6-18-2022

Visual Attention to Fake News Flags in Social Media News Posts: An Eye Tracking Study

Samuel Kießling
University of Innsbruck, samuel.kiessling@uibk.ac.at

Frederik M. Wiedmann
University of Innsbruck, frederik.wiedmann@uibk.ac.at

Kathrin Figl
University of Innsbruck, kathrin.figl@uibk.ac.at

VISUAL ATTENTION TO FAKE NEWS FLAGS IN SOCIAL MEDIA NEWS POSTS: AN EYE TRACKING STUDY

Research Paper

Samuel Kießling, University of Innsbruck, Innsbruck, Austria, samuel.kiessling@uibk.ac.at

Frederik Wiedmann, University of Innsbruck, Innsbruck, Austria, frederik.wiedmann@uibk.ac.at

Kathrin Figl, University of Innsbruck, Innsbruck, Austria, kathrin.figl@uibk.ac.at

Abstract

Given the widespread prevalence of fake news on social media, fake news warnings can play a decisive role in combating misinformation. However, research is still debating the extent to which readers of news on social media heed fake news warnings, which is important to evaluate their effectiveness. In this work, we focus on fake news flags with color gradients from green (verification) to red (warning) and investigate conditions under which they receive visual attention. In an eye tracking experiment, we assigned fake news flags to three social media post elements (user, source, news article) and manipulated the number of fake news flags that indicate a warning or verification. Our results reveal that fake news flags for the news article receive more visual attention than those for the user or source. In addition, we provide evidence that confirmation bias moderates the effect of unique flags (warning or verification) on visual attention.

Keywords: Fake News Flag, Eye Tracking, Visual Attention, Confirmation Bias.

1 Introduction

Fake news is especially widespread in the context of social media, which functions as a "new distribution mechanism for news" (Duffy, 2020, p. 51), where the implications are deemed to be particularly harmful, for example, impacting public health decisions (Islam et al., 2020; Scherer and Pennycook, 2020) or shifting political attitudes and outcomes (Murray et al., 2020; Zhuravskaya et al., 2020). Compared to traditional news media, social media platforms enable their users to share information without latency, and therefore accelerate the distribution of any content, including fake news (Zubiaga et al., 2018). Journalism's traditional role as gatekeeper has been shifted to friends and algorithms determining what news shows up in the social media feed (Duffy, 2020). This user-generated content on social media poses verification concerns, especially for Facebook, which has been accused of accelerating the spread of disinformation on political issues (Allcott and Gentzkow, 2017). Still, recent research misses a thorough investigation of protective measures to mitigate threats that could affect democracy itself and highlights that even a short exposure to fake news could unconsciously alter a person's behavior (Bastick, 2021).

Social media platforms have introduced warnings in the form of fake news flags. Although fake news warnings on social media posts have been under close research attention for a few years now (e.g., Allcott and Gentzkow, 2017; Moravec et al., 2020; Kim and Dennis, 2019; Pennycook et al., 2020a), there are still contradictory research findings regarding their effectiveness. On the one hand, there are studies that underpin the understanding that the display of explicit warnings reduces the believability of the flagged content and thus limits the spread of fake news (Lewandowsky et al., 2012; Figl et al., 2019; Moravec et al., 2020). On the other hand, a few empirical findings challenge this direct effect of fake news flags on believability (Ross et al., 2018). In the case of politically charged topics, even an opposite effect can occur, a kind of defiance and “backfire” effect that causes fake news flags to have a positive impact on believability when people are strongly convinced of their views (Flynn et al., 2017; Kim et al., 2019b). In addition, there is research that falls between the two poles in its view and has failed to find any significant mitigation of believability, as other effects, like confirmation bias, dominate in terms of their potential influence on believability (Kim et al., 2019b). Research has also addressed different designs or structures of the warnings as potential moderators of the effectiveness of fake news flags to lower the believability of fake news content (Moravec et al., 2020; Figl et al., 2019). Here, empirical studies revealed that there might be a need to attach not only warnings, but also verifications, since the absence of a warning “will have the unintended side effect of causing untagged headlines to be viewed as more accurate” (Pennycook et al., 2020a, p. 2).

Until now, the focus of research has been mainly on the direct effects that can emerge from different design elements of fake news flags. However, an aspect of fake news flag design that has been largely neglected so far, with only few exceptions, is whether readers of social media news posts actually attend the displayed warnings or verifications (Kim et al., 2021). If readers were not sufficiently considering fake news flags, they would not be reached by the warnings or verifications (Kim et al., 2021), which could explain ambivalent findings on the effectiveness fake news flags. To make readers’ visual attention to fake news flags accessible, we must apply neurophysiological measures like eye tracking. Currently, there remains a shortage of research that investigates fake news flag designs using modern neuropsychological methods. Electroencephalography (EEG) as a neuropsychological method has been applied to fake news flag research (Moravec et al., 2019) to measure cognitive activity evoked by fake news flags, but is not able to capture visual attention. Only eye tracking is able to shed light on the black box of how fake news and warnings are perceptually and cognitively processed (Rayner, 2009). A number of eye tracking studies on fake news in social media posts have been conducted so far, e.g., on fake news on different consumer brands (Ladeira et al., 2021), heath-related misinformation (Chou et al., 2020; Trivedi et al., 2020; Trivedi et al., 2021), new scientific results (Tanaka et al., 2019) and various other fake news topics (Simko et al., 2019; Simko et al., 2021), but only a few have considered any kind of fake news correction or warning (e.g. Kim et al., 2021; Tanaka et al., 2019; Shi et al., 2022). For example, Simko et al. (2021) investigated visual attention by letting study participants scroll through Facebook posts and let them judge the veracity of the posts, but they had not included any kind of fake news flags.

Eye tracking as a method to collect neurophysiological data opens up new research aspects of fake news flags that have remained unanswered so far, and concern readers’ visual attention to fake news flags and, therefore, fake news flag effectiveness. First, readers of social media news posts are confronted with content- and source-related cues to evaluate the believability of a news article (Reinhard and Sporer, 2010). Beyond the content-related news article, social media news posts typically include source-related information like the user who shared the news article or the original source of the article in a narrow sense (i.e., the news provider or author of the news article). Since readers can also infer the believability of a news article based on source-related social media news post elements, research has experimented with attempts to attach fake news flags to the source of the post (Kim et al., 2019b, 2019a; Kim and Dennis, 2019). However, we do not yet know whether readers pay more visual attention to fake news flags for source-related social media news post elements or to content-related social media news post elements, although attracting attention is considered fundamental for an effective warning (Laughery and Wogalter, 2006). Therefore, one goal of our study is to investigate differences in the visual attention to fake news flags that are either attached to content- or source-related social media post elements. Attaching fake news flags for source-related elements also means that readers are confronted by more than one fake news flag per social media news post, which could lead to contradicting fake news displays. For instance, a user who shared the article might be flagged with a warning, but the news article itself could be classified as credible. Therefore, the second goal of our study is to shed more light on the conditions under which fake news flags receive visual attention when a social media news post

is accompanied by multiple fake news flags due to the flagging of several elements. Given these two goals, the following research questions arise:
To what extent do readers visually attend fake news flags for different (content- and source-related) elements of a social media news post, and which conditions influence their visual attention to these flags?

To investigate this research question, we conducted an eye tracking experiment with artificially created Facebook news posts. We attached fake news scales to display either warnings or verifications regarding the user who shared the news article, the source of the news article, and the news article itself, to observe how users deal with similar or conflicting information cues. This provides the opportunity to directly compare the effect of fake news flags attached to individual social media news post elements and their inherent information cues, and to uncover users' visual attention to these flags. Visual attention to fake news flags is a prerequisite for influencing users' believability judgment of news posts.

In the following, we will first discuss pioneering steps and results in fake news flag research relevant for our study, from which we can then derive our hypotheses. Subsequently, we describe our eye tracking study design and then proceed to the results of the study. Finally, we will discuss the results and derive possible implications as well as limitations.

2 Background

A central point of fake news research revolves around the effectiveness of fake news flags. A major focus of prior research has been on the effect of flags on the believability of news articles, as a decisive antecedent for user behavior, such as sharing posts (Pennycook and Rand, 2021; Pennycook et al., 2020a; Moravec et al., 2019). However, past research overlooked visual attention to these fake news flags as an important intermediate step, which is potentially required for the fake news flag to exert any effect. Yet, other studies that relied on eye tracking considered visual attention to different social media post elements, but ignored distinct fake news flags attached to specific social media post elements (e.g., Simko et al., 2021; Trivedi et al., 2020; Kim et al., 2021; Tanaka et al., 2019).

As the research on fake news has progressed, it has become apparent that it is challenging to accurately differentiate between true and fake content on social media, which is why there is an entire field of research on the detection of fake news (e.g., Ozbay and Alatas, 2020; Shu et al., 2019; Shu et al., 2017; Batailler et al., 2021; Bunde et al., 2022). Warning messages are deemed to change people's perceptions and thus influence subsequent behavior (Moravec et al., 2018). A basic precondition for the effectiveness of fake news flags, which we will address, is that they receive attention from users. Focusing on visual attention is relevant for the context of fake news flags, because "an effective warning must initially attract attention" (Wogalter, 2005, p. 28) to urge users to process the information (Laughery and Wogalter, 2006).

2.1 Content- versus source-related fake news flags

Warnings could be used to draw attention to the content itself but also to social media users, as just 27% of news consumers access news content directly, with the majority accessing news via "side door" points of entry (Newman et al., 2021). The most prominent side door point of entry in 2021 was social media, which accounted for 26% of traffic across all age groups and as much as 34% of coveted traffic among those under 35 (Newman et al., 2021). Additionally, compared to humans, bots tend to play only a minor role in spreading fake news, because humans are motivated to share information that sounds unfamiliar and new, as it is often the case with fake news, and therefore gain social status (Vosoughi et al., 2018). Previous research has uncovered a variety of reasons why users share news content (e.g., Haug and Gewald, 2018; Heimbach and Hinz, 2018), even though in many cases, users may not bother to look at the article at all (Gabielkov et al., 2016).

A news post on Facebook can, thus, consist of three parts, the content (article and image), the source, and the user who shares the article. While prior studies have investigated the content and source as elements for attaching a fake news warning or a trust scale (e.g., Moravec et al., 2020; Kim and Dennis, 2019), the role of fake news flags for users who post (news) articles, has received less attention for evaluating believability yet. However, such user flags based on past posts and displayed next to the user would also be important for readers to help them determine the believability of a user, particularly if a user often shares false content.

We can draw on previous research findings on how fake news warnings for various post elements might be perceived by readers. On the one hand, results indicate that verifying the source of a news article and, consequently, displaying a warning can lead to a news article being believed less (Di Domenico et al., 2021; Kim et al., 2019a). On the other hand, this effect is estimated to be lower compared to article-related fake news warnings (Kim et al., 2019a). It, therefore, remains unresolved what effect a fake news flag of the user might have. Following research on processing information cues to judge believability more generally, Reinhard and Sporer (2010, p. 94) define “content information as central information and source cue information as peripheral information.” Making this distinction between central and peripheral information becomes crucial when monitoring visual attention in situations where there are multiple cues that have a similar appearance. Peripheral cues are received in a reflexive manner, whereas central cues are perceived consciously (Riggio and Kirsner, 1997).

Another factor that could have an influence on the evaluation of different information cues is the so-called cognitive authority (Bernier, 1984) that can be assigned to the source or the user. People attribute a certain amount of authority to other people or institutions depending on how capable they think they are (Rieh, 2002). In this case, a source might be regarded more as an expert who understands how to prepare content conscientiously, and the user more as a layperson who has only limited knowledge for evaluating content. If we apply this rationale further to assessing fake news flags, the result is a priority list for judging believability in a descending order, with article-related information (fake news flags) being considered first, followed by the source and the user. A higher priority should be reflected in higher visual attention; hence we hypothesize:

H1. Fake news flags that are attached to the news article receive more visual attention than fake news flags that are assigned to the source (H1a) or the user (H1b).

2.2 Warnings versus verifications

The different social media post elements (article, source, user) can be highlighted and evaluated for readers using warnings or verifications. Many different fake news flag designs have been researched in previous studies. One of the more obvious flags is a caution triangle embedded in a red frame attached to the article of a fake news post. This design was used by Facebook itself and was therefore examined in most of the first studies (e.g., Pennycook et al., 2020a; Pennycook et al., 2018; Ross et al., 2018; Moravec et al., 2020). Yet, research has considered many other fake news flag designs, such as a modified caution symbol (Kim et al., 2019b), a stop sign (Moravec et al., 2020), or the word “false” in red letters placed across the article image (Pennycook et al., 2020a). Such research has demonstrated that design plays a role in warning effectiveness (Moravec et al., 2020).

Gradually, researchers also recognized the need for verifications to mitigate potential effects, like the implied truth effect (Pennycook et al., 2020a), where news articles without any fake news flag are believed more when others have flags attached to them. That is why, for example, Zubiaga et al. (2018) developed a rumor classification system that includes a veracity component and thus does not exclusively consider the debunking of fake information. Another approach is to remind users, prior to their exposure to potentially false content, to pay attention to the accuracy of the information, so they can better identify fake news (Pennycook et al., 2020b). Additionally, research streams on algorithmic support in fake news detection advanced (Ozbay and Alatas, 2020; Shu et al., 2019; Janze and Risius, 2017), which could also be used to verify certain information instead of only highlighting fake content with warnings (Jin et al., 2016; Rubin et al., 2015). As verifications, for example, green symbols with a confirmation checkmark (Kim et al., 2019b) or the word “true” in green letters across the article image (Pennycook et al., 2020a) have been investigated in previous research studies.

Warnings alone are not sufficient when considering possible mitigation strategies against fake news. In fact, fake news research has documented the illusory truth effect (Henderson et al., 2021), whereby repetitive information is considered true after a certain period of time because it can be processed more fluently, even though this information was received in the context of a warning (Skurnik et al., 2005). Via priming for accuracy, it is possible to prevent the illusory truth effect (Brashier et al., 2020). Additionally, warnings combined with control questions to understand the warnings displayed, can mitigate the illusory truth effect, though not eradicate it entirely (Nadarevic and Aßfalg, 2017). Verifications could have a similar effect since different stimuli are displayed (warnings and verifications) and could prevent the illusory truth effect.

Yet, few studies have investigated the effects of both using verifications and warnings for posts on a social media platform, with the study of Pennycook et al. (2020a) being a rare exception, which has, however, not looked at visual attention. We can draw on diverse psychological theories to hypothesize how users might distribute their attention to verifications and warnings. For instance, it is known that the visual search for threatening stimuli (e.g., snakes) is faster than for neutral stimuli (e.g., flowers) and that threatening stimuli attract more visual attention based on evolutionary accounts (Öhman et al., 2001). Because associations between visual threats and warning symbols are quickly learned, fake news flags may be better at drawing visual attention to a potential threat, i.e., fake news posts, than to trustworthy news posts (Cave and Batty, 2006). Even if fake news is not considered a potentially dangerous threat, negative information has been shown to generally attract more attention than positive information (Smith et al., 2003). Therefore, we assume that readers of posts on social media are likely to pay more attention to warnings than to verifications:

H2. Fake news flags that indicate a warning receive more visual attention compared to those flags that indicate a verification.

2.3 Visual attention to fake news flags

Addressing visual attention as a research focus necessarily involves considering visual information processing and theories included therein. Usually, whenever fake news flags are assigned to social media posts or their source to highlight believability, they are often displayed using a particular color palette (e.g., Kim and Dennis, 2019) or shapes that are associated with attention (e.g., Pennycook et al., 2020a). In the following, we want to derive hypotheses for a situation in which not only one, but several warnings or verifications for elements of a social media news post would be available for a reader, which might even contradict each other, and should therefore be visually different.

In this context, the feature integration theory provides a substantial basis for hypothesizing how attention might be distributed in the presence of unique fake news displays, indicating either a warning or verification. The feature integration theory (Treisman and Gelade, 1980) is a prominent theory of human visual attention, and it distinguishes between two stages, namely preattentive and focused attention, to explain recognizing objects through visual attention. The former is fast, automatic, and effortless. During this stage, the perceptual process analyzes object details (or features), such as shape, color, orientation, and movement, which frame different dimensions (Wolfe and Horowitz, 2017). Attention is decisively biased by the dissimilarity between the target feature and the deviant feature, and by the homogeneity of the deviant features among themselves (Koehler et al., 2014). The former should therefore be particularly strong, while the latter particularly low.

Applying this rationale to the context of using several fake news displays, a warning or verification would visually pop out (preattentive attention) as soon as its color, shape, or orientation is unique, meaning all differing displays would indicate a different – but identical in comparison to each other – visual feature. In the case of color scales, as, for example, used by Kim and Dennis (2019), even the orientation of the indicator signaling the color would differ for the unique scale and therefore be visually highlighted. Moody (2009, p. 762) has called this criterion perceptual discriminability, which refers to the “ease and accuracy with which graphical symbols can be differentiated from each other.”

A difference compared to other close objects, according to the feature integration theory, would immediately receive more attention, even though it might be a cue with lower importance, because two other fake news flags indicate the opposite. Thus, we hypothesize:

H3. Fake news flags that uniquely indicate a warning or verification receive more visual attention compared to those that share their fake news indication with other fake news flags.

2.4 The effect of cognitive dissonance on visual attention

When fake news is consumed on social media, it rarely occurs objectively but rather with the reader's opinion as a filter (Figl et al., 2019; Knobloch-Westerwick and Lavis, 2017). In this case, the attitude that already exists before reading a news article can potentially influence how additional information about the news article, such as a fake news warning, is perceived and processed, even to the extent that warnings are completely ignored (Moravec et al., 2019). Confirmation bias – a human bias to prefer information that matches our own assumptions and block out contrary narratives (Nickerson, 1998) – could thus affect how we perceive fake news warnings or verifications. A similar effect has already been observed in other research fields, for example, in the perception of online reviews (Yin et al., 2016) or demand-based scarcity booking decisions (Kim et al., 2020).

Further, cognitive dissonance which can arise in situations when contradictory information has to be considered is often combined with conformation bias (Izuma et al., 2010). People experience distress when they are presented with information that contradicts their own beliefs, which can lead to the devaluation of contradictory evidence in order to restore their own equilibrium (Festinger, 1957). Kim et al. (2019b) previously concluded that confirmation bias is a dominant factor when considering the effectiveness of fake news warnings and linked this finding to the intuitive and hedonistic mindset that governs social media use. Hence, we assume that cognitive dissonance should translate to eye tracking measures (e.g., Guazzini et al., 2015), because to resolve the dissonance, it may be necessary to adopt a more analytical and reflective mindset (Moravec et al., 2020), which might lead to longer processing time than a quick and intuitive judgment (Bago and De Neys, 2017). We, therefore, assume that, due to cognitive dissonance, subjects' visual attention is longer for fake news displays that are inconsistent with their own views. That means if readers agree with an article, they will pay more attention to warnings; however, if readers disagree with the content, they will pay more attention to verifications. Thus, we hypothesize:

H4a. Confirmation bias moderates the relationship between the fake news indication and visual attention in so far that under a higher level of confirmation bias, a fake news flag indicating a warning receives more visual attention compared to a flag that indicates a verification.

H4b. Confirmation bias moderates the relationship between the uniqueness of the fake news indication and visual attention in so far that under a higher level of confirmation bias, a fake news flag that uniquely indicates a warning or verification receives more visual attention compared to a flag that shares its indication with another flag.

3 Method

3.1 Research design

In our study, we set up an eye tracking experiment with 12 social media news posts. Each post contained fake news flags that indicated the reliability of the user, source, and news article. The fake news flags were designed to indicate either a warning or a verification. Our manipulation included a between-subjects factor (uniqueness of warning vs. uniqueness of verification) which reflected the number of fake news flags on the social media news post showing warnings relative to verifications. In addition, we created three experimental (sub)groups with variations in the distribution of warnings and verifications among the three scales so that each possible constellation of verification or warning was covered and counterbalanced. In total, the manipulation of the fake news flags' uniqueness times the three variations led to six different experimental subgroups. *Table* 1 illustrates which fake news flags the six experimental subgroups would present for the first three social media news posts; the pattern continues for the rest of the posts.

Experimental Treatment	**Fake News Flag**	**Variation 1**			**Variation 2**			**Variation 3**		
		Post 1	Post 2	Post 3	Post 1	Post 2	Post 3	Post 1	Post 2	Post 3
1 (1 Warning, 2 Verifications)	**User**	V	V	W	V	W	V	W	V	V
	Source	V	W	V	W	V	V	V	V	W
	Article	W	V	V	V	V	W	V	W	V
2 (1 Verification, 2 Warnings)	**User**	W	W	V	W	V	W	V	W	W
	Source	W	V	W	V	W	W	W	W	V
	Article	V	W	W	W	W	V	W	V	W

Table 1. Experimental treatments and variations of green verifications (V) and red warnings (W).

3.2 Social media news posts and fake news flags

The structure of the social media news posts resembled Facebook's design for shared articles in the news feed (Figure 1). In doing so, the chosen design is intended to be a proxy for social media posts in general and was chosen as Facebook is still widespreadly used for news consumption (Newman et al., 2021) and due to the crucial role it has played over the past few years in spreading fake news on political issues (e.g., Allcott and Gentzkow, 2017). The post included the account picture and name of the user who shared the article, a picture, and name of the original source of the news article, and the shared news article itself. Each news article consisted of a headline, a picture, and a short sentence summarizing the article. We decided to make up the content of all news articles to ensure similar conditions among the social media news post. Hence, two authors of this paper and one volunteering student developed and refined the 12 news articles in several rounds. All news articles were inspired by real public debates from the years 2019 to 2020. However, we followed Pennycook et al. (2021) and opted for news articles that were emotionally charged but where their relevance was not too short for the time frame of the study. Below each social media news post were three buttons that indicated the option to like, share, or comment on the social media news post. It was possible to click multiple buttons. If participants decided to comment on the social media news post, a text field popped up below the buttons. Additionally, we designed colored fake news scales as fake news flags, which were either indicating green for verifications or red for warnings with a yellow color gradient in between. To keep the number of possible values of fake news scales on a manageable level, the fake news scales only provided binary values (i.e., either green or red). This means that the arrow on the fake news scale could only be in two positions: at the scale's green end or red end. From a perceptual perspective, a visual orientation in the direction of small black arrows is automatic, even involuntary, because "arrows reflexively activate an oculomotor response" (Kuhn and Kingstone, 2009, p. 324), meaning that readers would immediately look to the indicated colors. Since all users, sources, and news articles were invented, the fake news scales in the experiment did not show their actual reliability. Instead, their values were manipulated according to the experimental design. In the following, we use the term fake news flag for the scales.

We placed the fake news flags near the user, source, and news article of the social media news posts (see Figure 1). To manipulate the number of fake news flags that indicated a warning or verification respectively, we created two experimental groups. In one group, participants received the social media news posts with two fake news flags showing a warning and one fake news flag showing a verification. Conversely, in the other group, we had one fake news flag showing a warning and two fake news flags showing a verification.

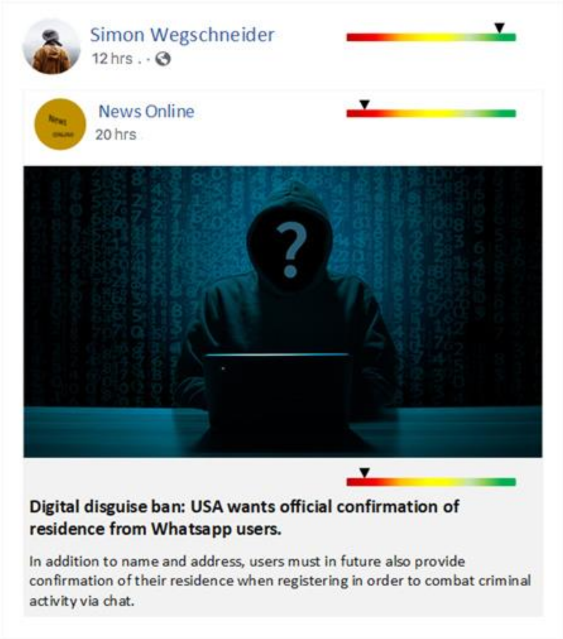

Figure 1: Example of a social media news post.

3.3 Eye tracking apparatus and areas of interest

In this study, we employed a Tobii TX300 eye tracking system. The eye tracker had a gaze sample rate of 300 Hz and an integrated 23-inch monitor with a screen resolution of 1920 x 1080. We defined a minimum fixation duration of 60 ms and a velocity-threshold identification fixation classifier (I-VT) of 30 degrees/second, which are Tobii's standard values for determining fixations (Tobii, 2021). To analyze the fixations, we defined areas of interest (AOIs) around the three fake news flags. To determine the buffers around each AOI, we used the average accuracy values from the calibration process in relation to the screen resolution, which resulted in an 11 pixel tolerance around elements.

3.4 Sample and procedure

Thirty-four participants (27 females and 7 males) followed our invitation and signed up for a study on news headlines on Facebook. Most of them (79%) were students at a European university. Out of the seven non-students, six were currently employed. Participants' highest level of educational qualification was either a secondary education degree (59%), a bachelor's degree (38%), or a master's degree (3%). In addition, participants' age ranged between 18 and 40 years (M = 24.38; SD = 4.47) and they indicated to spend between one and 36 hours (M = 10.47; SD = 7.34) per week on social networks.

Before the experiment started, we asked for the informed consent of each participant and emphasized that participation was voluntary and that the experiment could be quit at any time without giving any reasons. We conducted the experiment with one participant at a time. After receiving the informed consent, we instructed each participant to sit down on a chair and read the letters from an eye test poster hanging on the wall. The eye test was in place to ensure a correct vision on the eye tracking monitor. Next, we helped participants find a suitable and comfortable sitting position, where the distance between the eye tracker and a participant's eyes was at 65 cm with a tolerance of +/- 5 cm. Afterward, we adjusted the eye tracker to the participant and ran a 9-point calibration to realize high accuracy and precision levels of the recorded eye movements and reduce potential confounding factors (e.g., slightly turned head positions). We accepted accuracy deviations of up to 0.7° as the maximum tolerance level. After the calibration, participants began with the experiment. We implemented the experiment on the online survey tool Sosci Survey[1]. When participants clicked on the start of the survey, a build-in random number generator assigned participants into experimental groups. Subsequently, participants were provided with a trial page, which contained an example of a social media news post with fake news scales. Our intention was to give participants a situation to become familiar with the set-up and interface. Then, participants went through 12 pages, each containing a social media news post and fake news scales. The order of the presentation was randomized. Additionally, all pages had been consistently displayed for 30 seconds to allow for a fair comparison of attention distribution. After the time expired, a window popped up, which informed the participant to proceed by clicking "OK." Between pages, participants indicated their attitudes toward the previously seen news article by answering questions (7-point Likert scale from "extremely negative" to "extremely positive"). At the end of the experiment, participants filled out a survey, which collected demographic variables. Overall, participation in the experiment took between 15 and 26 minutes (M = 21 minutes).

3.5 Measures

Our operationalization included two measures. First, **confirmation bias** was measured by an adapted version of the self-reported items used by Kim and Dennis (2019). After each page with a social media news post, participants were directed to a page on which they rated the perceived importance of the news article (7-point Likert scale from "not at all important" to "extremely important") and their position on the news article (7-point Likert scale from "extremely negative" to "extremely positive"). Second, visual attention to the fake news flag was measured by the total fixation duration, which represents the total duration of all fixations on a particular stimulus (Peschel and Orquin, 2013) and which is a common operationalization for visual attention (van der Laan et al., 2015). We determined the total duration of fixation for each AOI placed over a particular fake news flag.

4 Results

Prior to the analysis, we checked how long participants stayed on the social media pages since participants had to click on an "OK" button after the time a page expired in order to be redirected to the next page. Therefore, we conducted an outlier analysis based on z-scores to reduce the possibility of biased data. Given a z-score threshold of 1.96, we excluded a total of 14 pages from 11 participants from the further analysis. The final dataset from 34 participants encompassed 466 stimuli (single social media pages). To consider that participants received multiple news articles during the experiment, we performed a linear mixed-effects analysis based on the R package lme4 (Bates et al., 2020) to test our hypotheses. Since we wanted to investigate effects between distinct fake news flags, we set the unit of analysis to individual fake news flags, i.e., each row of the dataset represents a single fake news flags.

To answer our research questions, we analyzed the effects of the position of the fake news flags (article, user, source), its fake news indication (warning or verification), the uniqueness of the fake news indication (unique or non-unique), as well as the moderating effect of confirmation bias for fake news indication and uniqueness of the fake news indication, on visual attention to the respective fake news flag. Confirmation bias was a continuous variable and centered and rescaled with mean normalization. The other variables were categorical and implemented as treatment or effect contrasts respectively (Schad et al., 2020). We used treatment contrasts to code the position of the fake news flag in the linear mixed model, where we set the news article as the baseline condition for the comparison with user and source. The fake news indication and the uniqueness of the fake news indication were implemented as effect contrasts. Each of these variables and the interaction terms between fake news indication and confirmation bias as well as between the uniqueness of the fake news indication and confirmation bias were entered as fixed effects into the linear mixed model. In addition, we included random effects for participants, social media news posts, and the order of the social media news post (first presented post

[1] www.soscisurvey.de

to twelfth post) in the model. We first tested the model with a complete set of random intercepts and slopes for each defined fixed effect. However, this model did not converge, so we followed the recommendation of Matuschek et al. (2017) and incrementally reduced the complexity of the set of random effects in the model by removing random slopes until the maximum power of the model was accomplished. At each step, we manually compared the model's Bayesian information criterion (BIC) with the prior model to consider possible information loss after removing random slopes. Eventually, we stopped this procedure at a set of random effects that included random intercepts and slopes for the position of the fake news scale depending on the participant and random intercepts depending on the social media news post and its position. We visually inspected a QQ-diagram of the distribution of residuals of our final linear mixed model, which revealed normally distributed residuals.

We show the descriptive statistics in Figure 2 and the linear mixed model results in Table 2. In our first hypotheses regarding the position of the flags, we had stated that fake news flags receive less visual attention when they are attached to (H1a) the *source* or (H1b) the *user* rather than the news article. Indeed, our data shows that participants visually attended the fake news flags significantly less when the scales were attached to the user relative to the news article ($b = -0.253$, $SE = 0.046$, $p < .001$) and to the source relative to the news article ($b = -0.211$, $SE = 0.051$, $p < .001$). Therefore, both (H1a) and (H1b) of our first hypothesis are supported. As for the fake news indication, our second hypothesis predicted that fake news flags receive more attention when they show a warning compared to a verification. However, the effect of fake news indication was not significant ($b = 0.039$, $SE = 0.024$, $p > .05$). Consequently, H2 must be rejected. Considering the uniqueness of the fake news indication, our hypothesis H3 stated that the visual attention to fake news flags is higher for a flag that uniquely indicates a warning or verification than for a flag that shares its fake news indication with another flag. Here, our data showed that participants visually attended the fake news flags more when their fake news indication was unique, e.g., a verification among two warnings or a warning among two verifications ($b = 0.062$, $SE = 0.024$, $p = .01$). Therefore, hypothesis H3 is supported. Our final hypotheses addressed the moderating role of confirmation bias regarding the fake news indication and the uniqueness of the fake news indication.

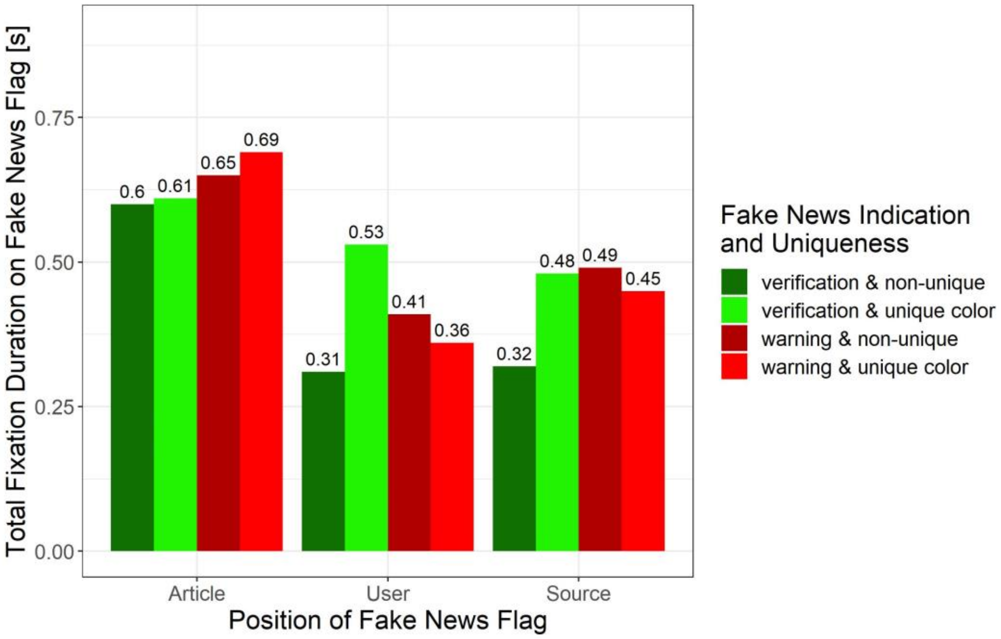

Figure 2: Visual attendance to the fake news scale.

Dependent Variable: Attention to Fake News Flag	**Fixed-Effects Estimates b**	**Standard Error**	**T**	**p**
Intercept	0.654***	0.048	13.558	<.001
Predictors				
User (Compared to baseline = article)	-0.254***	0.04574	-5.550	<.001
Source (Compared to baseline = article)	-0.212***	0.051	-4.171	<.001
Fake News Indication (warning vs. verification)	0.039	0.024	1.632	.103
Uniqueness of Fake News Indication (unique vs. non-unique)	0.061*	0.024	2.575	.010
Confirmation Bias	0.092	0.056	1.646	.100
Interaction				
Fake News Indication x Confirmation Bias	0.273**	0.098	2.789	.005
Uniqueness of Fake News Indication x Confirmation Bias	0.250*	0.103	2.430	.015
R^2	0.350			
Adj. R^2	0.069			
*p < 0.05, **p < 0.01, *** p < 0.001				

Table 2. Linear mixed-effects model.

In H4a, we had stated that confirmation bias positively moderates the relationship between fake news indication and visual attention to the fake news flag. Since the interaction between fake news indication and confirmation bias shows a significant positive effect (b = 0.273, SE = 0.098, p = .005), we find a positive moderation in so far that under a higher level of confirmation bias, a fake news flag displaying a warning receives more visual attention, while a flag that displays a verification receives less. As illustrated in Figure 3, warnings lead to more visual attention to the fake news flag at a higher level of confirmation bias and verifications at a lower level of confirmation bias. Hence, H4a is supported. Finally, H4b had suggested that confirmation bias positively moderates the relationship between the uniqueness of the fake news indication and the visual attention to the fake news flag. This hypothesis was tested in the interaction between the uniqueness of the fake news indication and confirmation bias, which was significant and positive (b = 0.250, SE = 0.103, p = .015). Thus, we also find a positive moderation in that under a higher level of confirmation bias, a fake news flag that uniquely displays a warning or verification receives more visual attention relative to a flag that shares its fake news indication with another flag, leading to support for H4b.

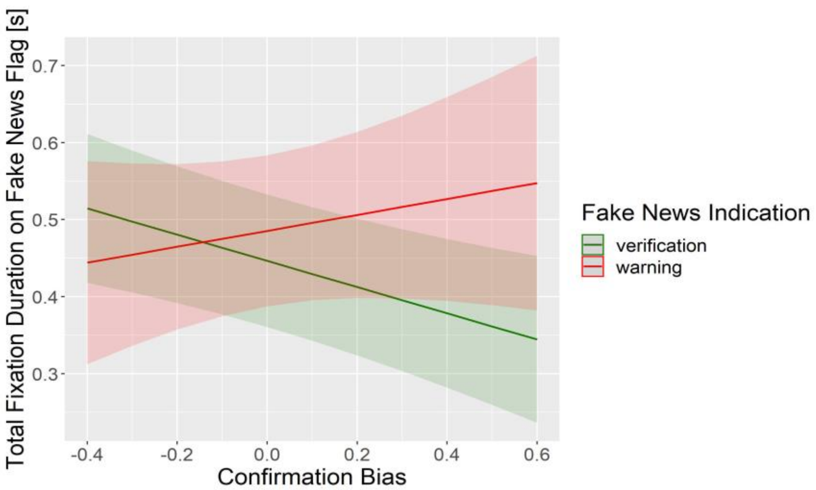

Figure 3: Interaction between confirmation bias and fake news indication.

5 Discussion

In this paper, we investigated the display of multiple fake news warnings and verifications in social media posts and their influence on visual attention. In detail, we compared three different fake news flag positions in social media posts, one for the user who shared the news article, the source of the news article, and the news article itself. Additionally, we examined the role of the fake news indication and the uniqueness of the fake news indication in a laboratory eye tracking experiment. First, our findings revealed that the position of the flags matters in that fake news flags attached to the news article receive more visual attention compared to those flags attached to the user or source. Although this design choice raises concerns about practical implementation strategies, such as discriminating users with bad ratings, it allowed us to assess multiple fake news flag positions distinguishing source and content and their inherent information cues. In addition, it would be quite conceivable for corporate and public figure accounts to be thoroughly evaluated and forced to position themselves towards other users, as their reach far exceeds that of ordinary users. Twitter, for example, has already embarked on such a path, when it has introduced labels for state-affiliated media accounts with the intention to indirectly highlight the lack of political independence of the content of posts from such accounts (TwitterSupport, 2020). Second, we investigated conditions under which fake news flags receive more visual attention. Surprisingly, a fake news flag that displayed a warning did not draw more visual attention compared to the indication of a verification. However, the results tended in the predicted direction. In addition, we identified a significant effect of the uniqueness of a flag's fake news indication. In cases where a fake news flag was the only flag displaying a warning or verification, while the two other flags indicated the opposite, the unique flag received more visual attention. Additionally, we discovered that participants' level of confirmation bias toward a news article moderated the relationships between the fake news indication and visual attention as well as the relationships between the uniqueness of the fake news indication and visual attention.

Our findings complement current fake news flag literature in various ways. First, we spotlight a potential prerequisite for the effectiveness of fake news flags, namely visual attention (Ball et al., 2006). Since our findings suggest that flags assigned to the news article itself attract more visual attention than those assigned to the user or the source, we can derive implications on how to assign fake news flags in news articles on social media. Due to their content-related quality, fake news flags associated with the news article seem to play a more important role in judging the believability of a news article than fake news flags associated with the source or user. We explain this finding based on credibility judgment research, which understands the fake news flags as information cues to assess the believability of news articles (Reinhard and Sporer, 2010). Cues on the believability of the user and source aid only as proxies, as believable users or sources are likely, yet not necessarily, to post believable articles. In contrast, fake news flags for the article are information cues that contain information that directly affects the news article, so it is plausible to perceive them as more important in that they receive more visual attention during the information search process of the credibility judgment.

Second, the relevance of visual information processing theories such as the feature integration theory (Treisman and Gelade, 1980; Wolfe and Horowitz, 2017) for the context of fake news flag research is clearly supported by the current findings. In our study, warnings or verifications were implemented as scales with an arrow pointing to distinct colors (red and green). As predicted by the feature integration theory, we observed more visual attention for unique warnings or verifications, where a unique fake news indication resembled a different color compared to all other flags that are attached to the same social media post. Hence, the unique color 'popped out' in the pre-attentive state, drawing the visual attention toward the flag. Such an effect is particularly interesting when considering that two out of the three scales indicate conflicting information, yet receive less visual attention, and, therefore, could potentially have a minor impact on social media users when assessing believability. That might be a serious challenge, when simultaneously displaying multiple but conflicting fake news flags at different hierarchical levels (user, source, news article) to provide social media users with an appropriate informational basis for evaluating news posts. We, therefore, contribute to the literature stream on fake news flag attached to social media posts to influence users in assessing the believability of news posts (e.g., Kim and Dennis, 2019; Moravec et al., 2020).

Third, our findings add to research on understanding confirmation bias in the context of fake news flags (Moravec et al., 2019; Kim et al., 2019b). Indeed, the moderating role of confirmation bias in the relationship between a flag and visual attention may be indicative of participants' cognitive dissonance. When participants had high levels of confirmation bias, which means they wanted the news article to be true, they fixated longer on those fake news flags that displayed a warning and thus did not support their opinion. This could be a sign of an increased cognitive effort, which is required to resolve the existing conflict (Colosio et al., 2017; Horstmann et al., 2009). Fake news flags evoking cognitive dissonance have been observed in prior experiments (Kim and Dennis, 2019; Figl et al., 2019), and measured via electroencephalography (Moravec et al., 2019), but to the best of our knowledge, this study is one of the first to find evidence based on eye tracking data.

6 Limitations and Future Research

Our experimental design, however, involved some limitations, which can offer avenues for future research. Firstly, although eye tracking is considered an unobtrusive measure (Glaholt and Reingold, 2011), an environment where participants know that their gazes are recorded is inevitably different from a setting where they are, for instance, at home and reading news articles at their own devices. All indicator scores were artificially counterbalanced, and we included no experimental group in which all three fake news scales indicated the same. In fact, all news articles, sources, and users were made-up. While we deemed the artificial nature of all articles to be an advantage to compare the different social media news posts, we cannot eliminate the possibility that participants questioned the indicator scores of the fake news scales or perceived the scores as artificial. Hence, future research could replicate our study, but use both true and fake news with a condition where the fake news scales actually indicate the respective believability. In addition, our findings could be replicated in an online experiment.

Additionally, we only considered binary fake news indications (warning or verification) for our fake news flags. In practice, however, the advantage of fake news flags is that they can depict a more nuanced evaluation of believability by exhausting the interval between the scale's extremes. Although this could be relevant for both research and practice, our design, particularly the counterbalancing of scores among each news article, forced us to use a manageable set of possible indicator scores. Future researchers could extend the current work by using research designs that can better handle indicator scores on an interval between verification and warning, which would unravel how these more fine-grained believability cues are perceived.

Finally, our experiment measured confirmation bias toward a news article after participants had seen the respective social media news post. We had to measure confirmation bias for each post separately. Although this has been used in prior research (Kim and Dennis, 2019; Moravec et al., 2020), we cannot exclude the possibility that the presence of the fake news flags in the social media posts could have influenced participants' responses to these questions. However, we expect such an effect to be minimal. Confirmation bias is based on attitudes toward a topic and should thus not change much after a short display of (conflicting) fake news flags.

7 Conclusion

One of the more significant findings to emerge from this study is that visual attention to fake news scales directly attached to the news article is higher relative to those attached to the user or source. In addition, we provide empirical evidence that fake news scales receive more visual attention when they are uniquely signaling a warning or a verification. Finally, we identified the recipients' confirmation bias to moderate the extent to which scale colors and the uniqueness of a scale color increase visual attention. These findings assist an understanding of effective fake news flags design, and, in turn, they can guide social media platforms in their effort to combat fake news.

References

Allcott, H. and M. Gentzkow (2017). "Social media and fake news in the 2016 election." *Journal of Economic Perspectives,* 31 (2), 211-36.

Bago, B. and W. De Neys (2017). "Fast logic?: Examining the time course assumption of dual process theory." *Cognition,* 158, 90-109.

Ball, L., P. Phillips, C. Wade and J. Quayle (2006). "Effects of Belief and Logic on Syllogistic Reasoning: Eye-Movement Evidence for Selective Processing Models." *Experimental Psychology,* 53, 77-86.

Bastick, Z. (2021). "Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation." *Computers in Human Behavior,* 116, 106633.

Batailler, C., S. Brannon, P. Teas and B. Gawronski (2021). "A Signal Detection Approach to Understanding the Identification of Fake News." *Perspectives on Psychological Science,* 17 (1), 78-98.

Bates, D. et al. (2020). *lme4: Linear mixed-effects models using 'Eigen' and S4. Version 1.1-23*.

Bernier, C.L. (1984). "Second-hand knowledge. An inquiry into cognitive authority. Patrick Wilson. Greenwood Press, 1983." *Journal of the American Society for Information Science,* 35 (4), 254-255.

Brashier, N.M., E.D. Eliseev and E.J. Marsh (2020). "An initial accuracy focus prevents illusory truth." *Cognition,* 194, 104054.

Bunde, E., N. Kühl and C. Meske (2022). Fake or Credible? Towards Designing Services to Support Users' Credibility Assessment of News Content. *Hawaii International Conference on System Sciences,* Hawaii.

Cave, K.R. and M.J. Batty (2006). "From searching for features to searching for threat: Drawing the boundary between preattentive and attentive vision." *Visual Cognition,* 14 (4-8), 629-646.

Chou, W.-y.S., N. Trivedi, E. Peterson, A. Gaysynsky, M. Krakow and E. Vraga (2020). "How do social media users process cancer prevention messages on Facebook? An eye-tracking study." *Patient Education and Counseling,* 103 (6), 1161-1167.

Colosio, M., A. Shestakova, V.V. Nikulin, E. Blagovechtchenski and V. Klucharev (2017). "Neural mechanisms of cognitive dissonance (revised): An EEG study." *The Journal of Neuroscience,* 37 (20), 5074-5083.

Di Domenico, G., D. Nunan, J. Sit and V. Pitardi (2021). "Free but fake speech: When giving primacy to the source decreases misinformation sharing on social media." *Psychology & Marketing,* 38 (10), 1700-1711.

Duffy, A. (2020). *Smartphones and the News*. Routledge.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford University Press.

Figl, K., S. Kießling, C. Rank and S. Vakulenko (2019). Fake news flags, cognitive dissonance, and the believability of social media posts. *International Conference on Information Systems*, Munich, Germany.

Flynn, D.J., B. Nyhan and J. Reifler (2017). "The Nature and Origins of Misperceptions: Understanding False and Unsupported Beliefs About Politics." *Political Psychology,* 38 (1), 127-150.

Gabielkov, M., A. Ramachandran, A. Chaintreau and A. Legout (2016). Social clicks: What and who gets read on twitter? *ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science,* Antibes Juan-les-Pins, France, 179-192.

Glaholt, M. and E. Reingold (2011). "Eye Movement Monitoring as a Process Tracing Methodology in Decision Making Research." *Journal of Neuroscience, Psychology, and Economics,* 4, 125-146.

Guazzini, A., E. Yoneki and G. Gronchi (2015). "Cognitive dissonance and social influence effects on preference judgments: An eye tracking based system for their automatic assessment." *International Journal of Human-Computer Studies,* 73, 12-18.

Haug, M. and H. Gewald (2018). Are you the Reason Fake News exists? Investigating News Sharing Attitude on Twitter. *International Conference on Information Systems,* San Francisco, 9.

Heimbach, I. and O. Hinz (2018). "The Impact of Sharing Mechanism Design on Content Sharing in Online Social Networks." *Information Systems Research,* 29 (3), 592-611.

Henderson, E.L., S.J. Westwood and D.J. Simons (2021). "A reproducible systematic map of research on the illusory truth effect." *Psychonomic Bulletin & Review,* 1-24.

Horstmann, N., A. Ahlgrimm and A. Glöckner (2009). "How distinct are intuition and deliberation? An eye-tracking analysis of instruction-induced decision modes." *Judgment and Decision Making,* 4 (5).

Islam, M.S. et al. (2020). "COVID-19–Related Infodemic and Its Impact on Public Health: A Global Social Media Analysis." *The American Journal of Tropical Medicine and Hygiene,* 103 (4), 1621-1629.

Izuma, K., M. Matsumoto, K. Murayama, K. Samejima, N. Sadato and K. Matsumoto (2010). "Neural correlates of cognitive dissonance and choice-induced preference change." *Proceedings of the National Academy of Sciences,* 107 (51), 22014-22019.

Janze, C. and M. Risius (2017). Automatic Detection of Fake News on Social Media Platforms. *Pacific Asia Conference on Information Systems,* Langkawi, Malaysia, 261.

Jin, Z., J. Cao, Y. Zhang and J. Luo (2016). News Verification by Exploiting Conflicting Social Viewpoints in Microblogs. *AAAI Conference on Artificial Intelligence,* Phoenix, Arizona USA.

Kim, A. and A.R. Dennis (2019). "Says who? The effects of presentation format and source rating on fake news in social media." *MIS Quarterly,* 43 (3), 1025-1039.

Kim, A., P. Moravec and A.R. Dennis (2019a). "Combating fake news on social media with source ratings: The effects of user and expert reputation ratings." *Journal of Management Information Systems,* 36 (3), 931-968.

Kim, A., P. Moravec and A.R. Dennis (2019b). "When do details matter? Source rating summaries and details in the fight against fake news on social media." *Kelley School of Business Research Paper No. 19-52*.

Kim, E.J., C. Choi and S. Tanford (2020). "Influence of scarcity on travel decisions and cognitive dissonance." *Asia Pacific Journal of Tourism Research,* 25 (7), 721-735.

Kim, S.C., E.K. Vraga and J. Cook (2021). "An Eye Tracking Approach to Understanding Misinformation and Correction Strategies on Social Media: The Mediating Role of Attention and Credibility to Reduce HPV Vaccine Misperceptions." *Health Communication,* 36 (13), 1687-1696.

Knobloch-Westerwick, S. and S.M. Lavis (2017). "Selecting Serious or Satirical, Supporting or Stirring News? Selective Exposure to Partisan versus Mockery News Online Videos." *Journal of Communication,* 67 (1), 54-81.

Koehler, K., F. Guo, S. Zhang and M.P. Eckstein (2014). "What do saliency models predict?" *Journal of Vision,* 14 (3), 14-14.

Kuhn, G. and A. Kingstone (2009). "Look away! Eyes and arrows engage oculomotor responses automatically." *Attention, Perception, & Psychophysics,* 71 (2), 314-327.

Ladeira, W.J., M. Dalmoro, F.d.O. Santini and W.C. Jardim (2021). "Visual cognition of fake news: the effects of consumer brand engagement." *Journal of Marketing Communications,* 1-21.

Laughery, K.R. and M.S. Wogalter (2006). "Designing Effective Warnings." *Reviews of Human Factors and Ergonomics,* 2 (1), 241-271.

Lewandowsky, S., U.K.H. Ecker, C.M. Seifert, N. Schwarz and J. Cook (2012). "Misinformation and Its Correction: Continued Influence and Successful Debiasing." *Psychological Science in the Public Interest,* 13 (3), 106-131.

Matuschek, H., R. Kliegl, S. Vasishth, H. Baayen and D. Bates (2017). "Balancing Type I error and power in linear mixed models." *Journal of Memory and Language,* 94, 305-315.

Moody, D.L. (2009). "The "physics" of notations: Towards a scientific basis for constructing visual notations in software engineering." *IEEE Transactions on Software Engineering,* 35 (5), 756-779.

Moravec, P., A. Kim and A.R. Dennis (2018). Flagging fake news: System 1 vs. system 2. *International Conference on Information Systems,* San Francisco, CA, USA.

Moravec, P.L., A. Kim and A.R. Dennis (2020). "Appealing to Sense and Sensibility: System 1 and System 2 Interventions for Fake News on Social Media." *Information Systems Research,* 31 (3), 987-1006.

Moravec, P.L., R.K. Minas and A.R. Dennis (2019). "Fake news on social media: People believe what they want to believe when it makes no sense at all." *MIS Quarterly,* 43 (4), 1343-1360.

Murray, S., M. Stanley, J. Mcphetres, G. Pennycook and P. Seli (2020). *"I've said it before and I will say it again...": Repeating statements made by Donald Trump increases perceived truthfulness for individuals across the political spectrum*. PsyArXiv. URL: https://doi.org/10.31234/osf.io/9evzc.

Nadarevic, L. and A. Aßfalg (2017). "Unveiling the truth: warnings reduce the repetition-based truth effect." *Psychological Research,* 81 (4), 814-826.

Newman, N., R. Fletcher, A. Schulz, S. Andı, C.T. Robertson and R.K. Nielsen (2021). *Reuters Institute Digital News Report 2021*. URL: https://reutersinstitute.politics.ox.ac.uk/digital-news-report/2021.

Nickerson, R.S. (1998). "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises." *Review of General Psychology,* 2 (2), 175-220.

Öhman, A., A. Flykt and F. Esteves (2001). "Emotion drives attention: Detecting the snake in the grass." *Journal of Experimental Psychology: General,* 130 (3), 466-478.

Ozbay, F.A. and B. Alatas (2020). "Fake news detection within online social media using supervised artificial intelligence algorithms." *Physica A: Statistical Mechanics and its Applications,* 540.

Pennycook, G., A. Bear, E. Collins and D.G. Rand (2020a). "The implied truth effect: Attaching warnings to a subset of fake news headlines increases perceived accuracy of headlines without warnings." *Management Science,* 66 (11), 1-14.

Pennycook, G., J. Binnendyk, C. Newton and D.G. Rand (2021). "A Practical Guide to Doing Behavioral Research on Fake News and Misinformation." *Collabra: Psychology,* 7 (1).

Pennycook, G., T.D. Cannon and D.G. Rand (2018). "Prior exposure increases perceived accuracy of fake news." *Journal of Experimental Psychology: General,* 147 (12), 1865-1880.

Pennycook, G., J. McPhetres, Y. Zhang, J.G. Lu and D.G. Rand (2020b). "Fighting COVID-19 Misinformation on Social Media: Experimental Evidence for a Scalable Accuracy-Nudge Intervention." *Psychological Science,* 31 (7), 770-780.

Pennycook, G. and D.G. Rand (2021). "The Psychology of Fake News." *Trends in Cognitive Sciences,* 25 (5), 388-402.

Peschel, A. and J. Orquin (2013). "A review of the findings and theories on surface size effects on visual attention." *Frontiers in Psychology,* 4 (902).

Rayner, K. (2009). "Eye movements and attention in reading, scene perception, and visual search." *Quarterly Journal of Experimental Psychology,* 62 (8), 1457-1506.

Reinhard, M.-A. and S.L. Sporer (2010). "Content Versus Source Cue Information as a Basis for Credibility Judgments." *Social Psychology,* 41 (2), 93-104.

Rieh, S.Y. (2002). "Judgment of information quality and cognitive authority in the Web." *Journal of the American Society for Information Science and Technology,* 53 (2), 145-161.

Riggio, L. and K. Kirsner (1997). "The relationship between central cues and peripheral cues in covert visual orientation." *Attention Perception & Psychophysics,* 59 (6), 885-99.

Ross, B., A. Jung, J. Heisel and S. Stieglitz (2018). Fake news on social media: The (in)effectiveness of warning messages. *International Conference on Information Systems*, San Francisco, CA, USA.

Rubin, V., N. Conroy and Y. Chen (2015). Towards News Verification: Deception Detection Methods for News Discourse. *Hawaii International Conference on System Sciences*, Kauai, Hawaii.

Schad, D.J., S. Vasishth, S. Hohenstein and R. Kliegl (2020). "How to capitalize on a priori contrasts in linear (mixed) models: A tutorial." *Journal of Memory and Language,* 110, 1-40.

Scherer, L.D. and G. Pennycook (2020). "Who Is Susceptible to Online Health Misinformation?" *American Journal of Public Health,* 110 (3), 276-277.

Shi, L., N. Bhattacharya, A. Das, M. Lease and J. Gwizdka (2022). The Effects of Interactive AI Design on User Behavior: An Eye-tracking Study of Fact-checking COVID-19 Claims. *ACM SIGIR Conference on Human Information Interaction and Retrieval,* Regensburg, Germany, 315–320.

Shu, K., L. Cui, S. Wang, D. Lee and H. Liu (2019). defend: Explainable fake news detection. *ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 395-405.

Shu, K., A. Sliva, S. Wang, J. Tang and H. Liu (2017). "Fake news detection on social media: A data mining perspective." *ACM SIGKDD Explorations Newsletter,* 19 (1), 22-36.

Simko, J., M. Hanakova, P. Racsko, M. Tomlein, R. Moro and M. Bielikova (2019). Fake news reading on social media: an eye-tracking study. *ACM Conference on Hypertext and Social Media*, Hof, Germany, 221-230.

Simko, J., P. Racsko, M. Tomlein, M. Hanakova, R. Moro and M. Bielikova (2021). "A study of fake news reading and annotating in social media context." *New Review of Hypermedia and Multimedia,* 27 (1-2), 97-127.

Skurnik, I., C. Yoon, D.C. Park and N. Schwarz (2005). "How Warnings about False Claims Become Recommendations." *Journal of Consumer Research,* 31 (4), 713-724.

Smith, N.K., J.T. Cacioppo, J.T. Larsen and T.L. Chartrand (2003). "May I have your attention, please: Electrocortical responses to positive and negative stimuli." *Neuropsychologia,* 41 (2), 171-183.

Tanaka, Y., M. Inuzuka and R. Hirayama (2019). Utilizing eye-tracking to explain variation in response to inconsistent message on belief change in false rumor. *41th Annual Conference of the Cognitive Science Society (CogSci)*, Montreal, Canada, 2926-2932.

Tobii, 2021. *Pro Lab User Manual* [Online]. Tobii Pro AB. URL: https://www.tobiipro.com/siteassets/tobii-pro/user-manuals/Tobii-Pro-Lab-User-Manual/ (visited on 16.11.2021).

Treisman, A. and G. Gelade (1980). “A feature-integration theory of attention.” *Cognitive Psychology,* 12 (1), 97-136.

Trivedi, N., M. Krakow, K. Hyatt Hawkins, E.B. Peterson and W.-Y.S. Chou (2020). “"Well, the Message Is From the Institute of Something": Exploring Source Trust of Cancer-Related Messages on Simulated Facebook Posts.” *Frontiers in Communication,* 5.

Trivedi, N., M. Lowry, A. Gaysynsky and W.-Y.S. Chou (2021). “Factors Associated with Cancer Message Believability: a Mixed Methods Study on Simulated Facebook Posts.” *Journal of Cancer Education,* 1-9.

TwitterSupport, 2020. *New labels for government and state-affiliated media accounts* [Online]. URL: https://blog.twitter.com/en_us/topics/product/2020/new-labels-for-government-and-state-affiliated-media-accounts (visited on 28.03.2022).

van der Laan, L.N., I.T.C. Hooge, D.T.D. de Ridder, M.A. Viergever and P.A.M. Smeets (2015). “Do you like what you see? The role of first fixation and total fixation duration in consumer choice.” *Food Quality and Preference,* 39, 46-55.

Vosoughi, S., D. Roy and S. Aral (2018). “The spread of true and false news online.” *Science,* 359 (6380), 1146-1151.

Wogalter, M.S. (2005). Factors that Influence the Effectiveness of Warning Signs and Labels. In MORAES, A.d. and AMADO, G. (eds.) ErgoDesign/USIHC: Colletanea de Palestras de Convidados Internacionais e Nacionais. Rio de Janeiro Brazil.

Wolfe, J.M. and T.S. Horowitz (2017). “Five factors that guide attention in visual search.” *Nature Human Behaviour,* 1 (3), 1-8.

Yin, D., S. Mitra and H. Zhang (2016). “Research Note—When Do Consumers Value Positive vs. Negative Reviews? An Empirical Investigation of Confirmation Bias in Online Word of Mouth.” *Information Systems Research,* 27 (1), 131-144.

Zhuravskaya, E., M. Petrova and R. Enikolopov (2020). “Political Effects of the Internet and Social Media.” *Annual Review of Economics,* 12, 415-438.

Zubiaga, A., A. Aker, K. Bontcheva, M. Liakata and R. Procter (2018). “Detection and Resolution of Rumours in Social Media: A Survey.” *ACM Computing Surveys (CSUR),* 51 (2), 32.