Short but critical? How “fake news” and “anti-elitist” media attacks undermine perceived message credibility on social media

Bos, L.; Egelhofer, J.L.; Lecheler, S.

DOI
10.1177/00936502231178432

Publication date
2023

Document Version
Final published version

Published in
Communication Research

License
CC BY

Citation for published version (APA):
Short but Critical? How “Fake News” and “Anti-Elitist” Media Attacks Undermine Perceived Message Credibility on Social Media

Linda Bos¹, Jana Laura Egelhofer², and Sophie Lecheler³,¹

Abstract
Citizens increasingly turn to social media for information, where they often rely on cues to judge the credibility of news messages. In these environments, populist politicians use “fake news” and “anti-elitist” attacks to undermine the credibility of news messages. This article argues that to truly understand the impact of these criticism cues, one must simultaneously consider additional contextual cues as well as individual-level moderators. In a factorial survey, we exposed 715 respondents to tweets by a politician retweeting and discrediting a news message of which topic and source varied. We find that both the fake news cue and the anti-elitist cue have limited across-the-board effects but decrease credibility if the message is incongruent with voters’ issue positions. Our results thus offer a more optimistic view on the power of populist media criticism cues and suggest that source and confirmation heuristics are (still) stronger influences on citizens’ credibility evaluations.

Keywords
social media, news and newsworthiness, online credibility, survey, experiment

¹University of Amsterdam, The Netherlands
²LMU Munich, Germany
³University of Vienna, Austria

Corresponding Author:
Linda Bos, Amsterdam School of Communication Research, University of Amsterdam, Nieuwe Achtergracht 166, Amsterdam 1018 WV, The Netherlands.
Email: l.bos@uva.nl
The role of news media as information providers for political decision-making in democratic societies is uncontested (Tsfati & Cohen, 2005; Van Aelst et al., 2017). However, news media can only fulfill this role, if citizens perceive the media information they consume as credible (Strömbäck et al., 2020). Yet, while “in the old days” news consumers would derive their information directly from the news medium and the public value of news was embedded in the relation between the journalistic news content and the news user, information production and consumption in the high-choice media environment (Van Aelst et al., 2017) follows a different pattern. Social media play an important role here: citizens follow news media on these platforms directly (Newman, 2020) but are also increasingly exposed to online news incidentally (Boczkowski et al., 2018).

In this digital media environment, the perceived credibility of news in general, and of a specific news message in particular, is affected by a variety of factors. While the perceived credibility of news content traditionally is determined by source and message factors, such as the original news outlet, in digitized and information-rich democracies, various actors play a role in shaping the meaning of political information. Especially on social network sites, everyone is able to reframe the content produced by professional journalists (Casero-Ripollés, 2018). Opinion leaders can positively affect the perception of news stories by recommending stories (Turcotte et al., 2015), but they can also reconstruct the news story in a negative way by adding a critical assessment. News criticism on social media mostly occurs as short attacks that can easily activate long-running media criticism narratives (Carlson, 2016).

Evidence suggests that populist politicians, especially, use social media to actively undermine the credibility of (certain) media as a strategy to reject critical reporting (e.g., Engesser et al., 2017; Hameleers, 2022). Most prominently, populists have employed two types of short attack cues in recent years, namely “anti-elitist” and “fake news” media attack cues. Criticizing news media as “anti-elitist” has a long tradition in populist communication, but with the recent surge of populism in many Western-European countries, these attacks have been increasing (Aalberg et al., 2017; Fawzi, 2020; Krämer, 2018). Anti-elitist cues refer to the populist viewpoint that news media are part of the “hypocritical elite” conspiring with the political establishment. In this view, journalists are “out of touch” with the public’s interest and biased in favor of the ruling elite (Fawzi, 2020; Krämer, 2018). More recently, a growing number of populist politicians accuse the media of spreading disinformation and label them as “fake news” (Hameleers, 2022).

As citizens rely on political elite cues when evaluating media content (Smith, 2010), exposure to critical attacks on social media likely has negative effects on credibility perceptions. Yet, while scholars acknowledge that these attacks have become a central part of political communication with potentially negative consequences, empirical studies on this matter are largely lacking (Carlson, 2016; Fawzi, 2020), offer inconclusive evidence, or are heavily focused on the U.S. (Guess et al., 2017; Tamul et al., 2020).

In this paper, we argue that to make well-founded statements about the impact of these types of media criticism, other factors likely influencing news message credibility evaluations on social media platforms, such as the source, the issue, and the extent
to which these factors interact with predispositions of citizens should be taken into account. In order to investigate the extent to which media criticism, as well as these other factors, impact news message credibility, we mimic the digital media environment in our research design by conducting a factorial survey (FS). This design allows us to manipulate a variety of factors simultaneously. It thus is arguably a better-suited approach to study message credibility judgments on social media compared to traditional survey experiments (e.g., Wallander, 2009). We focus on the Netherlands, where both left- and right-wing populist parties were successful in the last two decades. Especially the latter have been attacking the media, using both anti-elitist and fake news cues, on a regular basis (e.g., Hameleers, 2022). By studying two prominent cues of modern populist media criticism in combination with traditional factors influencing credibility judgments, this study offers an important contribution to the literature on how message credibility perceptions are developed on social media.

**Message Credibility Perceptions on Social Media**

One of the central democratic roles of news media is to provide citizens with the information they need to be “free and self-governing” (Strömbäck et al., 2020, p.148), which renders the question of how citizens perceive the credibility of information coming from news media a key focus in political communication and journalism research. While for a long-time citizens relied on a small selection of media outlets for information, today the number of people who turn to social media to get informed is growing (Newman, 2020). News use on social media has certain particularities, such as incidentalness (Kümpel, 2022). That is, many people do not actively visit websites of news providers but are exposed to news in their social media streams (i.e., “aggregated flows of content” Bayer et al., 2020, p. 481). There, other—sometimes even unknown—actors, such as politicians, share news messages accompanied by social information, such as likes and (critical) comments (Kümpel, 2022; Metzger et al., 2010). Furthermore, news exposure is increasingly “granualized” (Kümpel, 2022), meaning that instead of consuming the total content of one newspaper or even a complete article, on social media, citizens increasingly consume isolated news headlines (so-called “snack news”; Schäfer, 2020). Indeed, research suggests that the average social media user clicks on only 7% of news previews in their feed to read the whole story (Bakshy et al., 2015).

When evaluating the credibility of these news message previews on social media, citizens heavily rely on various cues (Sterrett et al., 2019). In this study, we consider the following factors that likely influence citizens’ message credibility evaluations in high-choice media environments: (a) criticism by the person sharing the news story, (b) contextual factors related to the news story, and (c) individual predispositions relevant for either (a) or (b).

**Critical Cues Accompanying News Stories**

As mentioned above, news on social media is often accompanied by social information (Kümpel, 2022), such as critical comments by the person sharing the news
(Anspach & Carlson, 2020; Carlson, 2016). Put differently, on social media, consuming news messages often means simultaneously consuming criticism of these messages (Carlson, 2016). Crucially, research shows that this criticism often receives more attention than the information provided in the shared news message (Anspach & Carlson, 2020).

Especially populist politicians frequently attack mass media on platforms (e.g., Carlson et al., 2021). While these criticisms are often only short utterances, they can easily activate comprehensive media criticism narratives (Carlson, 2016). This might be especially true for recent and well-known phrases, such as “fake news” or “hypocritical elite,” which have been frequently used in political discourse and thus been connected to meaning (Tamul et al., 2020; Van Duyn & Collier, 2019). In other words, these verbal cues might trigger social heuristics connected to the assessment of journalistic credibility. Heuristics are mental shortcuts or rules of thumb that people use to reduce the complexity of cognitive tasks (e.g., Bellur & Sundar, 2014). Determining the credibility of a news story is a cognitively demanding task as it requires knowledge related to the covered issues in order to estimate whether a given story is providing a fair and accurate representation of the real world (Smith, 2010). Therefore, abundant research shows that news users constantly rely on heuristics, which can be triggered by cues related to the specific story (Bellur & Sundar, 2014; Metzger et al., 2010) but also by accompanying cues of political elites (Smith, 2010; Zaller, 1992). This study investigates the effects of two distinct elite cues on message credibility: a “fake news” cue and an “anti-elitist” cue.

Criticizing the media as part of the “hypocritical elite” has a long tradition in political populism (Jagers & Walgrave, 2007). At the core of populism lies a “Manichean outlook” on society, which implies a binary division between two societal groups, that is, the “good” or “ordinary” people on the one hand and the “evil” or “corrupt” elites on the other hand (e.g., Mudde, 2004). The elite threatens and is blamed for all the problems of the people (Bos et al., 2020) and encompasses not only the political establishment but also the media. This “anti-media populism” describes the populist viewpoint that the media do not constitute a central democratic institute but are corrupt and malicious actors who are “part of an elite conspiracy” (Krämer, 2018, p. 453). In line with this, content analyses show that anti-elitist media attacks are a regular feature of populist communication (Aalberg et al., 2017; Engesser et al., 2017). Specifically, populists depict the media as being “out of touch” with the people, neglecting their interests, and instead serving the ruling elite (Fawzi, 2020). As many journalists are highly educated, they might be considered part of the “intellectual elite,” by which populists do not feel represented (Fawzi, 2020). Furthermore, populist criticism often centers around the topic that coverage is biased toward the liberal (Fawzi, 2020). This relates to a prominent and long-standing political discourse about (primarily left-wing) media bias (Ladd, 2012; Smith, 2010). The idea that news media coverage often is one-sided is thus arguably very accessible in people’s minds. This is underlined by research showing that verbal cues such as “bias” (Smith, 2010) or “media conspiracy” (Uscinski et al., 2016) trigger related beliefs about the news. It is, therefore, to be
expected that a well-known anti-elitist cue (“hypocritical elite”) can act as a heuristic, triggering beliefs about media partiality.

More recently, populist politicians have adopted another type of media attack, that is, accusing the media of spreading disinformation and “fake news” (e.g., Hameleers, 2022). The described “Manichean outlook” of populism encompasses a polarized conceptualization of “truth.” That is, populism “rejects the possibility of truth as a common normative horizon and collective endeavour in democratic life” (Waisbord, 2018, p. 25). Instead, “truth” is produced by the people, while the (media) elite is spreading lies (Waisbord, 2018, p. 25). Accusing news media of spreading “fake news” therefore aligns with populist ideology (Hameleers, 2022). Unsurprisingly, “fake news” attacks, coined by Donald Trump, have been adopted by many populist politicians around the globe (Hameleers, 2022).

Using the phrase “fake news” might be a particularly powerful lying accusation as it has risen to a very prominent buzzword used frequently in political discourse, journalistic debate, and social media conversations to signal that information is inaccurate (Brummette et al., 2018; Egelhofer et al., 2020). It is connected to increased worries of citizens about the threats of disinformation in online news (e.g., Newman, 2019). Therefore, we argue that “fake news” is a powerful cue that clearly signals that the veracity of some piece of information is questionable.

However, so far, there are only a few empirical studies on the effects of news criticism cues, and these predominantly focus on the U.S. and Donald Trump. On the one hand, results suggest that fake news cues negatively affect (at least some) citizens’ general media trust levels (Guess et al., 2017; Van Duyn & Collier, 2019). On the other hand, exposure to tweets by Donald Trump in which he accuses the media of being fake news has no effect on the perceived credibility of an unrelated, subsequently read news story (Tamul et al., 2020). However, as noted, on social media, citizens are increasingly incidentally exposed to isolated news headlines directly accompanied by a fake news cue or an anti-elitist cue, not necessarily in subsequent fashion. That is, the criticism is directed at the specific message at hand—not at the media (or individual outlets) in general. In addition, while citizens might be incidentally exposed to short news snippets, in most cases, they do not take the time to go to the website of a media outlet and read a news story at length (Bakshy et al., 2015), an activity that in itself might have profound effects on the perceived message credibility. If we aim to understand the impact of a media attack cue on message credibility, we thus require a design in which these cues question the credibility of the specific message (or a mere preview) in a setup that mimics real life on a social media platform.

In sum, we argue that an anti-elitist cue signals that a news message is not impartial, while the fake news cue signals that a news message is not accurate. Impartiality and accuracy are core criteria of credibility (e.g., Naab et al., 2020). Therefore, when attached to a news message, these cues likely have negative effects on citizens’ credibility perceptions. Hence, we expect that both the fake news cue (H1) and the anti-elitist cue (H2) have a negative effect on news message credibility.
Contextual Cues of the News Story

As we laid out, the fake news cue and the anti-elitist cue do not appear in isolation on social media. Rather, they occur in combination with a news message consisting of several elements that citizens rely on when evaluating news credibility.

**Source.** One of the most frequently used cues for the evaluation of information is its source (Metzger & Flanagin, 2013; Metzger et al., 2010). The “reputation heuristic” suggests that the assessment of credibility of news content depends on how familiar and reputable the news outlet is (Metzger & Flanagin, 2013; Metzger et al., 2010). Some news organizations have acquired a higher reputation and are valued for their strong journalistic standards and gatekeeping criteria. Their source cue might activate a sense of trustworthiness, professionalism, and expertise (Sundar et al., 2007), which subsequently increases perceived message credibility (Funt et al., 2016).

**Topic.** Another important factor in this context is the topic of a news message (Fogg, 2003). Here we consider the contextual effects of the topic. The information environment in which a specific cue is experienced, for example, the media landscape and national issue agenda, is likely to matter (Strömbäck, 2017). Topics that are salient or contentious on national news agendas may be evaluated differently across the board (Feinholdt et al., 2017).

It is important to take these more traditional cues into account when considering the influence of fake news and anti-elitist cues on news credibility evaluations. In other words, discrediting media cues might have differential impacts depending on the source or the topic:

*To what extent do the effects of the fake news cue (RQ1a) and the anti-elitist cue (RQ1b) on news credibility vary by the issue of the news message?*

*To what extent do the effects of the fake news cue (RQ2a) and the anti-elitist cue (RQ2b) on news credibility vary by the media outlet (as the source of the news message)?*

**Individual Predispositions**

Finally, we look at individual predispositions that might moderate the impact of message previews and accompanying criticisms on perceived message credibility.

**Confirmation Biases**

First, the effectiveness of elite criticism cues is likely dependent upon predispositions (Smith, 2010; Uscinski et al., 2016; Zaller, 1992), such as individuals’ pre-existing attitudes toward the topic of the specific news message under attack. The concept of confirmation bias proposes that attitude-congruent information is processed less critically compared to incongruent information (Klayman & Ha, 1987). In line with this, research suggests that citizens evaluate information that supports their preexisting beliefs as more credible than information that counters these opinions (Metzger et al.,
This tendency to view incongruent information as not credible has also been termed the “self-confirmation heuristic” (Metzger et al., 2010). Based on this, one can expect that citizens’ attitudes toward the topic of the news message influence their credibility perceptions in such a way that congruent messages are perceived as more credible, while incongruent messages are perceived as less credible. However, and most importantly, in the latter situation, critical cues might act as confirmation bias heuristics, confirming the expectation that the incongruent message is less credible. Therefore, we predict that the effect of the fake news cue (H3a) and the anti-elitist cue (H3b) on news message credibility are particularly strong when the news story’s issue stance is incongruent with citizens’ pre-existing attitudes.

**Populist Attitudes**

Another predisposition that likely determines the effectiveness of elite criticism cues is populist attitudes—attitudes measuring the extent to which citizens agree with the central tenets of populist ideology (Akkerman et al., 2014). We know that populist voters’ views of mainstream media are quite negative: voters with stronger populist attitudes are more likely to have hostile media perceptions, lower media trust, are less satisfied with the media’s performance, and perceive media quality to be lower than voters not holding these attitudes (Fawzi, 2019; Schulz et al., 2020). Furthermore, a populist worldview is related to the perception that news media spread mis- and disinformation (Hameleers, 2022). In other words, populist voters’ negative attitudes toward the media are in line with populist politicians’ anti-media populism and polarized construction of truth. We, therefore, hypothesize that the effect of the anti-elitist cue (H4) and fake news cue (H5) are particularly strong for citizens who hold populist attitudes.

**Methods**

**Experimental Design and Stimuli**

In our study, we focus on Twitter, a social medium frequently used by politicians, which arguably features more (hard) news content than other social media platforms, such as Instagram or Facebook (Kümpel, 2022). Accordingly, in a Twitter context, chances are high that individuals are exposed to news messages accompanied by media criticism.

In order to test our hypotheses and obtain insights into which principles guide respondents’ evaluations of a news story while taking into account subgroup differences, we make use of a FS design (Ausburg & Hinz, 2015). FS are increasingly used within communication science (Glogger & Otto, 2019; Helfer & Aelst, 2016; Kruikemeier & Lecheler, 2018), most notably in journalism studies focusing on factors explaining news selection by journalists (see Otto & Glogger, 2020 for an overview), or politicians (Helfer, 2016). Yet, to our understanding, its usage for studying evaluations by news consumers is limited (albeit see Kruikemeier & Lecheler, 2018). In a FS design, respondents are exposed to a number of vignettes in order to
disentangle the causal effects of relevant factors underlying judgments or evaluations of certain objects. This design allows us to simultaneously test the impact of the anti-elitist cue and the fake news cue, the source of the article—a traditional left-wing newspaper (Volkskrant) versus a traditional right-wing newspaper (De Telegraaf) versus a relatively new and digital news outlet (VICE)—as well as the topic of the article (climate/immigration/EU/integration), and the valence of the topic (positive vs. negative) while also taking into account individual-level characteristics. As noted by Auspurg and Hinz (2015), in FS, the “hypothetical descriptions create an opportunity to overcome the problems of highly correlated dimensions in reality and allow an examination of situations that rarely or never occur in reality” (p. 13). For instance, in real life, the occurrence of a fake news cue or an anti-elitist cue is highly correlated with the presence of specific other factors, such as a combination of a left-leaning issue stance and a left-leaning news source. A factorial survey design allows us to tease out the impact of these cues on credibility evaluations regardless of the slant or issue. Moreover, multidimensional factorial survey designs require respondents to evaluate all dimensions simultaneously and consider possible trade-offs. In that sense, “the description of a situation in a FS resembles real life much more closely because it incorporates more complex considerations of trade-offs” (p. 11). In other words, while the design and the analysis aim to examine the impact of specific factors, the factorial survey itself forces respondents to evaluate vignettes holistically, akin to how they evaluate social media messages in real life.

This leads to a total of five dimensions with up to four levels each and 96 possible vignettes (see Supplemental Table A1 for an overview of the vignette dimensions and levels). In FS, the efficiency of the design is crucial: “the goal is to select [a design] that provides the maximum statistical information, given sensible numbers of respondents and vignettes per respondent” (Auspurg & Hinz, 2015, p. 27), that is, a design with maximum “D-efficiency” is preferred. We made use of the SAS macro mktex to construct a D-efficient design, ending up with a sample of 48 vignettes, blocked in 12 decks of four vignettes each (D-efficiency = 97.19). Decks of four vignettes also maximize the variation in the tweets received by each respondent without compromising the credibility of the deck by repeating too many factors.

The vignettes are constructed as tweets1: in each vignette, an anonymous politician retweets a news article (a headline and a link). The politician is anonymized, and a blurred picture is used, the gender of which is unidentifiable. Anonymizing the politician is a deliberate choice in this study. First, on social media citizens encounter politicians and opinion leaders they do not know, from parties they are unfamiliar with. The current design departs from this premise. Second, while in two-party contexts partisanship is expected to have a profound influence on the reception of a message, literature shows that in fragmented multiparty systems like the Dutch, party cues are much less likely to be effective. Not only are there more non-partisans, in multiparty systems identification as well as negative partisanship is not necessarily confined to one party (Wagner, 2021). Finally, including partisanship in the Netherlands (with 17 parties elected to parliament in March 2021) is not feasible in a design also testing the impact of the news source, issue, and valence.
The neutral tweet states “there they are again.” In the anti-elitist cue conditions, the words “the hypocritical elite” are added to the neutral tweet as well as the hashtag “they look away” (wegkijkers in Dutch - a term used by Dutch populists to attack elites neglecting the peoples’ problems). The fake news cue is operationalized by adding the words “with their fake news” (in Dutch) and the hashtag #fakenews. The news message is presented as originating from one of three Dutch news sources: de Volkskrant, De Telegraaf, or VICE.nl. Both newspapers (de Volkskrant and De Telegraaf) were selected because of their high circulation and ideological slant. De Volkskrant is the most left-leaning and De Telegraaf is the most right-leaning newspaper in the Netherlands. In addition, both newspapers have a strong reputation as professional news media. VICE.nl was selected because of its lack of reputation. This news medium is new to the Dutch media market, making it less likely to be perceived as trustworthy. The Twitter handle, and the logo of the source are included in the manipulation, as is common in retweets on Twitter. In the retweet text the lead of a news item is given. The news lead was developed using existing news stories but did not resemble any salient news topics during the time of data collection. Table 1 gives an overview of the main claims in these news previews (see Supplemental Appendix C for the exact wording of the news leads).

### Table 1. Overview Main Claims in News Message Previews Sorted by Issues and Valence.

<table>
<thead>
<tr>
<th></th>
<th>EU</th>
<th>Integration</th>
<th>Immigration</th>
<th>Climate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Netherlands should not leave EU</td>
<td>Positive impact of Islamic schools for cultural integration</td>
<td>Crime numbers among refugees exaggerated</td>
<td>Predictions of climate crisis exaggerated</td>
</tr>
<tr>
<td>Negative</td>
<td>Netherlands should leave EU (Nexit)</td>
<td>Negative impact of Islamic schools for cultural integration</td>
<td>Crime numbers among refugees underestimated</td>
<td>Predictions of climate crisis falling short</td>
</tr>
</tbody>
</table>

Sample and Procedure

The study made use of a commercial survey panel and the data was collected late January and early February 2020. The survey agency was instructed to employ quota sampling on age, education, and gender so as to achieve a representative sample of the population. The complete dataset constitutes a diverse sample (n = 743) of Dutch citizens with regard to age (M = 45.62, SD = 14.93, Min = 18, Max = 74), education (M = 4.26, SD = 1.52, Min = 1, Max = 7), political interest (M = 4.21, SD = 1.61, Min = 1, Max = 7) and ideology (M = 5.50, SD = 1.83, Min = 1, Max = 10); 49.79% of the participants were female, and 0.42% non-binary.

Before exposing respondents to the vignettes, we measured socio-demographics, populist attitudes, issue positions on each of the issues in the vignettes, media use and trustworthiness, and political predispositions. After the pre-test, participants were given instructions. They were told to read each tweet carefully and evaluate the
credibility and objectivity of each news item. Participants were then randomly exposed to a deck of four tweets (the order of which was randomized).

**Measures**

Our dependent variable, *message credibility*, was operationalized by asking respondents to what extent they think the news item retweeted by the politician is credible on a scale from 1 (not at all) to 11 (very much) ($M=5.69$, $SD=2.62$).

Issue positions were tapped before the vignette study with four items, on each issue one (see Supplemental Appendix D for the exact question wording) and measured on an 8-point scale, running from 1 (completely disagree) to 8 (completely agree). In a next step, (in)congruence with the message issue stance was computed by matching the tapped opinion on an issue with the specific news story, again resulting in a scale measured from 1 (completely disagree) to 8 (completely agree)—with the news story in this case ($M=4.48$, $SD=2.19$). For instance, if a respondent was exposed to a news headline that was positive about immigration and had a positive opinion on immigration herself (above the midpoint of the scale), she was expected to agree with the news lead. However, if she had a negative opinion, she was expected to disagree. Because the constructed scale does not follow the normal distribution, we recoded the variable into three groups: respondents disagreeing strongly (i.e., scoring $M-1$ $SD$, constituting 21.96% of the sample) respondents agreeing strongly with the story they were exposed to (i.e., scoring $M+1$ $SD-21.36$% of the sample), and respondents not having a strong opinion, scoring around the midpoint of the scale.

To measure populist attitudes, we focused on anti-elitist and popular sovereignty attitudes (for the wording of the items, see Supplemental Appendix D)—the two main dimensions of populism. Since these sub-dimensions in populism are considered to be non-compensatory, meaning that scoring high on one sub-dimension should not be compensated by scoring low on the other sub-dimension, we follow the approach as suggested by Wuttke et al. (2020). They propose a “Sartorian” approach by categorizing respondents only as populist when they belong to the top 25% on both dimensions. This resulted in a variable “populist” by which 22% of the respondents were considered populist.

**Data Cleaning**

We collected data from a total of 849 respondents. Respondents who did not give their informed consent ($n=49$) were removed from the dataset, as well as 50 respondents who did not complete the questionnaire, and 3 respondents younger than 18.

**Analysis Strategy**

To simultaneously analyze the effects of vignette variables and respondents’ variables the data were reordered in long format. To consider the dependency of observations we run linear mixed models (*xtmixed*) in Stata SE 16.1. To test H1 and H2 we analyzed the data hierarchically by entering the different factors in the model, starting with both
cues, entering the source of the message in a next step, then the issue, and then the valence. In all models, we add several controls: we control for background characteristics—age, gender, and education, the order of the stimuli, and the deck respondents were exposed to. Finally, we control for speeders. We added random slopes for the main predictors, that is, the experimental factors, one by one in our models. LR tests were used to test whether the addition of each random slope to the model led to a significant increase in model fit. The reporting of the models follows the suggestions of Meteyard and Davies (2020). The explained variance (R-squared) of level 1 and level 2 was calculated using the mlt package in Stata 16.1, using the Snijders/Bosker method (resulting in a more conservative estimation).

To test the moderations (H3a-H5 and RQ1a-RQ2b), we added interaction terms to our models and ran marginal effects analyses. In their standard work on Factorial Surveys Auspurg and Hinz (2015, p. 97) note that the insignificance of an interaction term can be attributed to insufficient statistical power. In such a case, it is advised to run subgroup analyses, which is what we resort to testing H3a and H3b.5

**Results**

Table 2 presents the results of hierarchical MLM random effects models. In the first model both cues, separately, not combined, decrease news message credibility

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fake news cue</td>
<td>−0.316 (0.135)*</td>
<td>−0.314 (0.135)*</td>
<td>−0.325 (0.134)*</td>
<td>0.072 (0.139)</td>
</tr>
<tr>
<td>Populist cue</td>
<td>−0.400 (0.146)**</td>
<td>−0.399 (0.146)**</td>
<td>−0.406 (0.145)**</td>
<td>0.005 (0.149)</td>
</tr>
<tr>
<td>VICE</td>
<td>0.139 (0.110)</td>
<td>0.134 (0.109)</td>
<td>0.130 (0.109)</td>
<td>0.108 (0.107)</td>
</tr>
<tr>
<td>Telegraaf</td>
<td>−0.304 (0.103)**</td>
<td>−0.289 (0.103)**</td>
<td>−0.247 (0.101)*</td>
<td></td>
</tr>
<tr>
<td>Integration</td>
<td>−0.303 (0.098)**</td>
<td>−0.309 (0.098)**</td>
<td>−0.307 (0.096)**</td>
<td></td>
</tr>
<tr>
<td>EU</td>
<td>−0.253 (0.109)*</td>
<td>−0.254 (0.107)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climate</td>
<td>−0.371 (0.111)**</td>
<td>−0.389 (0.109)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Valence</td>
<td>−0.038 (0.110)</td>
<td>−0.044 (0.108)</td>
<td>−0.772 (0.085)**</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>6.541 (0.457)*****</td>
<td>6.745 (0.461)*****</td>
<td>6.964 (0.467)*****</td>
<td>7.788 (0.473)*****</td>
</tr>
</tbody>
</table>

**Random effects**

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Respondent (intercept)</td>
<td>1.437 (0.060)</td>
<td>1.440 (0.060)</td>
<td>1.407 (0.076)</td>
<td>1.415 (0.098)</td>
</tr>
<tr>
<td>Populist cue (slope)</td>
<td>1.339 (0.158)</td>
<td>1.172 (0.158)</td>
<td>1.133 (0.160)</td>
<td>1.030 (0.163)</td>
</tr>
<tr>
<td>Issue</td>
<td>0.133 (0.088)</td>
<td>0.133 (0.088)</td>
<td>0.111 (0.101)</td>
<td></td>
</tr>
<tr>
<td>Valence</td>
<td></td>
<td></td>
<td></td>
<td>0.148 (0.292)</td>
</tr>
</tbody>
</table>

perceptions, as expected in respectively H1 and H2. LR tests revealed that adding a random slope for the populist cue significantly improved model fit, meaning that the impact of an anti-elitist cue on news credibility perceptions differs significantly between respondents. In Model 2, we see that news stories from VICE and Telegraaf are perceived as less credible than news stories from de Volkskrant. In this (and subsequent) model(s) the addition of random slopes for source did not increase model fit. Model 3 shows that stories discussing cultural integration, or the EU are perceived as less credible than immigration stories. In addition, slopes of the issue factor vary significantly between respondents. When the valence of the issue is added to the model—in Model 4—we find that the negative effects of the fake news and anti-elitist cue disappear. Instead, we find that positively valenced stories (positive about immigration, integration, EU, and positive about the climate—i.e., skeptical about climate change) are perceived as less credible (albeit, again, this impact differs between respondents—as indicated by the random slope). In other words: we find limited support for H1 and H2: both cues decrease perceptions of message credibility as long as the specific slant of the news story is not taken into account. This implies that the valence of the issue overrides the impact of both discrediting cues. In assessing the credibility of news messages, source and content characteristics are more important than discrediting media cues.

In a next step, we investigate whether the effect of the fake news cue and the anti-elitist cue is dependent upon issue agreement (H3a and H3b). We interacted the cues with the three-level agreement scale and ran a marginal effects analysis. Figure 1 shows the results of these analyses. The findings show that the effectiveness of the fake news cue is not dependent upon agreement with the message. We do find a backlash effect in a marginally significant positive impact of the combined cues on message credibility, but only for respondents agreeing strongly with the message. These results show no support for our expectation that the impact of the fake news cue (H3a) and the anti-elite cue (H3b) is dependent upon agreement with the message.

Following up on the suggestion by Auspurg and Hinz (2015), in Table 3, we conduct subgroup analyses. Here the coefficients show that respondents disagreeing with a message perceive the message as less credible when it is paired with a fake news cue or an anti-elitist cue. What is also interesting in these findings is that for disagreeing respondents, other message elements, such as the source or the topic, do not decrease the credibility of the message. Only discrediting cues do. For agreeing respondents, the credibility of a news message is affected by the topic and the valence of the story. Because the subgroup analyses and the marginal effects analyses point to a weak cross-level interaction, we tentatively accept H3a and H3b.

Next, we investigate whether effects differ between issues (RQ1a and RQ1b). Again, we estimate interaction effects. The results of the marginal effects analyses (Figure 2) show that only the impact of the fake news cue is dependent upon the issue. It negatively affects the credibility of the integration messages, while it has a positive impact on the credibility of an EU message and on a climate change message.6 In addition, when the fake news cue is combined with the populist cue, it increases the credibility of the integration message. Again, we perceive backlash effects, even more than effective news criticism effects.
Differences between sources (RQ2a and RQ2b) are estimated similarly. Figure 3 shows the result of nine marginal effects analyses, testing the marginal impact for each cue as well as the combined cues in the absence and presence of a source effect. The results show that while the impact of a cue differs in most cases, the difference is insignificant in almost all cases. The only significant impact we find is a backlash effect: the combined anti-elitist and fake news cue has a positive impact on message credibility of a Telegraaf post.

H4 posited that the effect of the anti-elitist cue was contingent on whether the respondent was populist or not, and H5 posed the same hypothesis for the fake news cue. With 22% of the respondents categorized as populist, we ran a marginal effects analysis. Figure 4 shows that there is a marginally significant negative impact of the anti-elitist cue on news credibility for populist voters, while this effect is non-existent for non-populist voters. H4 is cautiously accepted. We see that the impact of the fake news cue on credibility perceptions does not differ between populist and non-populist voters, leading us to reject H5.

We investigated these effects further by analyzing how populist attitudes moderated the discrediting cues in the most likely scenario, which involved discrediting the newspaper with the reputation for higher quality reporting—de Volkskrant (Roggeband & Vliegenthart, 2007). For non-populist respondents, it can be predicted that discrediting
media cues are less likely to provoke a reaction than for populist respondents, who are less likely to be impressed by reputation and more affected by these cues. We find that the fake news cue indeed does not affect non-populist respondents’ assessments of the reliability of de Volkskrant’s news reporting. However, this effect is negative—albeit non-significant—for populists (see Supplemental Table A5 in the Appendix). Additionally, for non-populist respondents, the anti-elitist cue has no effect on credibility perceptions of high-quality news, but for populist respondents, credibility is significantly reduced. These results imply that populist citizens’ judgments of the reliability of news stories from reputable news organizations might be decreased, in particular by anti-elitist cues.

**Discussion**

In high-choice media environments, citizens are confronted with a multitude of “snack” news messages (Schäfer, 2020), the credibility of which they must judge in seconds. In these settings, short media attacks are somewhat ubiquitous (Carlson, 2016). Scholars are particularly concerned about possible detrimental consequences of two specific attacks, that is, anti-elitist and fake news cues (e.g., Egelhofer &
Lecheler, 2019; Fawzi, 2020). This article proposed that by triggering social heuristics concerning the fallibility and partiality of news, fake news and anti-elitist cues might impact credibility perceptions of the accompanying messages. We also argued that to make well-founded statements about the actual influence of these cues, one must mimic the digital environment in which these attacks occur. That means one must acknowledge that these cues never occur in isolation but always in combination with various factors that also affect news message perceptions. Accordingly, we made use of a factorial survey that allows us to manipulate several factors simultaneously and thus represents a methodological design that is more geared toward the high-choice media environment.

In the most general sense, our results suggest that both media criticism cues have limited stand-alone effects. Instead, citizens’ credibility perceptions of news messages mostly depend on the source and covered issue (stance). This suggests that our assumption that social media users rely on the most salient—attacking—cues to evaluate the fairness and accuracy of a news story is incorrect. Instead, this indicates that heuristics related to the source and the content of the message appear to be more salient in the voters’ minds and more efficient in assessing the credibility of a news preview.

The limited effects of the critical comments we do find depend largely on whether a message is congruent with the participants’ attitudes. That is, when citizens...
disagree with the message’s issue stance, a fake news cue or an anti-elitist cue is influential in (further) decreasing its credibility. Similarly, when respondents have strong populist attitudes, the anti-elitist cue is more likely to be effective. In other words, when news attacks confirm voters’ pre-existing attitudes, they are more likely to be affected by them.

A deeper look at the data allows us to postulate that there is an idiosyncratic effectiveness in both cues to discredit one’s opponent. While, at first, the two cues seem to function very similarly, we find that in our data, the effectiveness of the fake news cue cannot be predicted by strong populist attitudes, while this is the case for the anti-elitist cue. This suggests that while both media attacks cues are currently most prevalently used by populists (Engesser et al., 2017; Fawzi, 2020; Hameleers, 2022), only the anti-elitist cue can be seen as a purely populist narrative with a longstanding history (e.g., Mudde, 2004), rendering populist citizens more susceptible to it. The fake news cue, on the other hand, has been used to discredit the media for only a few years now (e.g., Egelhofer & Lecheler, 2019). It thus might not be as strongly connected to populist ideology (yet). It, furthermore, is part of a larger debate, connected to a multitude of concepts that somehow relate to falsehood, which renders it a more fuzzy form of criticism that is used by all sorts of actors in public life to discredit information from different sources (e.g., Brummette et al., 2018; Egelhofer et al., 2020).

Figure 3. Average marginal cue effects for different sources.
Note. Point estimates denote marginal effect of each cue. Based on random intercept regression model. 90% confidence interval.
Our findings regarding the effectiveness of the anti-elitist cue advance previous research laying out the complex relationship between populists, their voters, and the media (Schulz et al., 2020). They indicate that anti-elitist media attacks are not without consequences. In the long run, exposure to these attacks might foster an emerging polarization of media perceptions between populist and non-populist citizens (Fawzi, 2020), driving them away from legacy media and into using more partisan media (Stier et al., 2020).

Lastly, our findings relating to the effectiveness of fake news and anti-elitist cues in connection to the more traditional cues in our setting, that is, the topics and sources of news messages, provide a complex picture. Regarding the news source, in our data, we only find a significant positive backlash effect of the combination of both cues on the credibility of messages by the outlet Telegraaf, a right-leaning newspaper often linked to populist views (Rooduijn, 2014). As in real life, this outlet probably is not the target of such attacks. The combination of cues might have been perceived as exaggerated and thus unrealistic by participants. Regarding the different issues, we find that only the effectiveness of the fake news cue mostly differs between the topics of news coverage. That is, the fake news cue is only marginally effective in decreasing the credibility of news messages that report on integration, while it appears to have no effects on stories about immigration, and even a positive backlash effect on the credibility of

Figure 4. Average marginal cue effects for populist and non-populist voters.
Note. Point estimates denote marginal effect of each cue. Based on random intercept regression model. 90% confidence interval.
news on the EU and on climate change. A possible explanation for these findings could be that the believability of the fake news cue may depend on the general (dis-)information environment, that is, whether there is a high amount of perceived disinformation in a specific political debate (Humprecht, 2019). For instance, because in Europe there is general consensus on the existence of human-made climate change, citizens might perceive attempts to discredit messages on climate change as “fake” as rather blatant attempts to undermine the message’s credibility, which causes them to react with a more positive evaluation of this messages’ credibility. For the topics of immigration and integration, on the other hand, there is more opinion polarization (Eberl et al., 2018)—and fake news stories frequently target (Islamic) immigrants in Europe (Humprecht, 2019). This might explain why respondents accepted the fake news cue when it was attached to messages on integration, which focused on whether Islamic schools are beneficial for integration. However, it does not explain why the cue was not efficient in decreasing the credibility of news about immigration, which covered crime numbers among refugees. Overall, these results call for further research and stress the necessity of testing specific media effect mechanisms under multiple contextual settings using multiple messages, issues, and sources (Slater et al., 2015).

Finally, we want to highlight the advantages of FS for exploring individual-level media effects. While there is currently a clear preference to use survey experiments to test political communication effects, FS designs offer a more comprehensive view of what exactly predicts change in individual-level attitudes and opinions (Otto & Glogger, 2020). This design thus allowed us to juxtapose the effectiveness of the fake news cue vis-à-vis an anti-elitist cue while considering several contextual and individual-level moderator variables.

Naturally, this study also has limitations. First, while our experimental design, employing made-up tweets, offers a fruitful approach to studying the effects of specific message factors, it naturally sacrifices ecological validity. That is, our design does not allow for studying the way individuals are indeed exposed to and react to politicians’ news message criticism (Bayer et al., 2020). For example, the algorithmic composition of individuals’ social media streams is influenced by a multitude of factors, “many of which are in a state of constant and dynamic flux” (Bayer, p. 486; see also Kümpel, 2022). In these streams, (hard) news is only one minor component of the social media environment (Kümpel, 2022), which renders the possibility that users might overlook these posts in favor of more entertaining content quite likely (Kümpel, 2022). The chance for this might be even higher on less “news-heavy” platforms than Twitter, such as Instagram. Future studies should consider how much attention news and its criticism receives when accompanied by other formats.

Next, while we were able to consider several determinants in this FS, of course, additional factors are worth considering, such as the source of the media criticism. In this study, we used an anonymous politician and did not take into account the role of partisanship. Including partisan cues in the multiparty system of the Netherlands would have increased the complexity of our FS design substantially. However, based on our design, we cannot rule out that partisanship or ideological alignment with the source plays a role in explaining the impact of media criticism. At the same time, we
avoided integrating factors that unmistakably dominate all other factors in the current design (Auspurg & Hinz, 2015). Most importantly, the addition of a party cue, which in itself is very relevant to the wider question of media accusations, could in our study have overridden all other effects in our experiment, because party cues have extremely potent effects (Petersen et al., 2013), especially when they originate from extreme parties like the Dutch populist radical right (Harteveld et al., 2022). While the current experiment thus omits an important factor that might impact credibility evaluations and give more insights into the schemata at play in processing them, it does allow us to tease out the extent to which a fake news cue and an anti-elitist cue, impact credibility evaluations of journalistic content, regardless of the political source. We do urge future research to take party cues and ideological alignment into account in assessing the impact of discrediting media cues.

In addition, we study four contentious issues and assume that these issues are of importance to all our respondents because of a shared information environment. However, individual differences in issue importance might impact the extent to which discrediting news cues are effective. We advise future studies to further take this into account. Also, we did not test all possible combinations of our factors for this survey. Instead, we chose a maximum of four vignettes per participant to not render the design too depleting and to avoid exposing respondents to conflicting vignettes in the same deck. In addition, while we used a large sample, including more than the recommended number of respondents per deck (Auspurg & Hinz, 2015), the statistical power was too low to detect some interactions we were interested in. While still drawing tentative conclusions regarding these interaction effects, we advise future studies using a similar design to increase statistical power.

In sum, this study shows that the attacks we studied can effectively discredit a news story, especially among specific news consumers. When citizens evaluate a news story on its credibility, they do not only take into account the source of the story and the topic at hand but evaluative cues as well. Our results suggest that both cues are most effective when attached to incongruent messages, highlighting the importance of confirmation bias in today’s high-choice media environment (e.g., Van Aelst et al., 2017). However, beyond this similarity, they both have distinct underlying mechanisms that determine their effectiveness. The anti-elitist cue is inherently a criticism of the journalistic actors, accusing them of ill motives. As the more purely populist narrative, this cue is thus more effective for populist citizens. The fake news cue, on the other hand, is heavily connected to a broader disinformation debate and can thus be rather understood as a criticism of the veracity of information. Therefore, its effectiveness seems to vary at least partly between issues. While the found effects of the discrediting cues are small, they give cause for worry in light of the increase in fake news accusations and anti-elitist media attacks, as they may be characteristic of a larger shift in how information is evaluated in information-rich democracies.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Funding
The authors received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Linda Bos https://orcid.org/0000-0002-0160-3490

Supplemental Material
Supplemental material for this article is available online.

Notes
1. Examples can be found in Supplemental Appendix B.
2. After data cleaning (see below). Auspurg and Hinz (2015) suggest 50 respondents per deck as a rule of thumb for power in factorial surveys. We aimed at 60 to be on the safe side.
3. Replication materials are available at https://osf.io/ex46n/
4. The 25% respondents taking less than 5 min to complete the survey.
5. Additional analyses and robustness checks are available in the Supplemental Materials.
6. The moderation analysis shows that the impact of the fake news cue on the integration news headline differs significantly from the impact of the fake news cues on the EU and climate change issue (b immigration=0.571, se immigration=0.310, p=.065; b EU=1.033, se EU=0.279, p<.000; b climate=1.279, se climate=0.307, p<.000). The backlash effect of the fake news cue on the climate change message only exists in comparison to the EU message.
7. It has to be noted that our additional analyses (Supplemental Table A5, Model I and Figure A1–A3) show that the perceived party of the politician barely has an impact on our findings. Only when the Twitter account is perceived to belong to populist radical right leader Thierry Baudet the fake news cue has a small negative impact on message credibility.

References
Bellur, S., & Sundar, S. S. (2014). How can we tell when a heuristic has been used? Design and analysis strategies for capturing the operation of heuristics. Communication Methods and Measures, 8(2), 116–137. https://doi.org/10.1080/19312458.2014.903390


Schulz, A., Wirth, W., & Müller, P. (2020). We are the people and you are fake news: A social identity approach to populist citizens’ false consensus and hostile media perceptions. *Communication Research, 47*(2), 201–226. https://doi.org/10.1177/0093650218794854


**Author Biographies**

**Linda Bos** is an Associate Professor of Political Communication at the Amsterdam School for Communication Research (ASCoR), University of Amsterdam. She is an expert in the broad field of populist political communication, studying the role of communication in the success of populist parties. In the general sense, her work centers on the relation between political elites, media and voters, focusing on the content and effects of political communication for democracy. Linda published 30+ articles in ISI-ranked journals such as *Political Communication, Party Politics, European Journal of Political Research, International Journal of Press/Politics* and *Political Psychology*. 
Jana Laura Egelhofer is a Postdoctoral Researcher at the Department of Media and Communication of the Ludwig-Maximilian University (LMU) Munich with a research focus on political communication and science communication. Her main research topics include mis- and disinformation, media criticism, anti-science discourse, and audience perceptions of science and news media. She holds a PhD from the Department of Communication, University of Vienna. Her work has been published in international journals such as *Journal of Communication*, *Annals of the International Communication Association*, *Journalism Studies*, *Media and Communication* and *Journal of Language and Politics*.

Sophie Lecheler, PhD, is Professor of Communication Science with a focus on Political Communication at the Department of Communication at the University of Vienna. Her research focuses on digitalization of political journalism, emotions, and news processing. Her work has been published in a wide range of international journals, such as *Communication Research*, *Journal of Communication, New Media & Society*, *Journalism Studies*, *Journalism & Mass Communication Quarterly*, *Journal of Media Psychology*, *International Journal of Public Opinion Research*, *Journalism*, *Communication Monographs*, *Communication Yearbook*, and *the International Journal of Press/Politics*. 