

FULL PAPER

Detecting misinformation in online social networks: A think-aloud study on user strategies

Erkennung von Fehlinformationen in sozialen Onlinenetzwerken: Eine Think-Aloud-Studie zu Strategien der Nutzer*innen

Isabelle Freiling

Isabelle Freiling (M.A.), Department of Communication, University of Muenster, Domplatz 6, 48143 Muenster, Germany; Contact: i.freiling(at)uni-muenster.de. ORCID: <https://orcid.org/0000-0002-6046-4005>

Detecting misinformation in online social networks: A think-aloud study on user strategies

Erkennung von Fehlinformationen in sozialen Onlinenetzwerken: Eine Think-Aloud-Studie zu Strategien der Nutzer*innen

Isabelle Freiling

Abstract: Although online social networks (OSN) facilitate the distribution of misinformation, one way of reducing the spread of false information in OSN is for users to detect it. Building on the framework of how audiences act to authenticate information, this study provides a user perspective on which strategies people use in evaluating information in OSN. In 15 qualitative interviews, participants were asked to think aloud while evaluating whether the content of posts from their own newsfeeds and of interviewer-supplied posts was true or false. Their answers were analyzed to determine which evaluation strategies they used. Analyzing participants' thoughts as they evaluate information is more reliable than directly asking participants which strategies they think they use. Results show that users' strategies in information evaluation are *searching for more information*, *knowledge of account or content carries the most weight*, and *every detail needs to fit*. A comparison of strategy usage for posts from befriended versus unknown personal accounts as well as for posts from followed news outlets versus not followed news outlets shows that for posts from followed news outlets, *knowledge of the account* was the most-used strategy followed by *knowledge of the content*. For other types of posts, strategy usage varied more widely and depended on each post. This highlights the importance and possible higher ecological validity of research on posts from news outlets that users actually follow, as users' experiences with previous posts seem to play a major role in how they go about evaluating information in new posts.

Keywords: misinformation, online social networks, qualitative interviews, strategies, user perspective

Zusammenfassung: Soziale Onlinenetzwerke erleichtern die Verbreitung von Fehlinformationen. Eine Möglichkeit die Verbreitung von Fehlinformationen zu reduzieren ist, dass die Nutzer*innen sie als solche erkennen. Diese Studie baut auf dem Audiences' Acts of Authentication-Konzept auf und bietet eine Nutzer*innenperspektive bei der Informationsbewertung in sozialen Onlinenetzwerken. In 15 qualitativen Interviews wurden die TeilnehmerInnen gebeten, laut zu denken, während sie den Wahrheitsgehalt des Inhalts von Posts aus ihrem eigenen Netzwerk sowie vom Interviewer mitgebrachten Posts einschätzen. Die Methode des lauten Denkens, mit der die Gedanken der Teilnehmer*innen während der Informationsbewertung analysiert werden können, liefert verlässlichere Ergebnisse als eine direkte Abfrage der Strategien. Die Ergebnisse zeigen, dass Nutzer*innen folgende Strategien zur Bewertung des Wahrheitsgehalts eines Posts anwenden: *Nach weiterer Information suchen*, *Wissen über den Account oder Inhalt ist am Wichtigsten*, und *jedes Detail muss zusammenpassen*. Ein Vergleich der Nutzung der

Strategien für befreundete versus unbekannte persönliche Accounts sowie Nachrichtenmedien, denen die Nutzer*innen folgen versus nicht folgen, zeigt: Für Posts von Nachrichtenmedien, denen die Nutzer*innen folgen, ist *Wissen über den Account* die meistgenutzte Strategie, gefolgt von *Wissen über den Inhalt*. Für die anderen Arten von Posts variiert die Wahl der Strategie mehr und hängt vom jeweiligen Post ab. Das zeigt, wie wichtig Forschung über Posts von Nachrichtenmedien ist, denen die Nutzer*innen tatsächlich folgen, da die Erfahrung mit bisherigen Posts eine wichtige Rolle in der Informationsbewertung spielt und somit eine solche Forschung ökologisch validere Ergebnisse liefern kann.

Schlagwörter: Fehlinformation, soziale Onlinenetzwerke, qualitative Interviews, Strategien, Nutzer*innen-Perspektive

Author notes: This research was supported by the German Research Foundation [research training group “Trust and Communication in a Digitized World,” grant number 1712/2]. I would like to thank my colleagues at the research training group, as well as Annie Waldherr and Dietram A. Scheufele for their valuable input on previous versions of the manuscript and Sevda J. Narbutas for her help during data collection.

1. Introduction

In newsfeed-based online social networks (OSN) such as Facebook and Twitter, each user can easily post and share short messages to a broad audience. This allows information as well as misinformation (i.e., false information) on diverse topics – e.g., voting (Kuklinski, Quirk, Jerit, Schwieder, & Rich, 2000) and vaccines (Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012) – to spread quickly (Del Vicario et al., 2016; Vosoughi, Roy, & Aral, 2018).¹ If users receive false information repeatedly, they are more likely to believe the content to be true (Lewandowsky et al., 2012). Further, it is difficult to correct false information in users’ minds (Bode & Vraga, 2015; Nyhan & Reifler, 2010; Thorson, 2016; Walter & Tukachinsky, 2019).

In this context, the current study focuses on how to prevent the distribution of false information in OSN in the first place. One way is for users to detect false information as such, and, thus, to not disseminate it. Users are unlikely to disseminate information they detect as false (as long as such false information is not in line with the individual’s attitudes, i.e., partisan selective sharing; Shin & Thorson, 2017), because they tend to fear of how their OSN will react (Neubaum & Krämer, 2018).

Several lines of research provide insights into how users can try to detect false information. For example, competence-based research focuses on the medium (critical media literacy), whereas other research focuses on general information processing and evaluation or even on evaluating journalistic content specifically, since in OSN posts are created not only by friends but also by news outlets. Research on

1 For false information to be misinformation, it does not matter whether it is false due to a mistake or due to someone’s intent to deceive the receiver; conversely, for false information to be disinformation, an intent to mislead is necessary (Lazer et al., 2018). As this study focuses on detecting false information, it is irrelevant why the information is false, i.e., whether it is due to a mistake or whether it is intended to mislead the recipient. However, one could argue that when trying to detect false information users consider whether the author wanted to deceive the recipient. Thus, in this study, I use “false information” when only “false information” is meant.

critical media literacy indicates that when individuals have high critical media literacy, they are able to detect false information because they know how to evaluate which information is credible (e.g., Livingstone, 2004; Metzger, Flanagin, Markov, Grossman, & Bulger, 2015). Individuals with a high critical media literacy can transfer this knowledge to new media environments such as OSN, and they are more skeptical of information quality (Vraga & Tully, 2019). Thus, enhancing users' critical media literacy for OSN helps to prevent individuals from believing false information. The research on information processing and evaluation has aimed to explain generally how individuals process information and form an opinion on it, using, for example, their biases, past experiences, and motivations (e.g., Chaiken, 1980; Lewandowsky et al., 2012; for further elaboration see section 2). Finally, communication research gives insight into how recipients assess the credibility and trustworthiness of journalistic content (e.g., Chung, Nam, & Stefanone, 2012; Sundar, 2008). For example, when individuals want to form an opinion about a journalistic product, one way to do this is by evaluating the news outlet (Metzger, Flanagin, & Medders, 2010; Pennycook & Rand, 2019). When evaluating a news outlet, individuals can search for potential biases or whether general reporting standards are met (Newman & Fletcher, 2017). This approach works less well for posts from befriended, non-journalistic accounts; here, OSN users could consider the types of posts, comments, shares, likes, the account, content, etc., instead.

When considering how users identify false information in OSN, however, the above-described research on the evaluation of general information and on the evaluation of journalistic content needs to be brought together, since OSN have some characteristics similar to mass communication and some similar to interpersonal communication (Sundar, 2008). A rather bottom-up approach for determining how OSN users detect false information is to simply look at what they currently do to detect it (e.g., Tandoc et al., 2018) and then, in a further step, examine whether what they are doing leads them toward detecting false information successfully.

Based on the framework of the audience's acts of authentication (Tandoc et al., 2018), people assess whether something is true or false using up to two steps of authentication. The first step is internal evaluation, which encompasses strategies concerning the individual's knowledge and experience as well as their judgement of the characteristics of the post and the source. If strategies in this step do not allow the individual to decide whether the post contains true or false information, individuals continue with the second step, external authentication. This step can involve intentional or incidental authentication as well as interpersonal or institutional authentication (Tandoc et al., 2018).

The current study builds on the audience's acts of authentication framework (Tandoc et al., 2018). By using a more granular method than used by Tandoc et al. (2018), in which only one open question in an online survey was used to ask participants how they verify news posts on OSN, this study aims to provide more extensive insights into how OSN users try to detect false information. Furthermore, the current study examines whether users apply different detection strategies when assessing posts by personal accounts (a typical feature of OSN) that they are friends with versus personal accounts that they do not know or when assessing posts by news outlets that they follow versus news outlets that they do not follow. To do so, this study applied qualitative inter-

views in which participants thought aloud as they evaluated whether the content of posts from their own newsfeeds and of interviewer-supplied posts was true or false.

2. Information processing and evaluation

False information can be persuasive, especially when it is in line with attitudes held by the individual. Dual-processing models describe two types of information processing that may influence how easily persuaded an individual is (e.g., Chaiken, 1980; Petty & Cacioppo, 1986). For this study, the persuasion part of dual-processing models is less relevant, but these models inform the current study by providing insights into the two types of how information is processed. The different models have in common that one type of information processing is rather fast, whereby individuals do not have to invest much effort; an example is to assume that experts or known sources must be true. This type of information processing is often labelled as “heuristic,” “experiential,” or “peripheral.” The other type of information processing is less automatic and more sophisticated, and it requires individuals to invest more effort into evaluating the information. An example of this type of processing is checking the arguments in and details of a message. This type of processing is often called “systematic,” “rational,” or “central” (Chaiken, 1980; Kirkpatrick & Epstein, 1992; Petty & Cacioppo, 1986). Because it is impossible for individuals to use systematic processing for every piece of information, as this would require too much effort and time, individuals only use this type of processing if they have a motivation that is high enough to invest more in information processing. This motivation could come from, e.g., how important the topic is to them, whether it is in line with their attitudes, or whether they want to be accurate (Ajzen, 1996; Kunda, 1990; Nir, 2011).

Dual-processing models focus on whether individuals process information heuristically/peripherally or systematically/centrally, which would be interesting to see in OSN. However, OSN users probably do not clearly distinguish between their processing types when evaluating whether a post is true or false, perhaps because they may only pay attention to each post for a few seconds (Vraga, Bode, & Troller-Renfree, 2016). Thus, the following section focuses on internal versus external information evaluation.

The audience’s acts of authentication framework (Tandoc et al., 2018) distinguishes between whether an individual takes only one step (internal evaluation) in assessing information or two steps (first an internal evaluation followed by an external evaluation). Internal evaluation can encompass the individuals themselves (their knowledge, experience, instinct), the message characteristics such as tone or logic, the source credibility, or popularity cues that come along with the message.

A strategy that individuals employ when evaluating information in the context of their own knowledge is to check whether the information is compatible with what they currently know or believe (Lewandowsky et al., 2012; Tandoc et al., 2018). As checking whether incoming information that aligns with one’s own knowledge is effortful, it requires motivation. Conversely, identifying whether incoming information conflicts with one’s own knowledge is easier to spot, as this leads to negative feelings such as cognitive dissonance (Festinger, 1957), and the processing of inconsistent information is less fluent. In addition to comparing the

information with their own knowledge, users may also draw on their experience with what a certain account typically posts. If, for example, a user knows from their uncle's previous posts that he usually posts something they believe is false, this may influence their future evaluations.

When users do not have any expertise on the subject of the post or when they are not motivated enough to invest more effort, users may opt to look at the source credibility, which is a detection strategy that operates on the source level (Lewandowsky et al., 2012; Tandoc et al., 2018). A typical example is that an established news outlet, as a well-known institution, is deemed credible (Tandoc et al., 2018). But, as there are also personal accounts on OSN, users may consider other account characteristics to determine the credibility of personal accounts compared to professional news outlet accounts. In comparison to journalistic accounts, the operator of a personal account does not necessarily have a professional education in evaluating information themselves. If an account holder cannot determine whether information they are posting is true or false, this can be problematic because other users trying to evaluate such a post may assume that accounts only share information they consider as true.

Additionally, when users personally know the author of a post (which is often the case for befriended accounts), they may also have a personal relationship with them, and this relationship might influence how thoroughly the information is evaluated. Whereas a post from a weak tie (e.g., a casual acquaintance) may be evaluated carefully, a post from a strong tie (e.g., parent, best friend, etc.; Neubaum & Krämer, 2017) may not be scrutinized so closely.

To be able to evaluate source credibility, users need to identify the source of a post, which can be difficult as there are different kinds of posts: original posts, reposts, posts linking to a URL, posts quoting others, videos or pictures, and a combination of these elements. Especially when there are combinations of the different types of posts, source identification can be difficult, as users may identify several possible sources, e.g., the reposting account, the original account or the website behind the URL. When the source is identified but unknown (which happens when someone posts in a group the user follows), the user may evaluate the source credibility using incidental aggregate representations such as the poster's number of friends or followers (Walther & Jang, 2012), which closely connects this level to the popularity level. Other information can also be used to determine the credibility of a source, such as the profile picture, the account's name, the information provided by the account holder or, as mentioned above, previous posts.

On the message level, users' strategies for detecting false information may include checking whether the information itself is coherent, i.e., whether there are no inconsistencies within the message (Lewandowsky et al., 2012; Tandoc et al., 2018). In addition, individuals can interpret the tone of the message, for example, whether it is not too polemical (Tandoc et al., 2018).

On the popularity level, a lower-effort strategy is for users to check whether others consider the information to be true, i.e., whether there is consensus (Lewandowsky et al., 2012). For this strategy, OSN users may consider popularity cues such as likes and shares (Tandoc et al., 2018). Further, since OSN users may personally know the posting account, they may consider not only the number of likes and shares but also *who* liked and shared a post; this is because users have different

relationships with other users (weak vs. strong ties; Neubaum & Krämer, 2017). Since previous research has not focused on strategies that OSN users employ for evaluating the information in posts by their friends, users may adopt as yet unidentified strategies for detecting false information in their friends' posts.

All of the strategies discussed above involve the first step of authentication, namely internal evaluation. The second step, external evaluation, is used when internal evaluation does not lead to a satisfactory result, meaning that the individual still cannot identify whether a piece of information is true or false. External evaluation can be intentional (users actively search for further information, such as by a Google search or asking someone about the issue) or incidental (users may wait passively for information on the topic to come along) as well as interpersonal or institutional (Tandoc et al., 2018). As OSN users may be able to distinguish consciously between internal and external authentication, they also may be able to voice this as they go through their authentication process. Thus, the current study builds on the audience's acts of authentication framework.

As discussed above, users can employ different strategies in evaluating information in OSN posts. However, the above-mentioned strategies derive either from contexts outside of OSN or within OSN but with a focus on news items only. Importantly, news items are not the only kinds of posts in newsfeed-based OSN; instead, communication in newsfeed-based OSN falls in between mass-mediated and face-to-face communication. Newsfeed-based OSN "combine institutionally authored messages alongside individually authored messages, while social groups and networks of many sizes and natures are also frequently involved in the retransmission and reformulation of these messages" (Walther & Valkenburg, 2017, p. 416). While users may follow accounts owned by traditional mass media enterprises or politicians, OSN users also engage in elements of interpersonal communication such as discussion with friends. Thereby, a typical characteristic of OSN – having posts from befriended accounts that the user may personally know and have a certain relationship to – has not been examined in the above-mentioned studies. Since a personal relationship with the author of a post may change the strategies a user applies to evaluate the information, this study incorporates OSN posts from both, personal and news outlet accounts. This allows for determining whether users employ different evaluation strategies with befriended accounts and news outlets. Apart from the interpersonal character of some OSN posts, other characteristics of OSN posts may also change the strategies users employ to determine whether a post contains true or false information.

Thus, using a qualitative approach that allows participants to evaluate posts by befriended accounts in their own network as well as to evaluate interviewer-supplied posts from outside their network, this study will answer the following research questions (RQ):

RQ1: Which strategies do OSN users employ when evaluating whether the content of a post is true?

RQ2: Is there a difference in the strategies they use for posts by a) befriended versus unknown personal accounts, and b) news outlets that the users follow versus news outlets that the users do not follow?

3. Method

The example OSN used for this study was Facebook, as in Germany (where the study was conducted) this platform is the one used most for informational purposes (Hölig & Hasebrink, 2018). Participants were interviewed face to face in a lab containing computers, because participants used the computers to evaluate Facebook posts. They were asked to think aloud while evaluating whether the information in a Facebook post was true or false. Analyzing participants' thoughts as they evaluate the information is more reliable than directly asking participants which strategies they think they use to determine whether something is true. The interviewer was the author of this study, as this enabled the interviewer to ask for more details on relevant points and questions that popped up during the interview.

At first, the interviewer explained thinking aloud in detail to the participants, building on a formulation by Ericsson and Simon (1993, p. 378) that was adapted to the current study. Shown below is an English translation of the German version used in the study:

“Within this study, I am interested in what you think about when solving the tasks that I am going to give you. In order to do this, I am going to ask you to think aloud as you work on the tasks. What I mean by *thinking aloud* is that I want you to tell me everything you are thinking from the time I give you the task until you find a solution for it. I would like you to think aloud constantly, the whole time. I do not want you to try to plan out what you will say or try to explain to me what you are saying. Just act as if you are alone in the room speaking to yourself. Most importantly, do not stop talking. If you are silent for too long, I will ask you to continue talking. Do you understand what I mean?”

In order to practice the method of thinking aloud, participants received two training tasks (Ericsson & Simon, 1998). These were as follows: 1) ‘how many windows are in your parents’ house? You do not need to count them yourself; I will do that for you. Just tell me which room has how many windows’ (Ericsson, 1987). Whereas some participants were already very good at thinking aloud and gave the interviewer a mental walk-through of their parents’ house, others still needed practice in voicing their thoughts, as they just counted each window to get to the final number quickly. 2) ‘Open the city’s website and search for information about the traffic. Do this until you feel well informed. Please tell me what you are looking at and what you are thinking when looking at something. I do not want a description of what you see but rather your thoughts that result from what you are seeing. If you want to show something to me, just use the mouse cursor.’ The second task not only trained the participants in thinking aloud, but it also made them familiar with thinking aloud when being online. Moreover, this task trained participants in using the cursor for showing something to the interviewer. This was useful for the data analysis, as it allowed for recording (via recording the monitor) what participants wanted to show, which can be matched with the recorded speech. Thus, the researcher was able to see what participants were referring to.

After completing the training tasks, participants logged into their own Facebook account. The interviewer and participant searched together for the first post that was possible to discuss with respect to whether the content was true or not.

Participants were asked to evaluate whether the information of the post seemed true or false to them while thinking aloud: ‘Please take a look at this post and think about how you can determine whether you think the post is true or false. Please think aloud.’

Logging into their own networks, participants were able to see posts that were representative of posts they usually viewed in their network, which provides high ecological validity. For example, in their own networks, they may know the account holder (from face-to-face situations or just from previous interaction online) and have had experience with their previous posts. Familiarity with an account can influence how much a user trusts the account (Cheng, Fu, & de Vreede, 2017) as well as give the user an impression of how credible posts from that account are. If participants are only provided with interviewer-supplied posts from news outlets that participants may or may not know, or even from made-up accounts that they clearly cannot know, this creates a situation that is far less ecologically valid.

However, going into a participant’s own network does not ensure that there will actually be posts containing false information. Thus, in a second step, participants received five interviewer-supplied posts, of which three contained false information. Of course, here the participants (may/do) not know the accounts (for a detailed description of the posts, see section 3.2). For the situation with the interviewer-provided posts, the same interview procedure as described above was used.

Another reason the first part of the interview involved participants searching for posts in their own newsfeed was because participants may find it much more credible. The participant knew that the researcher could not have manipulated the posts from their own feed, as during the experiment, the participant and the researcher searched for them together. This is different from the posts the interviewer showed the participant in the second part of the interview; here, the participant saw those posts as images of screenshots. The participant then could not know whether the researcher manipulated the posts but could suspect it. Thus, participants were likely to be far more skeptical towards the interviewer-supplied posts, and this heightened skepticism should be considered in the interpretation of the data. However, by including a mixture of accounts that participants knew and did not know and posts that contained true or false information, this study allowed participants to employ all the strategies in their repertoire for detecting false information in OSN posts from both known and unknown accounts.

During a think-aloud interview, some of the participants’ thoughts may vanish before they can be voiced. Moreover, what participants are talking about is highly driven by their thoughts, which may result participants not answering all of the researcher’s questions without further prompting. Thus, retrospective interviews are commonly used in combination with thinking aloud (Charters, 2003). Immediately after the interviews, participants were surveyed retrospectively with regard to questions that they had not yet answered. Those questions were on, for example, how they decided on evaluating the information, how they define trust or credibility if they mentioned the words in the interview, what their relationships and experiences with the accounts were, and whether they would like or share the post. At the end of the interview, participants answered a written ques-

tionnaire concerning how much prior knowledge they had on the topic of each post, how important the topic was to them, how much they agreed with the content of each post, the level of their Facebook and media literacy, and their media use. Afterwards, participants were debriefed, as some of the posts contained false information. The debriefing made clear which information was false and which was correct in the interviewer-supplied posts, and participants were asked whether they understood this. Participants were compensated for their time with 25€.

During the interviews, both the participant's speech and the computer monitor were recorded. An ethical challenge did arise, as participants had to log into their Facebook account and thus reveal sensitive private information. To minimize the intrusion of participants' privacy, when participants were logged into their own account, only the part of the monitor where the posts were visible was recorded.

3.1 Sample

Sampling followed a grounded theory approach. In order to find all the strategies used by OSN users to evaluate whether a post is true or false, the sample should be as diverse as possible in the characteristics that are assumed to influence the evaluation, which are age, sex, education, occupation, and Facebook use. Hence, participants were sampled after answering an online survey on age, sex, education, occupation, and Facebook use. They were selected according to their answers such that the sample was as diverse as possible; sampling was stopped at $N = 15$ due to theoretical saturation.

All participants were Facebook users, as they were asked to evaluate posts from their own newsfeed. Participants were between 18 and 57 years old, about half of them ($n = 7$) were female. Most participants used Facebook daily, although some participants only used it once a month or even less than once every 2–3 months. Participants were relatively evenly distributed in their occupation: a third worked full or part time, about another third was studying or was in professional education, and the last third was retired, job seeking, or a housewife/-husband. Slightly more participants had a high level of education than had a middle or low level of education.

3.2 Newsfeed posts/interviewer-supplied posts

The posts used for assessing participants' strategies for evaluating information during the interview did not cover a specific topic. In fact, no topic was restricted, as false information can exist on any topic and is not restricted to arenas such as politics or health communication. Furthermore, users' strategies may differ because of the topic (e.g., due to their prior knowledge or because they think it is important or interesting). The topics of the posts included politics, refugees, security, work, education, health, travel, entertainment, sports, society, culture, nature, and local news.

In the first part of the interview, the posts stemmed from the participants' own network, i.e., either from their newsfeed (see Fig. 1 for an example) or from befriended accounts that post presumably false content from time to time (see Fig. 2 for an example). This created a situation that is much more natural and valid

compared to most experiments (where participants usually do not know the posting accounts), with respect to the user's relationship and experience with an account: In their own network, users usually know the account holders, except perhaps for accounts that are posting in public groups or for posts by friends of friends. Having experience with an account's posts may influence the evaluation of trust and credibility. This may also be the case for news outlets that users follow, in that users may trust them due to experience with their posts, and they probably follow that account for a reason.

The posts in the first part of the interview consisted of about three to four posts. The interviewer selected the posts while the participant was scrolling through the posts. It was searched for the first posts that were possible to discuss regarding whether the content was true or not. Thus, posts with no content (e.g., just pictures without any context) or sponsored content were excluded.

Figure 1. Example of a post from a participant's Facebook newsfeed.



Figure 2. Example of a post of an account that a participant expects to post false information.



In the second part of the interview, five interviewer-supplied posts were shown to the participants. Two of them were screenshots of real Facebook posts containing true information. The other three posts contained false information (see Fig. 3 for an example) and were manipulated to make sure the information was definitely false. This false information was based on true stories that were modified enough so that they were clearly false. These five posts were different in nature (type of post, social endorsement, expert accounts, etc.) to allow participants to employ different strategies when evaluating whether a post was true. The true posts were from a TV channel's news program and from a local newspaper of another German city. Hence, participants could be familiar with both of these accounts, although several participants confused the TV channel's news program with one of their satirical shows. Moreover, not every participant knew the local newspaper from the other city. The false posts were from made-up accounts so that false information was not included in some real account's post. As the condition 'true' was always affiliated with a news outlet as the account and 'false' information was affiliated with made-up accounts, this might lead to confounds: One cannot be sure whether users employed a certain strategy because the account was from a news outlet or because the content was true. The same goes the other way around for a false post and a made-up account that participants could not know. However, as many studies show how difficult it is to fully correct false information once it is embedded in users' minds (Chan, Jones, Hall Jamieson, & Albaracín, 2017; Lewandowsky et al., 2012; Walter & Murphy, 2018), one could never be sure – even with a very detailed debriefing – whether participants really remembered after the study that the false information was made up and did not really come from the account. Thus, ethical reasons led to the decision to not put false information into some real account's post. Additionally, to make sure that news outlets were represented in both the real and made-up accounts, one of the made-up accounts allegedly quoted a newspaper for doctors. Thus, this wrong quote was made by the account and not by an account of the newspaper itself. In the debriefing, participants were explicitly told that the newspaper did not write that false information.

The topics of the interviewer-supplied posts were drought, active volcanos, housing and construction plans, minimum price for alcohol, and car parking. The posts on the active volcanos and car parking were true posts, while the others contained false information.

Figure 3. Example of an interviewer-supplied post containing false information.

Andy P
2 Std. · 🌐

Die Ärzte Zeitung tätigt die Äußerung, dass Belgiens Einführung eines Minimalpreises auf Alkohol (60 Euro-Cent per Zentiliter purem Methanol) bereits positive Resultate verzeichnen konnte. Denn „[i]m annuellen Vergleich hat sich seit dem Datum der Einführung des Minimum-Preises, dem 01. Mai 2018, der Verkauf von Spirituosen um die Hälfte halbiert.“

🙄🤔 18 4 Kommentare 3 Mal geteilt

👍 Gefällt mir 💬 Kommentieren ➦ Teilen

Kommentieren ... 😊 🗨️

Hans Dresko Solange die nicht die Schokolade höher besteuern, nur zu!!! Aber bei Schokolade hört der Spaß auf!!!
Gefällt mir · Antworten · 2 Std.

Anne Fischmann Sophie Beepunkt gut, dass wir nicht in Belgien wohnen :D
Gefällt mir · Antworten · 50 Min

Weitere Kommentare anzeigen 2 von 4

3.3. Data analysis

In September 2018, the 15 qualitative interviews were conducted. The same researcher who conducted the interviews analyzed the statements from the interviews using qualitative content analysis. Having the same researcher conduct the interviews and do the coding gave the coder all the context information that might be needed for interpreting the statements correctly, i.e., as the participants meant them. The statements were coded inductively and deductively. Deductively, categories were created according to a) the results of a study on the characteristics users pay attention to when evaluating information on websites (Horstmann, Rösner, Conrad, & Heidemann, 2018), b) the strategies identified by Lewandowsky et al. (2012) and Tandoc et al. (2018), and c) certain heuristics that people may employ when processing and evaluating information on OSN, such as consistency, endorsement, self-confirmation, or expectancy violation (e.g., Metzger & Flanagin, 2013). Inductively, OSN-specific categories were added throughout the coding process. When a new category was added, all interviews were coded again to ensure they were also coded for the new category. Based on the categories and statements made by the partici-

pants concerning their general approach to evaluating information on OSN, the last coding step involved deriving the actual strategies.

4. Findings

The interviews showed that asking participants to evaluate a post as true or false is something they did not always find a suitable question. This is because OSN posts do not always contain information but often opinions. Although some of these opinions are based on information that can be true or false, participants found that opinions can be contrary to their own opinion but can rarely be wrong. Thus, they tried to separate opinion from information.

To answer RQ1, which asks about strategies OSN users employ when evaluating whether the content of a post is true, the following strategies were found in the qualitative content analysis of the interviews: *searching for more*, *knowledge carries the most weight*, and *every detail needs to fit*. As not all approaches fit into these strategies, the remaining approaches are discussed below the strategies. Afterwards, the usage of the strategies is described (RQ2).

4.1 Strategies

4.1.1 *Searching for more*

A strategy that participants commonly used in the interviews was to gather more information before deciding how to evaluate a post. This was done by googling the issue of a post or by clicking on a link provided in the post. Participants used this strategy especially when they were torn:

Because this does not look like something trustworthy. Maybe it is trustworthy anyway, [...] firstly, I would go to the [newspaper] or I'll google it or I'll just ask [refers to the account of the newspaper]. (P3)

There is even a video, I think. Or no, not exactly a video, but a link. In my opinion, somehow it makes it always a bit more credible when you can have a look at something more. (P11)

With this strategy, though, interest was key. Participants only used this strategy when the content was important or interesting to them. Otherwise, they said, they would just continue scrolling through their newsfeed without evaluating whether the post was true or false.

4.1.2 *Knowledge carries the most weight*

Another main strategy was based on the participants' knowledge. This was either knowledge about the content (expertise) or about the account (experience with it).

Knowledge of the content carries the most weight. If users had knowledge of or expertise on the content of a post and were very sure about their knowledge, they could then weigh the content of the post as being the decisive factor for evaluating it. In this case, the post's other characteristics did not play a big role in

their decision. Sometimes, when participants were completely certain that the content of a post was true, they did not even mention considering the account, an otherwise often mentioned characteristic in evaluating information:

[Post is on drinking methanol] Methanol. [...] Well. Pure nonsense. [...] Pure nonsense, but only because I know that methanol is utterly toxic. (P12)

In other cases, even participants' trust in the account increased when they read content that they knew was true:

Schulz against Gauland.² I know all of it. I was yesterday already, I mean, on the website of n-tv.³ I scroll through that website every noon. The information was there already. In the evening, I also watched a video about it, with the exact same sequence as here, showing how Mr. Schulz reacts and how the AfD reacts. Thus, I am a 100 percent sure that this is correct news and that, well, somehow increases my trust in the [local newspaper, which is the source of the post]. (P5)

This strategy is about knowledge, but it is impossible to have knowledge on every topic. Thus, when users did not have knowledge on the topic, they could not employ this strategy. In such cases, they instead tried to evaluate the post either by determining whether they found the content credible (which could be done using the strategies *search for more* or *every detail needs to fit*) or by determining whether the account might be trustworthy. Users can determine an account's trustworthiness, in turn, by having knowledge of the account or by clicking on the account and evaluating it, based on the information given there and by examining the account's previous posts.

Knowledge of the account carries the most weight. If users have knowledge of and experience with an account, this knowledge can also outweigh the content in users' evaluation of a post. This might be the case when the account holder has strong ties to the user (e.g., family member, good friend) or when the account is run by a news medium that the user trusts completely. In this case, the user would trust the content to be true just because it came from a certain account; at the same time though, the content cannot be too unrealistic.

And my friend is highly educated and she is researching *everything*, as well as I do. There is nothing we do not research before we talk about it. Thus, when I see what she is posting [...] I do not always have the same opinion as she does. This happens, but then I know 1,000 percent that what she is posting is well researched and that I can trust it. (P14)

I count on it that it is the truth. I just know it, because I used to work at this news agency myself [...] when someone was reporting incorrectly, I mean not on purpose, but just translating it incorrectly, then they were fired. Thus, I know for sure that it is the truth. (P15)

This also works the other way around, when the account is not trustworthy at all. In this case, participants would not even read posts from untrustworthy accounts:

I wouldn't even read it, if it wasn't for this study, huh? I wouldn't even have a look at [this friend's] page anymore, because I know what he writes. (P13)

2 Martin Schulz and Dr. Alexander Gauland are two German politicians. Martin Schulz is in the Social Democratic Party (SPD). Dr. Alexander Gauland is in the Alternative for Germany (AfD).

3 n-tv is a TV news channel.

Thus, to use this strategy, users must know the account; this strategy does not work with unknown accounts. When participants did not know the account, some tried to evaluate it by clicking on the profile.

Overall though, as the name of the strategy implies, participants using this strategy primarily relied on their knowledge (knowledge *has the possibility* of outweighing the other characteristics), but it means neither that knowledge *always* overruled everything nor that participants failed to consider anything else. Although participants gave priority to their knowledge in this strategy, they might consider other strategies.

4.1.3 *Every detail needs to fit*

Another strategy was that participants considered everything they saw. This included the content, account, social endorsement, type of post, date and time of post, writing style, etc. Participants using this strategy found it important that everything was consistent and plausible (also referring to their own knowledge on the issue). For example, they checked to make sure that the content was realistic, that there were no typos, and that there were no mistakes in citations (no expectancy violation). In addition, participants paid attention to whether the content of the post fit the account (e.g., that there was no persuasive intent) and whether the picture fit the content:

This one [refers to the post] is done a bit nicer, with some numbers, a picture, and, well, a name [refers to the name of the account] that sounds intelligent. [...] But, well, when you take a closer look, profile picture doesn't fit. Source is missing. Thus, it is not really reputable. [...] Most of all, one should've had heard of it somewhere else as well. (P3)

I can't really imagine that anything would be built on the lawns of this square. [...] Well, the absurdness [of the content] would be a reason for me (laughs) to perceive it as false. (P11)

4.2 Ways of evaluating information that are not covered by the strategies

As the strategies mentioned above represent broad means of arriving at a decision, they do not encompass every single way that users evaluate a post.

Besides evaluating if every detail fit, users also considered general cues for credibility that they might have learned through evaluating information in other media formats or in school. This included verifying tags and quotes or searching for pro and con arguments within a post.

I mean, it surely is trustworthy, because they have this blue hook [refers to the check mark]. It's the same as on Instagram, huh? (P2)

I'm looking at those posts like we did in school [...] with pros, cons and so on. Well, this guy seems relatively reliable to me as he is quoting correctly [...]. So, this would always influence me, the way it's written, and if there was something done the way I learned in school. (P11)

The last example shows that this approach of evaluating general credibility cues is close to the strategy *every detail needs to fit*. Although in this case certain par-

ticipants did consider the same content (here: the quote), they used it for a different mode of evaluation. In the strategy *every detail needs to fit*, they used it to evaluate the whole picture they had of a post, whereas here participant eleven used it as a general cue for credibility based on what he learned in school.

Participants also showed a bias in information processing not covered by the strategies, namely a bias due to their political preferences.

Maybe there are different political directions, which you don't trust, or let's say that you don't know certain pages. Of course, I would take a closer look at those. Yes. Perhaps you would assume that it's a false report first. (P4)

4.3 Usage of the strategies

Before addressing RQ2, this section elaborates on the general usage of strategies according to the characteristics of the post and the participants, and it describes some of the most often used strategies.

With the exception of one participant who only considered the content and had either knowledge of the topic or found the content to be unrealistic, all other participants used several strategies to evaluate the posts. For some posts, some participants even used all four strategies.

Even when participants did use *knowledge of account or content carries the most weight*, they often used this strategy along with other strategies. As this is a knowledge-based strategy, users can only employ it when they have knowledge regarding the content or account. Regarding the strategy *searching for more*, this one is used when a user has interest in a topic; they are only motivated to invest more time and effort via this strategy if the topic is interesting to them. When users lack motivation to search for more information or knowledge on the content or account, the only remaining strategy (of those identified in this study) is checking whether every detail fits. Hence, which strategy participants employed largely depended on their motivation and knowledge.

For the usage of strategies, neither the topic of a post (including knowledge on the topic and how important it is to users) nor the participants' willingness to share a post mattered. This is probably due to the study design, which aimed at covering diverse posts (and topics) without explicitly testing whether participants use different strategies to identify false information for different topics. In order to test whether the topic influences how users attempt to detect false information, several posts on one topic would have been needed. Further, the finding that willingness to share did not matter for participants' strategy use may be because users showed a very low willingness to share the posts used in the study: Only two participants were willing to share one or two of the posts they evaluated.

RQ2 asks whether the use of strategies differed for posts by news outlets that participants followed versus did not follow or for posts by befriended accounts or unknown personal accounts. For posts by news outlets, the results show that the evaluation strategies chosen differ when participants follow the news outlet in their own network and have them in their newsfeed, as compared to the news outlet post supplied by the interviewer. The strategy participants used for every

“in-network” news outlet post was *knowledge of account*. The use of this strategy sometimes aligned with the *knowledge of content* strategy, although, of course, participants had to decide then which *knowledge carries the most weight*. For the interviewer-supplied news outlet post, participants’ use of strategies was different. For those posts, the choice of strategy depended on the post.

Participants evaluated posts from personal accounts – befriended or unknown accounts – in the same way that they evaluated the interviewer-supplied news outlet posts; depending on the post, they used one or a set of the identified strategies. For unknown accounts, participants were not able to use the strategy *knowledge of account*.

Overall, to answer RQ2, the study showed that for evaluating the posts by news outlets that participants followed in their OSN, they relied heavily on their knowledge of the account; conversely, for evaluating posts by befriended or unknown personal accounts or by unfamiliar news outlets (that they do not follow), they used up to all three of the identified strategies.

5. Discussion

When comparing the findings of this study, namely the strategies participants used to evaluate whether information on an OSN is true or false, to findings from other contexts, such as how people generally evaluate information or how people evaluate the information in media content or on websites, the results show a large overlap. The strategy *searching for more* can be interpreted as an umbrella strategy for the four external strategies found by Tandoc et al. (2018): incidental, intentional, interpersonal, or institutional acts of authentication. Participants explicitly referred to an intentional search for more information, likely because participants followed the study design and tried to evaluate the information. Participants also mentioned incidental external authentication when they said they would only try to search for more information if the topic was interesting to them. Otherwise, they would just wait to get more information on the respective topic incidentally. This is in line with propositions made in dual-processing models, as interest in the topic (i.e., motivation) leads individuals to use more effortful strategies (e.g., Chaiken, 1980; Petty & Cacioppo, 1986). When wanting to search for more information, participants said they would google it, search for it on YouTube, go to the website of a newspaper (institutional authentication), or they would discuss it with friends or family and see what they considered to be correct or wrong about the topic (interpersonal authentication). This *search for more* strategy can also be interpreted as delaying “the arrival at a conclusion” (Kunda, 1990, p. 482), meaning that participants do not make an immediate judgement of the credibility of information but want to wait for more evidence. Hence, it goes in line with literature on motivated reasoning, which suggests that the motivation to be accurate drives a more elaborate search for information (e.g., Nir, 2011).

The other strategies rather refer to internal acts of authentication. *Knowledge of content carries the most weight* goes in line with the strategy of relying on oneself (Tandoc et al., 2018) and maintaining compatibility with one’s own beliefs (Lewandowsky et al., 2012) as well as the heuristics of self-confirmation and consistency (Metzger & Flanagin, 2013). The information in a post should be con-

sistent and compatible with what a user already knows. Regarding relying on oneself, this also means that the participants judged the information using their intuition, which, in turn, is based on knowledge and experience.

Knowledge of account carries the most weight goes in line with source credibility (Lewandowsky et al., 2012; Metzger & Flanagin, 2013; Tandoc et al., 2018) and source experience (Lucassen & Schraagen, 2011). Whereas the different definitions of source credibility focus on which indicators make a source credible (e.g., trustworthiness, expertise, believability, truthfulness; Metzger & Flanagin, 2015; Roy, Huh, Pfeuffer, & Srivastava, 2017), knowledge about the account is to a large extent based on the individual's experience with the account. This can be based on previous posts or on the personal relationship that the individual might have to the account holder outside the OSN. When using this strategy, users sometimes paid attention to whether the post is compatible with what the account usually posts (otherwise the post violates their expectation of the account). Hence, the heuristic expectancy violation fits to this strategy as well (Metzger & Flanagin, 2013). Further, the current study supports the finding by Lucassen and Schraagen (2011) that the experience with the source can lead people to disregard any other characteristic of the message.

The *every detail needs to fit* strategy encompasses diverse heuristics as well as the strategies coherence and popularity (Lewandowsky et al., 2012; Tandoc et al., 2018). When employing this strategy, users may consider everything that is not solely based on knowledge. This can include persuasive intent, coherence of a message, and characteristics of the post such as spelling and grammar or neutrality of images (Horstmann et al., 2018; Metzger & Flanagin, 2013). Additionally, characteristics such as plausibility, which also refers to one's own knowledge, can also be considered within this strategy, as everything has to fit.

When the post was from a news outlet that participants followed, *knowledge of account carries the most weight* was the most-used strategy; this indicates that participants relied on these accounts. This goes in line with selective exposure, as users actively chose to follow these news outlets (Djerf-Pierre & Shehata, 2017). It may be that they follow a news outlet because they expect that outlet to post true news. In this case, the use of the *knowledge of account carries the most weight* strategy can be interpreted as heuristic or peripheral processing in the sense of dual-processing models. Of course, this does not apply if users follow a news outlet's account because they find it entertaining and do not take its content seriously or because they want to get insight into what an untrustworthy news outlet posts. However, it seems that for the news outlets discussed within this study, this was not the case. Instead, participants may have based their decision to rely on the news outlet because the journalists who post on them are assumed to be professionals in evaluating information and writing news items.

6. Limitations

Several limitations of the study design should be considered when interpreting the results. One concerns ecological validity. As participants were asked to think aloud, they were a "captive audience" (Hovland & Weiss, 1951). This may make their information processing and evaluation far more systematic than it would be in a real situation, where they scroll through their Facebook newsfeed and spend only a few

seconds on most posts (Vraga et al., 2016). The setting of the study, however, aimed at giving participants enough time to mention everything they thought about. Thus, they might have looked at and mentioned more post characteristics than they usually would. This may explain why for many posts, participants used several strategies. However, the aim of this study was to find all the strategies that users employed consciously in evaluating posts for true or false content. Thus, giving the participants enough time, which might lead to an over-reporting of the characteristics they considered, was important to follow this study's aim. But, users may also employ strategies unconsciously (Fiske, Kinder, & Larter, 1983); those, of course, cannot be voiced by the participants in a think-aloud study.

Another limitation may be socially desirable answers. As found in previous research, participants of studies tend to underreport that they simply give lower ratings of credibility to counter-attitudinal information to preserve their own attitude; instead, they tend to report that they use arguments against the content (Zuwerink Jacks & Cameron, 2003). A similar social desirability bias might be present in the current think-aloud study: Participants might have described what they thought they should be thinking about the posts in question (especially since the researcher was present in the room). Moreover, they might have mentioned those characteristics that they thought should be relevant to consider and not only those that they actually considered. To try to minimize this possible bias, during the interview the interviewer constantly reassured participants that what they were saying was correct by nodding or saying something supportive ("yes," "right," etc.).

The aim of this qualitative study was to investigate which strategies OSN users employ in evaluating the information in posts. Thus, the design was focused on posts that actually appear in users' newsfeeds. This way, participants were likely to have a relationship with the posting account, which was important for the strategy that used knowledge of the account. However, with a sample of 15 participants, this design did not allow for quantitatively testing whether these strategies helped users successfully detect false information in OSN. This is an important point, because if these strategies do not help users detecting false information sufficiently, educators of media literacy might want to consider ways of improving the OSN users' media literacy that is practical for them to employ. Users will always need lower-effort strategies, as they do not have an endless capacity for evaluating all the information they get through OSN. Thus, future research could build on the results of this study and test whether these strategies help users to detect false information successfully.

7. Conclusion

Keeping these limitations in mind, what are the implications of each strategy for the spread of false information? As the strategy *searching for more* showed, OSN users are willing to search for further information when they cannot determine whether a post contains false information, which can be seen as a positive sign for reducing the spread of false information. However, as this is impossible to do for every post, and users rarely verify information they find online (Metzger, 2007), users will likely only employ this accuracy-driven, high-effort strategy if the topic is interesting to them. Nonetheless, this strategy is not likely to lead to

further distribution of false information, because users will not have strong reasons to share or react to content that they do not find interesting, as long as they are not aiming to support someone (e.g., good friend, family member).

A strategy that might be problematic for the spread of false information is *knowledge of account carries the most weight* if it overrules the content. For example, if a highly trustworthy account makes a mistake, the trusting user may rely on the trustworthy account and not question the content. It is then problematic if the user reacts to or shares the post.

This is different for the strategy of when *knowledge of content carries the most weight*. If the user's knowledge is correct, there are two options: Either the user does not react to the post, which would have no effect on the distribution of false information, or, if the user reacts to a post and refutes it, this may reduce the distribution of false information. However, users' overconfidence about the topical knowledge might be a problem if the knowledge is false.

With the third strategy, *every detail needs to fit*, the implications are not that clear and completely differ from situation to situation. If the post is inconsistent and false, this strategy helps users to detect false information; then the user's actions determine the effect on the spread of false information. If users do not react to the post there is no effect on the spread of false information, and if they refute the post this action helps to reduce the spread of false information. If the post is consistent and true, then the user recognizes that the information is true. If the post is inconsistent but true, users may think the true content is false information, which should also have no effect on the distribution of false information. However, if the post is consistent but false, this may cause problems: In this case, users employing this strategy may think the false information is true and then react to it. This could increase the distribution of false information.

One main finding of this study, that users evaluate information from an "in-network" news outlet (a news outlet that they follow) by heavily relying on the account and their experiences with previous posts of the account, shows that for future research on news posts in OSN, researchers must consider whether the participating OSN users follow the news outlet or not.

References

- Ajzen, I. (1996). The social psychology of decision making. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 297–325). New York: Guilford Press.
- Bode, L., & Vraga, E. K. (2015). In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication*, 65(4), 619–638. <https://doi.org/10.1111/jcom.12166>
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology*, 39(5), 752–766. <https://doi.org/10.1037//0022-3514.39.5.752>
- Chan, M.-P. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, 28(11), 1531–1546. <https://doi.org/10.1177/0956797617714579>

- Charters, E. (2003). The use of think-aloud methods in qualitative research: An introduction to think-aloud methods. *Brock Education*, 12(2), 68–82.
- Cheng, X., Fu, S., & de Vreede, G. J. (2017). Understanding trust influencing factors in social media communication: A qualitative study. *International Journal of Information Management*, 37(2), 25–35. <https://doi.org/10.1016/j.ijinfomgt.2016.11.009>
- Chung, C. J., Nam, Y., & Stefanone, M. A. (2012). Exploring online news credibility: The relative influence of traditional and technological factors. *Journal of Computer-Mediated Communication*, 17(2), 171–186. <https://doi.org/10.1111/j.1083-6101.2011.01565.x>
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., . . . Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences of the United States of America*, 113(3), 554–559. <https://doi.org/10.1073/pnas.1517441113>
- Djerf-Pierre, M., & Shehata, A. (2017). Still an agenda setter: Traditional news media and public opinion during the transition from low to high choice media environments. *Journal of Communication*, 67(5), 733–757. <https://doi.org/10.1111/jcom.12327>
- Ericsson, K. A. (1987). 6. Theoretical implications from protocol analysis on testing and measurement. In R. R. Ronning, J. A. Glover, J. C. Conoley, & J. C. Witt (Eds.), *The influence of cognitive psychology on testing* (pp. 191–226). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (2nd ed.). Cambridge, MA: MIT Press.
- Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity*, 5(3), 178–186.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.
- Fiske, S. T., Kinder, D. R., & Larter, W. M. (1983). The novice and the expert: Knowledge-based strategies in political cognition. *Journal of Experimental Social Psychology*, 19, 381–400.
- Hölig, S., & Hasebrink, U. (2018). *Reuters Institute Digital News Report 2018: Ergebnisse für Deutschland. Arbeitspapiere des Hans-Bredow-Instituts: Nr. 44* [Reuters Institute Digital News Report 2018: results for Germany. Working papers of the Hans-Bredow-Institute: no. 44]. Hamburg, Germany: Hans-Bredow-Institut für Medienforschung.
- Horstmann, A. C., Rösner, L., Conrad, L., & Heidemann, R. (2018, September). *Fake News or Real Truth?! Ergebnisse einer Think Aloud Befragung zur Erkennung von Falschnachrichten* [Fake news or real truth?! Results of a think aloud study on the detection of false news]. Paper presented at the 51st conference of the DGPs, Frankfurt am Main, Germany.
- Hovland, C., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly*, 15, 635–650. <https://doi.org/10.1086/266350>
- Kirkpatrick, L. A., & Epstein, S. (1992). Cognitive-experiential self-theory and subjective probability: Further evidence for two conceptual systems. *Journal of Personality and Social Psychology*, 63(4), 534–544.
- Kuklinski, J. H., Quirk, P. J., Jerit, J., Schwieder, D., & Rich, R. F. (2000). Misinformation and the currency of democratic citizenship. *The Journal of Politics*, 62(3), 790–816.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480–498. <https://doi.org/10.1037/0033-2909.108.3.480>
- Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., . . . Zittrain, J. L. (2018). The science of fake news: Addressing fake news requires a multidisciplinary effort. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.aao2998>

- Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest: a Journal of the American Psychological Society*, 13(3), 106–131. <https://doi.org/10.1177/1529100612451018>
- Livingstone, S. (2004). Media literacy and the challenge of new information and communication technologies. *The Communication Review*, 7(1), 3–14. <https://doi.org/10.1080/10714420490280152>
- Lucassen, T., & Schraagen, J. M. (2011). Factual accuracy and trust in information: The role of expertise. *Journal of the American Society for Information Science and Technology*, 62(7), 1232–1242. <https://doi.org/10.1002/asi.21545>
- Metzger, M. J. (2007). Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078–2091. <https://doi.org/10.1002/asi.20672>
- Metzger, M. J., & Flanagin, A. J. (2013). Credibility and trust of information in online environments: The use of cognitive heuristics. *Journal of Pragmatics*, 59, 210–220. <https://doi.org/10.1016/j.pragma.2013.07.012>
- Metzger, M. J., & Flanagin, A. J. (2015). Psychological approaches to credibility assessment online. In S. S. Sundar (Ed.), *The handbook of the psychology of communication technology* (Vol. 32, pp. 445–466). Chichester, UK: Wiley-Blackwell. <https://doi.org/10.1002/9781118426456.ch20>
- Metzger, M. J., Flanagin, A. J., Markov, A., Grossman, R., & Bulger, M. (2015). Believing the unbelievable: Understanding young people's information literacy beliefs and practices in the United States. *Journal of Children and Media*, 9(3), 325–348. <https://doi.org/10.1080/17482798.2015.1056817>
- Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60(3), 413–439. <https://doi.org/10.1111/j.1460-2466.2010.01488.x>
- Neubaum, G., & Krämer, N. C. (2017). Opinion climates in social media: Blending mass and interpersonal communication. *Human Communication Research*, 43(4), 464–476. <https://doi.org/10.1111/hcre.12118>
- Neubaum, G., & Krämer, N. C. (2018). What do we fear? Expected sanctions for expressing minority opinions in offline and online communication. *Communication Research*, 45(2), 139–164. <https://doi.org/10.1177/0093650215623837>
- Newman, N., & Fletcher, R. (2017). *Bias, bullshit and lies. Audience perception on low trust in the media*. Retrieved from <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/inline-files/Nic%20Newman%20and%20Richard%20Fletcher%20-%20Bias%2C%20Bullshit%20and%20Lies%20-%20Report.pdf>
- Nir, L. (2011). Motivated reasoning and public opinion perception. *Public Opinion Quarterly*, 75(3), 504–532. <https://doi.org/10.1093/poq/nfq076>
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, 32(2), 303–330. <https://doi.org/10.1007/s11109-010-9112-2>
- Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences of the United States of America*, 116(7), 2521–2526. <https://doi.org/10.1073/pnas.1806781116>

- Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of persuasion. *Advances in Experimental Social Psychology*, 19, 123–205. [https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)
- Roy, A., Huh, J., Pfeuffer, A., & Srivastava, J. (2017). Development of trust scores in social media (TSM). Algorithm and application to advertising practice and research. *Journal of Advertising*, 46(2), 269–282. <https://doi.org/10.1080/00913367.2017.1297272>
- Shin, J., & Thorson, K. (2017). Partisan selective sharing: The biased diffusion of fact-checking messages on social media. *Journal of Communication*, 67(2), 233–255. <https://doi.org/10.1111/jcom.12284>
- Sundar, S. S. (2008). The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility. In M. J. Metzger & A. J. Flanagin (Eds.), *Digital Media, Youth, and Credibility. The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning* (pp. 73–100). Cambridge, MA: The MIT Press. <https://doi.org/10.1162/dmal.9780262562324.073>
- Tandoc, E. C., Ling, R., Westlund, O., Duffy, A., Goh, D., & Zheng Wei, L. (2018). Audiences' acts of authentication in the age of fake news: A conceptual framework. *New Media & Society*, 20(8), 2745–2763. <https://doi.org/10.1177/1461444817731756>
- Thorson, E. (2016). Belief echoes: The persistent effects of corrected misinformation. *Political Communication*, 33(3), 460–480. <https://doi.org/10.1080/10584609.2015.1102187>
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146–1151. <https://doi.org/10.1126/science.aap9559>
- Vraga, E. K., Bode, L., & Troller-Renfree, S. (2016). Beyond self-reports: Using eye tracking to measure topic and style differences in attention to social media content. *Communication Methods and Measures*, 10(2–3), 149–164. <https://doi.org/10.1080/19312458.2016.1150443>
- Vraga, E. K., & Tully, M. (2019). News literacy, social media behaviors, and skepticism toward information on social media. *Information, Communication & Society*. Advance online publication. <https://doi.org/10.1080/1369118X.2019.1637445>
- Walter, N., & Murphy, S. T. (2018). How to unring the bell: A meta-analytic approach to correction of misinformation. *Communication Monographs*, 85(3), 423–441. <https://doi.org/10.1080/03637751.2018.1467564>
- Walter, N., & Tukachinsky, R. (2019). A meta-analytic examination of the continued influence of misinformation in the face of correction: How powerful is it, why does it happen, and how to stop it? *Communication Research*. Advance online publication. <https://doi.org/10.1177/0093650219854600>
- Walther, J. B., & Jang, J.-w. (2012). Communication processes in participatory websites. *Journal of Computer-Mediated Communication*, 18(1), 2–15. <https://doi.org/10.1111/j.1083-6101.2012.01592.x>
- Walther, J. B., & Valkenburg, P. M. (2017). Merging mass and interpersonal communication via interactive communication technology: A symposium. *Human Communication Research*, 43(4), 415–423. <https://doi.org/10.1111/hcre.12120>
- Zuwerink Jacks, J., & Cameron, K. A. (2003). Strategies for resisting persuasion. *Basic and Applied Social Psychology*, 25(2), 145–161. https://doi.org/10.1207/S15324834BASP2502_5