rated and is declared uncheckable, then the criterion is fulfilled". To avoid unnecessary repetition, we state this qualification explicitly only here.) | |
| 11 | Coders assess whether the checked claim is too ambiguous or vague for the purpose of assigning a truth-value to it. If the claim is not considered too ambiguous or vague , then the criterion is fulfilled . Otherwise, it is not. | .5315 |
| 12 | Coders assess whether the checked claim is about non-existent things (e. g. "The current king of France is bald"). If the claim is not about non-existent things , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 13 | Coders assess whether the checked claim is a liar sentence. If the checked claim is not a liar sentence , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 14 | Coders assess whether the checked claim concerns aesthetic, moral or ethical values. If the claim does not concern any aforementioned values, then the criterion is fulfilled . Otherwise, it is not. | .0000 |
| 15 | Coders assess whether the claim concerns the supernatural or matters of faith. If the checked claim does not concern the supernatural or matters of faith, then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 16 | Coders assess whether the checked claim refers to personal experience. If the checked claim does not refer to personal experience, then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 17 | Coders assess whether the checked claim is a tautology. If the checked claim is not a tautology , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 18 | Coders assess whether the checked claim is a performative. If the checked claim is not a performative , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 19 | Coders assess whether the checked claim is a figure of speech. If the checked claim is not a figure of speech , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 20 | Coders assess whether the checked claim could be considered as a speaker's personal opinion. If the checked claim is not an opinion , then the criterion is fulfilled . Otherwise, it is not. | -.0121 |
| 21 | Coders assess whether the checked claim is a goal set by the speaker. If the checked claim is not a goal set by the speaker , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 22 | Coders assess whether the checked claim is an interrogative sentence. If the checked claim is not an interrogative sentence , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 23 | Coders assess whether the checked claim is an exclamatory sentence from which a proposition cannot be deduced. If the checked claim is not an exclamatory sentence , or it is an exclamatory sentence from which a proposition can be deduced , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |
| 24 | Coders assess whether the checked claim is an imperative sentence. If the checked claim is not an imperative sentence , then the criterion is fulfilled . Otherwise, it is not. | Complete agreement between coders |

10 were quite high, but for criteria 11, 14 and 20, the alphas turn out to be problematic. This is probably due to the rareness of the criteria violations as well as opinion differences between coders on defining which criteria the fact-check in question violates. Sometimes there was agreement that the claim in question was not checkable, but the views about why, were dissenting. In the case of criterion 14, for instance, both coders agreed on all

the fact-checks, except one. The first coder considered that the fact-check violated criterion 14, while the other thought that it violated criterion 20. This dropped the alpha to .0000.

To tackle this problem, we combined criteria 10–24 into one dichotomous variable. This new variable tells whether the claim is checkable, regardless of the reason. It receives a value of “1” if the coder interpreted that the fact-check violated none of criteria 10–24 and it receives a value of “0” otherwise. Therefore, in the case mentioned above, the coders agreed regarding this new variable, because both defined the claim uncheckable, albeit for differing reasons. This combination is justified, since the fact-checked claim is uncheckable if any of criteria 10–24 are violated. Krippendorff’s alpha for this new variable was close to .7 (.6953), so we included it to our analysis. As such, we do not take a stand on criteria 10–24 as separate cases, but comment only on whether the fact-checks were about checkable claims in general.

Before moving forward, we would like to address the issue of potential overlap in our criteria (pointed out by anonymous reviewers), especially in the case of criteria 10–24. We admit that drawing the line between e.g., claims considering ethical values and personal opinions is not a straightforward task. Nevertheless, we think that in the case of criteria 10–24 this is not a colossal problem. For example, the difference between a vague claim and an opinion could sometimes be a viewpoint dependent issue (e.g., how much is “much” in a particular context) but that does not change the fact that there is no reliable way to fact-check that claim. We return to this issue in the discussion part.

Analysis

We look first at the criteria that all the fact-checks fulfilled. In our sample, we observed that no fact-check had serious problems concerning the immediate context. Thus, criterion 2 is satisfied. PolitiFact itself offers details about the immediate context: who says what, where, what else is said, is the claim a response to a question, part of a debate, etc. When we could not check the immediate context ourselves (e.g., due to an outdated link), we ruled these in favor of PolitiFact because of their practice to explain the immediate context.

Criterion 3 is satisfied, because PolitiFact does not accuse any speaker of lying. This is not surprising as PolitiFact itself notes that it does not use the word “lie”, except when defining claims they consider to be the “Lie of the year”. This “nomination” seems not to be actually about lying, because PolitiFact states that this practice is about choosing the most egregious *falsehood* (Drobnic Holan 2018b). Neither do we interpret their “Pants on Fire” rating to refer to lying, since PolitiFact defines this category in a way that does not involve intention (Drobnic Holan 2018a).

We could not determine any blatant inconsistencies between fact-checks, so we consider criterion 4 to be fulfilled. However, this result comes with an asterisk. The way PolitiFact defines its Truth-O-Meter’s categories leaves considerable leeway when determining claims ratings. For example, PolitiFact rates a claim as “True” if “*The statement is accurate and there’s nothing significant missing*” and “Mostly true” if “*The statement is accurate but needs clarification or additional information.*” (Drobnic Holan 2018a). Deciding whether a particular missing nuance or detail is “significant” usually requires subjective consideration. Take, for example, Nancy Pelosi’s claim about the House of Representatives

never suing the sitting president. PolitiFact judges it “True”, while noting that the statement understates the clashes between these government branches (Jacobson 2014). Compare this with Donald Trump’s statement about the murder rate increase within U.S. cities, which he claims to be the biggest in 45 years. PolitiFact judges the claim “Mostly true”. According to PolitiFact, the biggest problem with this claim is that it omits the downward trend in murder rates before the recent increase (Jacobson 2016). It is hard to say whether these truth ratings are consistent, because of the subjectivity of significance. Another matter complicating the consistency estimations is PolitiFact’s custom of combining multiple claims, making it difficult to sort out the rating of each individual statement.

There were no problems with criteria 5–9 either. PolitiFact fact-checks claims from Democrats and Republicans as well as from members of other parties and sometimes even from politicians of other countries. PolitiFact also checks chain emails, blog- and Facebook posts, tweets and claims from different organizations and groups. Our sample did not contain fact-checks where PolitiFact would have advocated for policy positions, and each fact-check is based on evidence. The sources are also comprehensively published; so too are their methodologies and the definitions of Truth-O-Meter categories. Complying with criteria 4–9 is not surprising, because PolitiFact is an IFCN code of principles signatory (though their status was “expired” at the time of writing, IFCN 2020b).

Thus, in the light of our criteria 2–9, PolitiFact fared well. However, part of PolitiFact’s fact-checks violated some of our criteria. Checking complex propositions was the most infringed upon criterion. In our sample, there were 279 fact-checks (33% of our sample) containing a complex proposition. This practice is problematic, since people reading only the claim and the truth rating might misinterpret the accuracy of individual claims.

Take, for example, Amy Klobuchar’s claim about millions of Americans needing mental healthcare, while the number of public mental healthcare beds in Iowa is only 64 and the count is similar throughout the country. The statement contains three claims: (1) number of American’s in need of mental healthcare is in the millions, (2) Iowa has 64 public mental healthcare beds and (3) the number of beds in other states is similar to Iowa. PolitiFact rated the statement as “Mostly True” (Austin 2019). If the reader treats this rating as a summary, they might think that all claims are “Mostly True”, which seems not to be the case. PolitiFact writes that an estimated 11.2 million adults were living with serious mental illness in 2017 and that 7.5 million received treatment in the course of the preceding year. Thus, to us the first claim seems to be wholly true, not “mostly”. The number of beds in Iowa is either 64 or more, depending on how they are counted. As for the similarities between states, according to the statistics PolitiFact consulted, there are two beds per 100,000 people in Iowa compared to the nationwide number of 11.7. This difference seems to be quite wide. We are not at all sure whether this is the association people will get when reading that the comparison of Iowa’s number of public mental healthcare beds to nationwide numbers is “Mostly True”. This problem could have been avoided with three different fact-checks.

The second issue from the point of view of our criteria was PolitiFact’s tendency to fact-check claims that cannot be checked in practice. Whether the reason for this is due to personal experience references, future projections or too ambiguous or opinion-like statements, 92 fact-checks in our sample (11%) were about claims that, in theory, are

beyond the reach of fact-checking. One frequent pattern was checking claims about what will happen in the future if some particular policy is implemented, continued, etc., usually based on estimates (see e.g., Farley and Drobnic Holan 2010) or expert opinions (see e.g., Drobnic Holan 2009) about the issue. However, no matter how good the estimate is or how reasonable the expert's opinion might be, there is always at least a theoretical chance that the reality turns out to be something unexpected (this is a tendency that Uscinski and Butler 2013, 171 also noted in their critique). Sometimes PolitiFact even checks claims concerning what would have happened if a policy proposal that did not pass had been implemented (see e.g., Kruzel 2018). In other words, checking claims about a future that did not actualize. (Fact-checks about claims that describe the content of some particular policy proposal are nevertheless checkable, see e.g., Richert 2009).

Another pattern within this variable was checking claims containing terms and expressions that are semantically vague (these are expressions whose truth-conditions are not constant across speakers and contexts, such as *many*, Egré and Icard 2018) or whose boundary values are basically up to the fact-checker to decide. Instances of these are fact-checks taking a stand on what kind of poll numbers reflect a policy's "high unpopularity" (Greenberg 2017), what kind of decrease in ratings could be described as "way down" (Tobias 2017) and what kind of electoral defeat could be called "shellacking" (Jacobson 2010). However, we did not consider every fact-check about claims containing a vague expression uncheckable, if the fact-check focused on a general tendency instead of determining the vague term. Thus, a fact-check about crime trends in Germany (Greenberg 2018), for example, was not considered to infringe our criteria, as PolitiFact checked whether there is an upward trend in crime (there was not) and did not take a stand on when an increase in crime could be considered "Way up".

Here are three more detailed examples about claims that we considered uncheckable. The first example is a fact-check of Mitt Romney's statement about Pakistan's nuclear warhead capability exceeding the capability of Great Britain in the near future. The claim was classified as "Mostly True" (Jacobson 2012). PolitiFact consulted experts on the topic and they agreed that Pakistan could have more nuclear warheads than Great Britain in the future. However, being plausible does not make the claim "mostly" true. Estimates can lend credibility to the claim, but one still cannot know what the future holds.

The second example of an uncheckable claim was Donald Trump's statement that he has been tougher on Russia than former president Barack Obama was (Jacobson and Kruzel 2018). However, there is no agreed upon metric to gauge a politician's toughness toward other countries. While PolitiFact's pondering about the issue is reasonable as such, there is no unambiguous way to determine the claim's truthfulness.

However, the variable of the uncheckable claims also contained some problematic borderline cases. Our third example illustrates such a case. PolitiFact checked Donald Trump's statement where he claims that he has seen ICE liberating towns from MS-13's control. PolitiFact rated the claim "False", because they could not find evidence of these "liberation" operations (Valverde 2018). The fact-check infringes upon our criterion 16, as the claim in the form that PolitiFact quotes is about Trump's sensory perceptions, i.e., what Trump has seen. In principle, a fact-checker could try to refute claims like this by proving that the state of affairs has never actually occurred. Nevertheless, the claim is still about (personal) sensory experience, and the fact-checker cannot confirm in practice

what the speaker has seen. Therefore, despite the fact that the information PolitiFact offered was significant, the claim is still considered uncheckable, though we admit that drawing the line in cases like this is not easy. These borderline cases reflect the fact that occasionally evaluating the verifiability of the claim is a somewhat subjective process. It requires a bit of discretion, and thus justifiable differences of opinion are sometimes possible. For example, one of the reviewers noted that Trump's claim about toughness on Russia could be seen as a borderline case, as "*fact-checkers make reasoned evaluations in many cases where 'no agreed upon metric' exists, and no metric is perfectly agreed-upon.*" Therefore these evaluations should not always be considered as an absolute judgment of a fact-checker crossing the limits of fact-checking, but rather as statements that the fact-checker is moving in a gray area.

Discussion

We examined how well PolitiFact's fact-checks comply with criteria based on and inspired by fact-checking literature and the IFCN's code of principles. We found that PolitiFact fared quite well in general. Its practices have many strong and valid elements: it checks claims in their immediate context, it does not accuse speakers of lying and we found no blatant inconsistencies between ratings. In addition, PolitiFact does not focus its fact-checks on any one particular side, it does not advocate for policy positions in its fact-checks, it bases its fact-checks on evidence and it publishes the sources of this evidence as well as the methodology used and the definitions of Truth-O-Meter categories.

Nevertheless, PolitiFact's practices also contain aspects that are problematic from the point of view of our criteria. Firstly, PolitiFact sometimes checks complex propositions. In our sample, we found 279 fact-checks (33% of cases) where PolitiFact checked multiple statements under one truth rating. We consider this practice problematic, as it leaves room for misinterpretations about the accuracy of each individual claim if people read only the claim and rating, but not the actual text part. The issue could be tackled by checking different parts of the complex proposition in separate fact-checks or by giving a separate truth rating for each individual claim. And if it is unavoidable to compound elements from separate propositions, then a reformulation of the content could be tried in a way that allows what is being checked to be seen as a single proposition.

The case of complex propositions is worth some discussion, though. Firstly, the first criterion is based on a hypothesis that people reading complex propositions and their ratings might misinterpret the accuracy of each individual claim (e.g., consider each claim "mostly true", while some of them are wholly true and others false to some degree) and that separating the complex claim into multiple fact-checks could reduce this risk. We are not aware of any empirical studies on how people interpret the ratings of complex claims in the context of fact-checking (if they read only the claim and the rating).

Additionally (as an anonymous reviewer noted to us), an argument could be made that since many fact-checkers assumingly expect the people to read the whole text, it is not meaningful to criticize fact-checkers for checking complex propositions. There might also be good (journalistic) reasons for publishing the evaluation of a complex proposition in an individual fact-check (e.g., there might be costs involved when publishing separate fact-checks). These arguments are reasonable, but we still believe that our approach to

the complex propositions is justified. Firstly, while fact-checkers might expect people to read beyond the “headline” (claim quote & rating), there might still be consumers who read only the rating and continue no further. Thus, the risk for misinterpretations might still exist. Secondly, it is not guaranteed that a reader who goes through the whole fact-check gets the right impression. In their study of subtle misinformation in news article headlines, Ecker et al. (2014) found that misleading headlines could impair readers’ memories of the article in question and affect their inferential reasoning. They state that misleading headlines can lead to misconceptions and do damage despite a reader’s actual attempt to grasp the article. If this kind of tendency also applies for fact-checking articles (considering the claim & rating as a headline and assuming that the reader considers the rating to apply to each individual claim), then there is a possibility that some readers might misinterpret the rating of separate parts of a complex claim even after reading the whole article.

Another issue regarding our sample was the checking of uncheckable claims. In our sample there were 92 fact-checks (11% of our sample) that we considered to be beyond fact-checking’s reach. The sample contained fact-checks of claims about the future, claims concerning personal experience, claims that were too vague as well as claims that were more opinion-like than factual statements. However, as noted earlier, the criteria compiled here is not the only way to approach uncheckable claims (or the fact-checking in general). With other principles (e.g., ones that prioritize the journalistic aspect of fact-checking), the question of “uncheckable” claims might look different. Additionally, even in the framework of the criteria used here, there are borderline cases and room for justifiable disagreement. Therefore, the determination of the verifiability of a claim requires some subjective judgement.

In addition, in many cases where an uncheckable claim was checked, the information PolitiFact offers is meaningful and significant from a journalistic perspective. When examining a claim about the future, it is reasonable to bring up estimations and projections about the issue, even though the truthfulness of the claim cannot (yet) be defined. And in cases of ambiguous claims, opinions and claims concerning personal experience, it is justifiable to give contextual information and possible approaches to the issue, even though the claim cannot be truth-rated in practice. If one wants to address these issues in the framework of our criteria, then a separate category for uncheckable claims could be worth considering. That would provide a framework to reflect on these claims, devoid of any pressure to judge them as true or false.

As final remarks, some aspects of the criteria as well as lessons from using it are worth discussing. Firstly, while we believe that the criteria offer a meaningful approach to fact-checking and especially to its limitations, we admit that it is somewhat narrow. The purpose of the criteria is to sketch the borders of fact-checking, despite the fact that sometimes there might be good reasons for the fact-checkers to cross these borders. As previously noted, e.g., in the case of complex propositions, there might be legitimate journalistic reasons for a fact-checker to publish only one fact-check, and in the case of claims about the future, for example, there might be justifiable incentives for a fact-checker to give the claim a truth-rating. The criteria do not consider issues like this, which could be seen as a limitation.

In addition, the criteria do not take into account the quality of evidence, whether the fact-check is of high-quality content-wise and whether the fact-checking method and

format are as efficient as possible to combat misperceptions and fake news. However, these are important points to ponder if one wants to comprehensively appraise fact-checks. A framework that considers every relevant aspect of fact-checking (limits, journalistic aspects, quality of content, maximal efficacy, etc.) might not be easy (or even possible) to model.

The analysis (as well as comments from anonymous reviewers) also invoked some thoughts considering the future use of these criteria, as there are potential overlaps. Firstly, it is difficult in practice to make a difference between claims about values (criterion 14) and opinions (criterion 20), as expressions of values could also be interpreted as statements of opinion. This problem also applies to vague claims (criterion 11) and figures of speech (criterion 19), since figures of speech tend to be vague. In future analyses, it might be reasonable to combine criterion 14 into criterion 20, and criterion 19 into criterion 11.

Disclosure Statement

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

Funding

This work was supported by Kone Foundation and The Foundation for Municipal Development.

References

- Amazeen, Michelle. 2013. *Making a Difference?: A Critical Assessment of Fact-Checking in 2012*. Research Paper. New America Foundation.
- Amazeen, Michelle. 2015. "Revisiting the Epistemology of Fact Checking." *Critical Review* 27 (1): 1–22.
- Amazeen, Michelle. 2016. "Checking the Fact-Checkers in 2008: Predicting Political Ad Scrutiny and Assessing Consistency." *Journal of Political Marketing* 15 (4): 433–464.
- Ashoorkhani, Mahnaz, Reza Majdzadeh, Saharnaz Nedjat, and Jaleh Gholami. 2017. "Promoting the Quality of Health Research-Based News: Introduction of a Tool." *International Journal of Preventive Medicine* 8 (1): 87.
- Austin, J. L. 1962. *How to Do Things with Words*. Cambridge: Harvard University Press.
- Austin, Sophie. 2019. "Amy Klobuchar Right About State Psychiatric Bed Shortages." PolitiFact. June 13. Accessed February 28, 2020. <https://www.politifact.com/factchecks/2019/jun/13/amy-klobuchar/amy-klobuchar-right-state-psychiatric-bed-shortage/>.
- Brandtzaeg, Petter Bae, Marika Lüders, Jochen Spangenberg, Linda Rath-Wiggins, and Asbjørn Følstad. 2016. "Emerging Journalistic Verification Practices Concerning Social Media." *Journalism Practice* 10 (3): 323–342.
- Buccioli, Alessandro. 2018. "False Claims in Politics: Evidence from the US." *Research in Economics* 72 (2): 196–210.
- Charnley, Mitchell. 1936. "Preliminary Notes on a Study of Newspaper Accuracy." *Journalism Quarterly* 13 (4): 394–401.
- Drobnic Holan, Angie. 2009. "Jindal Says Unemployment Measures in Stimulus Will Cost the State Money." PolitiFact. February 24. Accessed July 22, 2020. <https://www.politifact.com/factchecks/2009/feb/24/bobby-jindal/jindal-unemployment-measure-stimulus/>.
- Drobnic Holan, Angie. 2018a. "The Principles of the Truth-O-Meter: PolitiFact's Methodology for Independent Fact-Checking." PolitiFact. February 12. Accessed December 12, 2020. <https://www.politifact.com/truth-o-meter/article/2018/feb/12/principles-truth-o-meter-politifact-methodology-i/>.

- Drobnic Holan, Angie. 2018b. "Why PolitiFact Doesn't Use the Word 'Lie' — Except Once a Year." PolitiFact. December 13. Accessed February 28, 2020. <https://www.politifact.com/truth-o-meter/article/2018/dec/13/why-politifact-doesnt-use-word-lie-except-once-yea/>.
- Ecker, Ullrich, Stephan Lewandowsky, Ee Pin Chang, and Rekha Pillai. 2014. "The Effects of Subtle Misinformation in News Headlines." *Journal of Experimental Psychology: Applied* 20 (4): 323–335.
- Ecker, Ullrich, Ziggy O'Reilly, Jesse Reid, and Ee Pin Chang. 2019. "The Effectiveness of Short-Format Refutational Fact-Checks." *British Journal of Psychology*. doi:10.1111/bjop.12383.
- Egré, Paul, and Benjamin Icard. 2018. "Lying and Vagueness." In *The Oxford Handbook of Lying*, edited by Jörg Meibauer, 354–369. Oxford: Oxford University Press.
- Ellis, Gavin. 2012. "Journalism's Road Codes: The Enduring Nature of Common Ethical Standards." *Pacific Journalism Review* 18 (2): 118–130.
- Farley, Robert, and Angie Drobnic Holan. 2010. "Barack Obama Tells Jon Stewart the Health Care Bill Cuts the Deficit by Over \$1 Trillion Dollars." PolitiFact. October 28. Accessed July 22, 2020. <http://www.politifact.com/truth-o-meter/statements/2010/oct/28/barack-obama/barack-obama-health-care-bill-cuts-deficit/>.
- Godler, Yigal, and Zvi Reich. 2017. "Journalistic Evidence: Cross-Verification as a Constituent of Mediated Knowledge." *Journalism: Theory, Practice & Criticism* 18 (5): 558–574.
- Goldman, Alvin. 1999. *Knowledge in a Social World*. New York: Oxford University Press Inc.
- Graves, Lucas. 2016. *Deciding What's True: The Rise of Political Fact-Checking in American Journalism*. New York: Columbia University Press.
- Graves, Lucas. 2017. "Anatomy of a Fact Check: Objective Practice and the Contested Epistemology of Fact Checking." *Communication, Culture & Critique* 10 (3): 518–537.
- Graves, Lucas. 2018. "Boundaries Not Drawn: Mapping the Institutional Roots of the Global Fact Checking Movement." *Journalism Studies* 19 (5): 613–631.
- Greenberg, Jon. 2017. "How Unpopular is the Obamacare Individual Mandate?" PolitiFact. November 14. Accessed July 23, 2020. <https://www.politifact.com/factchecks/2017/nov/14/donald-trump/how-unpopular-obamacare-individual-mandate/>.
- Greenberg, Jon. 2018. "Donald Trump Wrong About Crime in Germany." PolitiFact. June 18. Accessed July 23, 2020. <https://www.politifact.com/factchecks/2018/jun/18/donald-trump/donald-trump-wrong-about-crime-germany/>.
- Hayes, Andrew, and Klaus Krippendorff. 2007. "Answering the Call for a Standard Reliability Measure for Coding Data." *Communication Methods and Measures* 1 (1): 77–89.
- IFCN. 2020a. "The Commitments of the Code of Principles." Poynter. Accessed December 9, 2020. <https://ifcncodeofprinciples.poynter.org/know-more/the-commitments-of-the-code-of-principles>.
- IFCN. 2020b. "Verified Signatories of the IFCN Code of Principles." Poynter. Accessed December 9, 2020. <https://ifcncodeofprinciples.poynter.org/signatories>.
- Jacobson, Louis. 2010. "Do Presidents Always Get 'Shellacked' in Midterm Elections?" PolitiFact. September 7. Accessed July 23, 2020. <https://www.politifact.com/factchecks/2010/sep/07/mary-jordan/do-presidents-always-get-shellacked-midterm-electi/>.
- Jacobson, Louis. 2012. "Mitt Romney says Pakistan is on a Path to Overtake the U.K. in Nuclear Weapons." PolitiFact. October 23. Accessed February 28, 2020. <https://www.politifact.com/factchecks/2012/oct/23/mitt-romney/mitt-romney-says-pakistan-path-overtake-uk-nuclear/>.
- Jacobson, Louis. 2014. "Nancy Pelosi says U.S. House 'Has Never Sued a Sitting President in All of U.S. History.'" PolitiFact. July 31. Accessed February 28, 2020. <https://www.politifact.com/truth-o-meter/statements/2014/jul/31/nancy-pelosi/nancy-pelosi-says-us-house-has-never-sued-sitting-/>.
- Jacobson, Louis. 2016. "Donald Trump Largely Accurate that U.S. had Biggest Increase in Murders in 45 Years." PolitiFact. October 10. Accessed February 29, 2020. <https://www.politifact.com/truth-o-meter/statements/2016/oct/10/donald-trump/trump-claims-us-has-biggest-increase-murders-45-ye/>.
- Jacobson, Louis, and John Kruzel. 2018. "Has Donald Trump 'Been Much Tougher on Russia' than Barack Obama?" PolitiFact. February 20. Accessed February 2, 2020. <https://www.politifact.com/factchecks/2018/feb/20/donald-trump/has-donald-trump-been-much-tougher-russia-barack-o/>.

- Kruzel, John. 2018. "In Montana Senate Race, Long-Standing Battle Between Developers, Farmers Looms Large." *PolitiFact*. August 9. Accessed July 22, 2020. <https://www.politifact.com/factchecks/2018/aug/09/jon-tester/montana-senate-race-long-standing-battle-between-d/>.
- Lim, Chloe. 2018. "Checking How Fact-Checkers Check." *Research and Politics* 5 (3): 1–7.
- Lowrey, Wilson. 2017. "The Emergence and Development of News Fact Checking Sites: Institutional Logics and Population Ecology." *Journalism Studies* 18 (3): 376–394.
- Maier, Scott. 2005. "Accuracy Matters: A Cross-Market Assessment of Newspaper Error and Credibility." *Journalism & Mass Communication Quarterly* 82 (3): 533–551.
- Makkonen, Kimmo, and Matti Wiberg. 2017. "Faktantarkistuksen perusteet." In *Faktat ja politiikka*, edited by Matti Wiberg, 37–53. Helsinki: Kunnallisan kehittämissäätiö, publication 7. (Available only in Finnish).
- Mantzaris, Alexios. 2015. "Introducing Poynter's International Fact-checking Network." Poynter. October 21. Accessed September 30, 2020. <http://www.poynter.org/2015/fact-checkers-of-the-world-unite/379716/>.
- Mantzaris, Alexios. 2016. "Fact-Checkers Around the World Agree on Shared Code of Principles." Poynter. September 15. Accessed September 30, 2020. <https://www.poynter.org/fact-checking/2016/fact-checkers-around-the-world-agree-on-shared-code-of-principles/>.
- Marietta, Morgan, David Barker, and Todd Bowser. 2015. "Fact-Checking Polarized Politics: Does the Fact-Check Industry Provide Consistent Guidance on Disputed Realities?" *The Forum* 13 (4): 577–596.
- Merpert, Ariel, Melina Furman, Maria Victoria Anauati, Laura Zommer, and Inés Taylor. 2018. "Is That Even Checkable? An Experimental Study in Identifying Checkable Statements in Political Discourse." *Communication Research Reports* 35 (1): 48–57.
- Nieminen, Sakari, Miira Raiskila, and Matti Wiberg. 2017. *Faktat politiikassa*. Helsinki: Kunnallisan kehittämissäätiö. (Available only in Finnish).
- Nieminen, Sakari, and Lauri Rapeli. 2019. "Fighting Misperceptions and Doubting Journalists' Objectivity: A Review of Fact-Checking Literature." *Political Studies Review* 17 (3): 296–309.
- Nieminen, Sakari, and Matti Wiberg. 2018a. "Faktantarkistajat harhateillä." *Yhteiskuntapolitiikka* 83 (5-6): 599–606. (Available only in Finnish).
- Nieminen, Sakari, and Matti Wiberg. 2018b. "Noudattaako Faktabaari faktantarkistuskriteereitä?" *Kriittinen Arvio. Media & Viestintä* 41 (3): 214–227. (Available only in Finnish).
- Nyhan, Brendan, and Jason Reifler. 2015. "The Effect of Fact Checking on Elites: A Field Experiment on U.S. State Legislators." *American Journal of Political Science* 59 (3): 628–640.
- Porlezza, Colin, Scott Maier, and Stephan Russ-Mohl. 2012. "News Accuracy in Switzerland and Italy: A Transatlantic Comparison with the US Press." *Journalism Practice* 6 (4): 530–546.
- Rauchfleisch, Adrian, Xenia Artho, Julia Metag, Senja Post, and Mike Schäfer. 2017, July. "How Journalists Verify User-Generated Content During Terrorist Crises. Analyzing Twitter Communication During the Brussels Attacks." *Social Media + Society* 3 (3). doi:10.1177/2056305117717888
- Richert, Catharine. 2009. "McConnell Gets it Right on Medical Malpractice Reform." *PolitiFact*. December 2. Accessed July 22, 2020. <https://www.politifact.com/factchecks/2009/dec/02/mitch-mcconnell/mcconnell-gets-it-right-medical-malpractice-reform/>.
- Rögener, Wiebke, and Holger Wormer. 2017. "Defining Criteria for Good Environmental Journalism and Testing Their Applicability: An Environmental News Review as a First Step to More Evidence Based Environmental Science Reporting." *Public Understanding of Science* 26 (4): 418–433.
- Sankari, Valtteri, and Matti Wiberg. 2019. "Tarkistuksessa vielä parannettavaa: Ylen faktantarkistuksen tarkistus." *Media & Viestintä* 42 (4): 235–245. (Available only in Finnish).
- Shapiro, Ivor, Patrizia Albanese, and Leigh Doyle. 2006. "What Makes Journalism "Excellent"? Criteria Identified by Judges in Two Leading Awards Programs." *Canadian Journal of Communication* 31 (2): 425–445.
- Stencel, Mark. 2017. "International Fact-Checking Gains Ground, Duke Census Finds." *Duke Reporter's Lab*. February 28. Accessed September 30, 2020. <https://reporterslab.org/international-fact-checking-gains-ground/>.
- Stencel, Mark, and Joel Luther. 2019. "U.S. Fact-Checkers Gear up for 2020 Campaign." *Duke Reporter's Lab*. November 25. Accessed March 2, 2020. <https://reporterslab.org/u-s-fact-checkers-gear-up-for-2020-campaign/>.

- Tobias, Manuela. 2017. "Trump's Mostly False Claim that NFL Ratings are 'Way Down'." PolitiFact. September 24. Accessed July 23, 2020. <https://www.politifact.com/factchecks/2017/sep/24/donald-trump/trumps-mostly-false-claim-nfl-ratings-are-way-down/>.
- Uscinski, Joseph. 2015. "The Epistemology of Fact Checking (Is Still Naïve): Rejoinder to Amazeen." *Critical Review* 27 (2): 243–252.
- Uscinski, Joseph, and Ryden Butler. 2013. "The Epistemology of Fact Checking." *Critical Review* 25 (2): 162–180.
- Valverde, Miriam. 2018. "Donald Trump Falsely Says ICE Liberated Towns from MS-13 Control." PolitiFact. July 9. Accessed February 28, 2020. <https://www.politifact.com/factchecks/2018/jul/09/donald-trump/did-ice-liberate-towns-donald-trump-claimed-have-w/>.
- Young, Dannagal, Kathleen Hall Jamieson, Shannon Poulsen, and Abigail Goldring. 2018. "Fact Checking Effectiveness as a Function of Format and Tone: Evaluating FactCheck.org and FlackCheck.org." *Journalism & Mass Communication Quarterly* 95 (1): 49–75.