True or false? Cognitive load when reading COVID-19 news headlines: an eye-tracking study

Li Shi
IX Lab, School of Information
The University of Texas at Austin
Austin, Texas, USA
lilylashi@utexas.edu

Nilavra Bhattacharya
IX Lab, School of Information
The University of Texas at Austin
Austin, Texas, USA
nilavra@ieee.org

Anubrata Das
School of Information
The University of Texas at Austin
Austin, Texas, USA
anubrata@utexas.edu

Jacek Gwizdka
IX Lab, School of Information
The University of Texas at Austin
Austin, Texas, USA
jacekg@utexas.edu

ABSTRACT

Misinformation is an important topic in the Information Retrieval (IR) context and has implications for both system-centered and user-centered IR. While it has been established that the performance in discerning misinformation is affected by a person's cognitive load, the variation in cognitive load in judging the veracity of news is less understood. To understand the variation in cognitive load imposed by reading news headlines related to COVID-19 claims, within the context of a fact-checking system, we conducted a within-subject, lab-based, quasi-experiment (N=40) with eye-tracking. Our results suggest that examining true claims imposed a higher cognitive load on participants when news headlines provided incorrect evidence for a claim and were inconsistent with the person's prior beliefs. In contrast, checking false claims imposed a higher cognitive load when the news headlines provided correct evidence for a claim and were consistent with the participants' prior beliefs. However, changing beliefs after examining a claim did not have a significant relationship with cognitive load while reading the news headlines. The results illustrate that reading news headlines related to true and false claims in the fact-checking context impose different levels of cognitive load. Our findings suggest that user engagement with tools for discerning misinformation needs to account for the possible variation in the mental effort involved in different information contexts.

CCS CONCEPTS

• **Information systems → Search interfaces**; • **Human-centered computing → User studies**; **Empirical studies in interaction design**.

KEYWORDS

fact checking, misinformation, cognitive load, pupil dilation

ACM Reference Format:
Li Shi, Nilavra Bhattacharya, Anubrata Das, and Jacek Gwizdka. 2023. True or false? Cognitive load when reading COVID-19 news headlines: an eye-tracking study. In *ACM SIGIR Conference on Human Information Interaction and Retrieval (CHIIR '23), March 19–23, 2023, Austin, TX, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3576840.3578290

1 INTRODUCTION

With increasing access to information technologies, misinformation becomes more easily and widely disseminated through social media [Del Vicario et al. 2016]. Research shows that misinformation spreads more rapidly than true information. Widespread misinformation can potentially disrupt public health, democratic processes, and public discourse [Hochschild and Einstein 2015; Xie et al. 2020]. Consequently, there is a growing public and academic interest in tackling the challenges associated with misinformation. Specifically, fact-checking as a method for curbing misinformation has gained attention in research and practice [Arnold 2020; Graves and Amazeen 2019].

While we have seen a surge of fact-checking organizations in the past few years, addressing the gigantic scale of misinformation online may be intractable with expert fact-checkers alone. Consequently, we have seen interest in automating the fact-checking process, either towards building end-to-end automation, or as decision support systems for fact-checkers [Graves 2018; Guo et al. 2022; Nakov et al. 2021].

Automated fact-checking research are often evaluated using automated metrics. However, the primary goal of automated fact-checking tools are to help fact-checkers with their decision making. Recent research has started looking into several human factors associated with fact-checking such as usability, trust, and intelligibility [Das et al. 2023; Jiang et al. 2020; Nguyen et al. 2018; Shi et al. 2022]. Additionally, some research has also analyzed the factors affecting human judgment during the interaction with misinformation, including the reading behaviors and beliefs in the misinformation [Ecker et al. 2022; Roozenbeek et al. 2020]. However, a detailed understanding of users cognitive load and fact-checking system have not been studied in great detail.

CHIIR '23, March 19–23, 2023, Austin, TX, USA

ACM ISBN 979-8-4007-0035-4/23/03…$15.00
https://doi.org/10.1145/3576840.3578290

Previous research in cognition and discerning misinformation typically displayed one news headline per trial. However, in realistic information search scenarios, people often encounter several pieces of information in one search. Moreover, the headlines shown in the previous experiments were mostly related to political topics. Currently, misinformation not only threatens democracy but also public health regarding the situation of the COVID-19 pandemic. As stated in [Zarocostas 2020], "we're not just fighting an epidemic; we're fighting an infodemic". Therefore, we designed a lab-based experiment where users checked COVID-19-related health claims in a fact-checking system and were shown multiple news headlines related to the claim simultaneously in one screen. During the experimental session, an eye-tracker recorded their pupillary response. The aim of the study was to examine how the cognitive load is impacted when users read the news headlines in a fact-checking context, and how it is related to their belief change and misinformation judgment. The contributions of this study include: (1) examining the effectiveness of pupil dilation measurements as an indicator of cognitive load in reading news headlines; (2) comparing the cognitive load imposed by reading news headlines under varied conditions of claim correctness, headline-belief stance, and evidence correctness; and (3) developing the understanding of the cognitive processes in discerning misinformation by investigating the cognitive load in a more realistic scenario.

2 BACKGROUND

2.1 Human Evaluation of Automated Fact-Checking

Several studies have evaluated automated fact-checking from a human factor perspective. Such factors include understanding, usability, intelligibility, and trust in those systems [Das et al. 2019; Mohseni et al. 2021; Nguyen et al. 2018; Shi et al. 2022]. Nguyen et al. [2018] studied the effect of intelligibility in automated fact-checking on user trust and task accuracy. Mohseni et al. [2021] examined the effect of intelligibility in calibrating user trust on automated fact-checking systems. Complementary to that, Das et al. [2019] investigated the role of confirmation bias in using automated fact-checking systems. The studies mentioned above focused on measures related to the fact-checking task and not on the user behavior while interacting with the system.

In contrast, Shi et al. [2022] examined user interaction with the fact-checking interface. They studied the effect of interactivity on several factors associated with user interaction, e.g., dwell time, attention, and mental resources, with the help of eye tracking. Our work extends such research and employs similar eye-tracking methodology in investigating users' cognitive load while interacting with automated fact-checking systems..

2.2 Cognitive load and discerning misinformation

Kahneman [2011]'s dual-process theory states that human cognition can be conceptualized as two systems, System 1 and System 2. System 1 "operates automatically and quickly, with little or no effort and no sense of voluntary control", i.e., these are autonomous and intuitive processes. System 2 "allocates attention to the effortful mental activities that demand it, including complex computations. The operations of System 2 are often associated with the subjective experience of agency, choice, and concentration", i.e., these are deliberative and analytic processes. Previous research found that people who engage analytical thinking perform better on rational thinking tests [Stanovich et al. 2011]. Similarly, the heuristic-systematic model explains that System 1 uses heuristics, while System 2 uses analysis, which makes people using System 1 more susceptible to decision-making biases [Chen et al. 1999]. People who are engaging System 2 put conscious effort into thinking and think analytically, and thus are more likely to identify misinformation. To investigate it, [Pennycook and Rand 2019] used Cognitive Reflection Test (CRT) as a measure of the willingness of activating System 2 and found that CRT is positively correlated with the capability to discern fake news. Another study investigated the impact of deliberation on correcting intuitive mistakes [Bago et al. 2020]. Researchers found that when people have more time to reconsider false claims, they are less likely to trust them [Moravec et al. 2020].

Based on the dual-process theory [Kahneman 2011], cognitive load can be used as an indication of System 2 activation, and thereby to study the performance and cognitive processes engaged in identifying fake news articles. Mirhoseini et al. [2022] found that higher cognitive load was imposed when users have better performance in discerning misinformation. Users with higher cognitive load utilize more System 2 resources, and deliberate and rationally examine the information correctness and ultimately discern misinformation. Additionally, pupillary response was shown to be a reliable physiological measure of cognitive load [Hossain and Yeasin 2014] since pupil dilation is associated with the amount of load on memory [Kahneman and Beatty 1966].

2.3 Cognitive load and pupil dilation

Cognitive load refers to the amount of working memory resources required to perform a cognitive task [Paas et al. 2016]. Typically there are three types of cognitive load measurements: task performance, subjective, and physiological [Gwizdka 2021; O'DONNELL 1986]. Task performance measures capture how well the user is performing a given task, such as task completion time, and the number of errors. Subjective measures use self-rating scales of cognitive load, such as NASA-TLX questionnaire [Hart 2006]. These measures are simple to collect but cannot reflect rapid and dynamic cognitive load changes [Palinko et al. 2010]. Physiological measures include heart-rate variability (HRV), galvanic skin response (GSR), Electroencephalography (EEG), and eye-tracking measurements [Antonenko et al. 2010; Shi et al. 2007; Urrestilla and St-Onge 2020]. Multiple eye-tracking measures, such as blink frequency and pupil dilation, have been shown to correlate with cognitive load levels [Siegle et al. 2008]. In the past decades, researchers have found that the behavior of the pupil is a direct reflection of neurological and mental activity [Hess and Polt 1964]. [Kahneman and Beatty 1966] showed that the changes of pupil diameter are related to task difficulties, and pupil dilation is associated with the amount of load on memory. Therefore, the pupillary response could be utilized as a reliable physiological measurement of cognitive load [Hossain and Yeasin 2014] in the misinformation studies.

Various metrics are used to process pupil diameter data, and therefore to estimate mental workload. A common approach is to measure pupil dilation relative to a baseline. The baseline could be the average pupil diameter measured during a baseline trial [Kruger et al. 2013] or during a baseline measurement made at the beginning of each trial [Krejtz et al. 2018], or during whole experimental session of each participant [Gwizdka 2014]. Then the pupil size difference which is calculated with respect to the baseline is so called Relative Pupil Dilation (RPD) [Gwizdka et al. 2017; Wang et al. 2021]. Another kind of alternative metric is proposed based on the moment-to-moment change in pupil diameter. This method estimates the frequency of pupil oscillation and fluctuation of pupil dilation while separating the effects of illumination. It was first proposed by Marshall with the measurement called the Index of Cognitive Activity (ICA) [Marshall 2002]. Since ICA is closed source, Duchowski et al. [2018] offered a similar and open-source, fully-detailed measurement called the Index of Pupil Dilation (IPA). The researchers further proposed the Index of High/Low Pupillary Activity (LHIPA), by taking the tonic component (LF) into account, in addition to pupil phasic response (HF), which suggested to be a more reliable indicator of cognitive load [Duchowski et al. 2020]. In this paper, we attempted to use the LHIPA and RPD as the indicators of cognitive load.

In previous research, eye-tracking was employed to investigate the effect of misinformation on cognitive activities. It was found that people fixated more frequently, had longer fixation duration, and increased pupil diameter when reading fake news compared to real news [Hansen et al. 2020; Ladeira et al. 2022; Sümer et al. 2021]. This is because reading false news imposed higher cognitive load on account of the reduced heuristic availability [Ladeira et al. 2022]. Furthermore, researchers measured pupil dilation in investigating the performance of judging the accuracy of the headlines and demonstrated that pupils dilate more when people perform better on the misinformation judgment task [Mirhoseini et al. 2022]. This study showed that higher cognitive load was associated with identifying misinformation.

Therefore, in our research, we measured pupil dilation as indication of cognitive load in information processing. We extended previous works to more realistic search scenarios, in which users encountered several relevant news headlines related to a single claim, identify misinformation, and determine the correctness of the claim. We aimed to explore how cognitive load is impacted in the fact-checking context (i.e., by the evidence correctness and users' prior beliefs), and if it was related to users' belief change. We hypothesized that:

H1: Reading news headlines that provide incorrect evidence imposes higher cognitive load.

H2: Reading news headlines that are inconsistent with their prior beliefs imposes higher cognitive load.

H3: Changing one's beliefs, and especially correcting beliefs, imposes higher cognitive load.

3 METHODS

3.1 Experimental Design

A controlled, within-subjects eye-tracking study was conducted in a usability lab at a university, with $N = 40$ participants (22 females). Participants interacted with a mock fact-checking system containing claims and news-article headlines in English language (Figure 1). Participants were pre-screened for native-level English familiarity, 20/20 vision (uncorrected or corrected), and non-expert topic familiarity of the content being shown in the fact checking system. Upon completion of the study, each participant was compensated with USD 25.

3.2 Apparatus

A Tobii TX-300 eye-tracker was used to record participants' eye movements and pupil dilation. Commercial usability and eye-tracking software was used to conduct the study, record raw gaze data, and perform partial data-cleaning and filtration for downstream analyses. Data analysis was performed in Python and R languages.

3.3 Mock Fact Checking System

Participants interacted with a mock fact checking system (Figure 1), and examined 24 COVID-19 related claims in the system. Each claim was shown at the top of the interface. Surrogates of five related news articles were presented below the claim, each with its corresponding news source, source reputation, news headline, and the article's stance towards the claim. Based on the article's stance and news source reputation, the system provided a prediction of the claim's correctness at the bottom. The news headlines were clickable, and upon clicking, opened the news article in a new browser tab. Each claim examination consisted of viewing the claim, the headlines of the news articles, and, optionally, clicking the news articles to read them in detail. To mitigate the effect of background luminance of pupil dilation, the color and luminance of the fact-checking-system interface was kept constant during the experimental session.

The claims and corresponding news-articles were on the topic of the COVID-19 pandemic. They were handpicked by the researchers to simulate a COVID-19 fact-checking system for usability analysis. Each claim was selected so as to have a pre-assigned ground-truth correctness value of TRUE, FALSE, or UNSURE (for claims that are partially true, or not totally proven at the time of data collection). The TRUE and UNSURE claims were handpicked from reputed websites in the medical domain, such as World Health Organization, WebMD, Mayoclinic, Johns Hopkins University, US State Government webpages, and others. The FALSE claims were sourced by searching for "coronavirus myths" on popular search engines. The relevant news articles for each claim were collected by manually searching the web. The source reputations for news articles were collected from existing datasets [Gruppi et al. 2020; Nørregaard et al. 2019], while the stance values of each news article towards each claim were labelled by the researchers. Two example claims are "wearing masks is not required during exercising", and "asymptomatic people can transmit COVID-19". In total there were 24 claims (8 TRUE, 8 FALSE, 8 UNSURE). The order of presenting the claims during each study session was randomized.

3.4 Procedure

The overall procedure of the experimental session is illustrated in Figure 2. Each session started with training task for participants to get familiar with the interface of the fact-checking system, and the procedure. Then the participants started the 24 trials. Each trial

Figure 1: Interface of the mock fact-checking system used in the study. Red frames represent the areas of interest (AOIs) around the news headlines, that are used in the analysis.

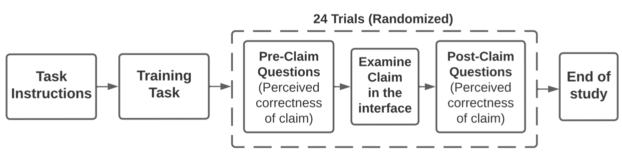

Figure 2: Flowchart of the experimental procedure.

consisted of three parts: *(i)* Pre-Claim Questions *(ii)* Examining the claim in the mock fact-checking interface, and *(iii)* Post-Claim Questions.

Pre-Claim Questions asked the following:

- Pre-perceived Correctness: *Do you think the claim is: False / Probably False / Neutral / Probably True / True*

For **examining the claim**, participants interacted with the interface freely without a time limit. Participants were also instructed to click on news headlines to open the underlying news articles in a new browser tab, and read it, if they considered it necessary for evaluating the claim.

Post-Claim Questions asked the following:

- Post-perceived Correctness: *After seeing the output of the fact-checking system, do you think the claim is: False / Probably False / Neutral / Probably True / True*

3.5 Measures

Our aim was to study cognitive load involved in reading news headlines. Previous research [Shi et al. 2022] found that most fixations on this type of fact-checking interface fell into the headline AOIs. This supports the plausibility of studying pupil dilation only on the news headline AOIs. So we marked each news headline area (Figure 1) as an *area of interest* (AOI) for eye-tracking analysis. Thus there were five AOIs in the fact-checking interface (i.e., from the first news headline to the fifth headline). Javascript function `Element.getBoundingClientRect()`[1] was used to get the coordinates for the AOIs. These coordinates were appropriately adjusted to match the coordinates recorded by the eye-tracker.

3.5.1 Claim Correctness. Each claim was selected so as to have a pre-assigned ground-truth correctness value of TRUE, FALSE, or UNSURE (denoted in UPPERCASE). This is the defined as the claim correctness. In this research, we wanted to understand the "definitive" behavior on TRUE and FALSE claims first, before trying to tease apart the more complex behavior that may be associated with UNSURE claims. Therefore, the analyses in this study only include trials in which users examined TRUE or FALSE claims.

3.5.2 Headline Stance. For each claim, we collected relevant news articles, which could be supporting or not-supporting the claim. Researchers labeled the news Headline Stance based on whether the news article supported the claim or denied the claim, on a 5-item scale: -1 (strong deny), -0.5 (partially deny), 0 (neither support nor deny), 0.5 (partially support), 1 (strong support).

3.5.3 Pre- and Post-Perceived Correctness. Participants' perceived correctness regarding each claim was collected before (Pre-) and

[1]https://developer.mozilla.org/en-US/docs/Web/API/Element/getBoundingClientRect

after (Post-) they viewed each claim in the fact-checking interface (Section 3.4). Responses to these Pre- and Post-perceived Correctness were on a five-item scale ranging from false to true (denoted in lowercase).

3.5.4 Evidence correctness. Evidence correctness denotes the relationship between the headline stance and the claim correctness. If the news supports a TRUE claim or denies a FALSE claim, it is categorized as correct evidence. In contrast, if the news denies a TRUE claim or supports a FALSE claim, it is categorized as false evidence. In this paper, we consider only those news articles that fully supported or fully denied a claim.

- **correct evidence**: headline stance is -1 (strong deny) and claim correctness is FALSE, or, headline stance is 1 (strong support) and claim correctness is TRUE.
- **incorrect evidence**: headline stance is 1 (strong support) and claim correctness is FALSE, or, headline stance is -1 (strong deny) and claim correctness is TRUE.

3.5.5 Headline-Belief-Consistency.

- **headline-belief-consistent**: headline stance is -1 (strong deny) and Pre-Perceived Correctness is false or probably-false; or, headline stance is 1 (strong support) and Pre-Perceived Correctness is true or probably-true.
- **headline-belief-inconsistent**: headline stance is 1 (strong support) and Pre-Perceived Correctness is false or probably false; or, headline stance is -1 (strong deny) and Pre-Perceived Correctness is true or probably true.

3.5.6 Belief Change. We measured participants' beliefs before and after they checked each claim in the fact-checking system. We grouped their belief change into five categories based on their Pre- and Post-Perceived Correctness, and the Claim Correctness:

- **stay-right**: claim correctness is TRUE and pre-trial and post-trial perceived correctness are both true or probably true; or, claim correctness is FALSE and pre-trial and post-trial perceived correctness are both false or probably false.
- **to-right**: claim correctness is TRUE and post-trial perceived correctness is more towards true than pre-trial; or, claim correctness is FALSE and post-trial perceived correctness is more towards false than pre-trial.
- **stay-neutral**: pre-trial and post-trial perceived correctness are both neutral.
- **to-wrong**: claim correctness is TRUE and post-trial perceived correctness is more towards false than pre-trial; or, claim correctness is FALSE and post-trial perceived correctness is more towards true than pre-trial.
- **stay-wrong**: claim correctness is TRUE and pre-trial and post-trial perceived correctness are both false or probably false; or, claim correctness is FALSE and pre-trial and post-trial perceived correctness are both true or probably true.

3.5.7 Cognitive Load: The Low/High Index of Pupillary Activity (LHIPA). We attempted to apply the LHIPA metric to process pupil dilation data. Pupil dilation during blinks was replaced by 0 according to the eye-tracker detection. After the pre-processing step, we then computed the LHIPA on the raw pupil diameter signal for each headline AOI visits. We observed that the LHIPA value changed significantly when the visit duration was around 1.67s and 6.67s. Because of the high variability associated with the AOI visit duration, the LHIPA metric was not an appropriate indicator of cognitive load, and was not applicable to compare the pupil pupil dilation when processing various news headlines.

3.5.8 Cognitive Load: Relative Pupil Dilation (RPD). We calculated pupil dilation based on the raw, high-resolution pupil data recorded at 300Hz. To eliminate individual variability in pupil sizes, we calculated a relative change in pupil diameter from a baseline for each participant. We first excluded the low-quality data (ET-Validity = 4) and the blink data (Blink detected (binary)=1) based on the blink detection algorithm implemented in our eye-tracking software. Then we took an average pupil size over all the experimental trials as the pupil diameter baseline ($P^i_{baseline}$) and calculated the relative change in pupil diameter (RPD^i_t) from each pupil measurement (Eq. 1) [Gwizdka et al. 2017; Wang et al. 2021]. We removed data records with diameters that exceeded ±3 SDs of the participants' total session average.

$$RPD^i_t = \frac{P_t - P^i_{baseline}}{P^i_{baseline}} \tag{1}$$

To calculate the *RPD* for each AOI, we first downsampled the *RPD* to 50Hz using a median filter to minimize the influence of the outliers. Then we excluded the RPD within 0.5 seconds after the interface visit started to reduce the influence of the variability of luminance across the web pages. We assumed that two fixations on a headline did not represent reading and, accordingly, kept the AOI visit which had more than 2 fixations and calculated the *RPD* median for all the AOI visits to a single AOI in each trial.

4 RESULTS

4.1 Testing assumptions

All assumptions were checked according to the type of statistical testing in this paper. Normality test was conducted before conducting t-test and ANOVA analysis. Bartlett's test was conducted to check for sphericity. The results of these tests indicated that no assumptions were violated.

4.2 AOI position

Figure 3 shows that relative pupil dilation (*RPD*) was largest when participants were reading the news headlines in the first row. Then the *RPD* decreased as they read the headlines in the following rows. A one-way ANOVA showed that the effect of headline position (i.e. the rank of headline in the interface) was significant, $F(4, 3004) = 321.9, p < .05$. A post hoc Tukey's HSD test showed that all groups differed significantly at $p < .05$.

4.3 Claim correctness and headline stance

Figure 4(a) indicates that larger *RPD* was on the news headline AOIs that denied the claims. *RPD* was generally larger when checking TRUE claims compared to FALSE claims. A two-way ANOVA was conducted to examine the effects of headline stance and claim correctness on *RPD*. Both claim correctness, $F(1,34)=15.54$, $p<.05$, and headline stance, $F(1,34)=31.98$, $p<.05$, had significant main effects on *RPD*. However, the interaction effects were not significant. A post

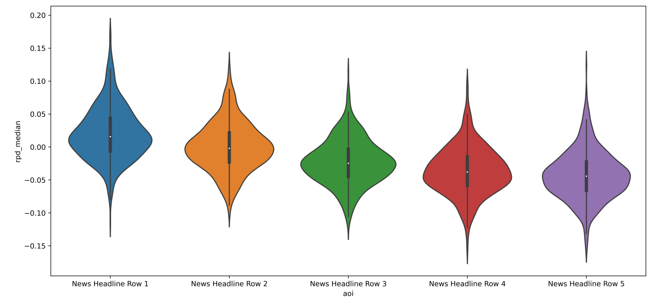

Figure 3: Relative pupil dilation (*RPD*) in the areas of interest (AOIs). From the left to the right it represents the AOIs of the news headlines that are from the first row to the fifth row of the fact-checking interface.

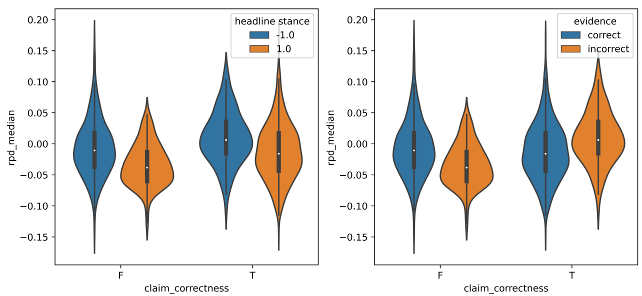

Figure 4: (a) Distribution of *RPD* of the news headline AOIs as a function of headline stance (-1: headline denies the claim; 1: headline supports the claim) and claim correctness (TRUE or FALSE claim). (b) Distribution of *RPD* of the news headline AOIs as a function of the evidence correctness (correct or incorrect) and claim correctness (TRUE or FALSE claim).

hoc Tukey's HSD test showed that all groups differed significantly at p<.05. The *RPD* was larger when participants were reading the news headlines that denied the claim, in both TRUE and FALSE claim group. The *RPD* was larger when participants were checking TRUE claims, no matter if they were reading news headlines denying or supporting a claim.

Figure 4(b) illustrates that *RPD* was larger when participants were checking the incorrect evidence for TRUE claims, and when checking the correct evidence for FALSE claims. *RPD* was tested by a two-way ANOVA with two levels of claim correctness (TRUE, FALSE) and two levels of evidence correctness (correct, incorrect). The main effect of the claim correctness was significant, F(1,34)=15.54, p<.05. The main effect of the evidence correctness was not significant. However, the interaction of claim correctness and evidence correctness was significant, F(1,34)=31.98, p<.05. A post hoc Tukey's HSD test showed that all the groups differed significantly at p<.05. When participants were checking TRUE claims, the *RPD* was larger in the incorrect evidence group compared to the correct evidence group, while checking FALSE claims, the *RPD* was lower in the incorrect evidence group compared to the correct evidence group. When participants were reading correct evidence, the *RPD* was larger in the FALSE claim group compared to the TRUE claim group, while when participants were reading incorrect evidence, the *RPD* was larger in the TRUE claim group compared to the FALSE claim group.

4.4 Prior belief and headline stance

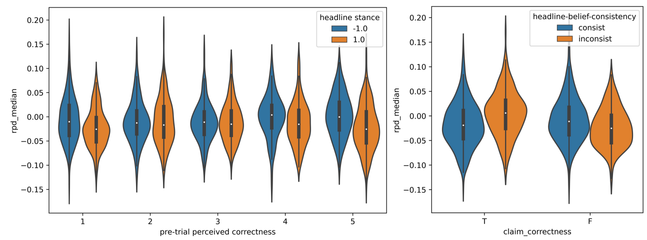

Figure 5: (a) Distribution of *RPD* on the news headline AOIs as a function of the perceived correctness before examining the claim in the system (1 to 5: False to True) and the headline stance (-1: headline denies the claim; 1: headline supports the claim). (b) Distribution of *RPD* of the news headline AOIs as a function of the headline-belief-consistency (consistent or inconsistent) and claim correctness (TRUE or FALSE claim).

Figure 5(a) shows that the largest difference in *RPD* between the news headlines that supported and those that denied the claim was when participants' prior belief was true or false. That difference was smaller when their prior belief was neutral. Therefore we looked further into the relationship between headline stance and the prior belief in two directions (i.e., the perceived correctness of the claim was either true or false). Figure 5(b) shows that the *RPD* was higher when headline-belief was inconsistent in TRUE claim groups and when headline-belief was consistent in FALSE claim groups. A two-way ANOVA was conducted to examine the effects of the headline-belief consistency and the claim correctness on *RPD*. The claim correctness had a significant main effect, F(1,35)=8.42, p<.05, while the headline-belief consistency had no significant effect. The interaction effects of claim correctness and headline-belief consistency were significant, F(1,35)=23.31, p<.05. A post hoc Tukey's HSD test showed that all the groups differed significantly at p<.05. When checking TRUE claims, the *RPD* was larger in the headline-belief inconsistent group compared to the headline-belief consistent group, while when checking FALSE claims, the *RPD* was lower in the headline-belief inconsistent group compared to the headline-belief consistent group. When headline-belief was consistent, the *RPD* was larger in the FALSE claim group compared to the TRUE claim group, while when headline-belief was inconsistent, the *RPD* was larger in the TRUE claim group compared to the FALSE claim group.

4.5 Belief change

A one-way ANOVA ($F(4, 2175) = 0.61$) indicated that the *RPD* was not significantly different between belief change conditions. Therefore, the change of the belief did not significantly influence the *RPD*. In our lab experiment, participants maintained their correct beliefs (stay-right) in 44.22% of the trials, and corrected their beliefs (to-right) in 46.56% of the trials. Only in 9.22% of the trials,

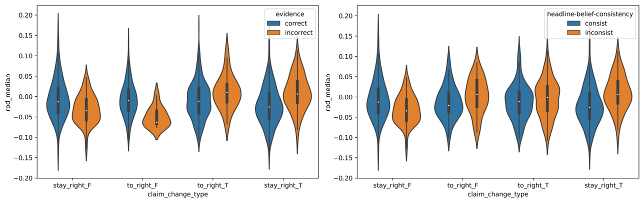

Figure 6: (a) Distribution of *RPD* of the news headline AOIs as a function of the evidence correctness (correct or incorrect) and the belief change in both claims (stay/to-right in TRUE/FALSE claims). (b) Distribution of *RPD* of the news headline AOIs as a function of belief-headline-consistency (consistent or inconsistent) and the belief change in both claims (stay/to-right in TRUE/FALSE claims).

participants stayed neutral, or remained incorrect (stay-wrong), or changed their beliefs to incorrect (to-wrong).

Furthermore, we checked the impact of the evidence correctness and the headline-belief-consistency on the *RPD* of the participants whose beliefs were corrected (to-right) or remained correct (stay-right). Figure 6(a) shows that *RPD* differed between reading the headlines that are correct evidence and incorrect evidence within each belief change and claim correctness combination group (i.e., to-right in TRUE claims, to-right in FALSE claims, stay-right in TRUE claims, stay-right in FALSE claims). *RPD* when reading correct evidence had small differences between belief change and claim correctness combination groups, while *RPD* when reading incorrect evidence had larger differences between belief change and claim correctness combination groups. when reading the incorrect evidence, the *RPD* differed more between change trend groups. A paired-sample t-test was conducted to compare *RPD* correct evidence and incorrect evidence conditions within each belief change and claim correctness combination group. There were significant differences in *RPD* between reading correct and incorrect evidence in all the belief change groups: stay-right for FALSE claims, $t(28) = 4.44, p < .05$, to-right for FALSE claims, $t(8) = 3.97, p < .05$, to-right for TRUE claims, $t(13) = -2.36, p < .05$, stay-right for TRUE claims, $t(29) = -5.25, p < .05$.

Table 1: Mann-Whitney U tests to determine whether there were significant differences between maintaining correct beliefs (stay-right) vs. correcting beliefs (to-right) (Section 4.5), (a) for different combinations of Claim-Correctness and Evidence Correctness, and (b) for different combinations of Claim-Correctness and Headline-belief Consistency. ($^*p < .05$, $^{**}p < .01$, $^{***}p < .001$)

Claim Correctness	Evidence Correctness	U	P
TRUE	Correct	24077	0.003**
TRUE	Incorrect	1224	0.960
FALSE	Correct	41633	0.991
FALSE	Incorrect	458	0.043*

(a)

Claim Correctness	Headline-belief Consistency	U	P
TRUE	Consistent	9960	0.042*
TRUE	Inconsistent	7280	0.056
FALSE	Consistent	20406	0.185
FALSE	Inconsistent	845	<0.001***

(b)

Figure 6(b) shows that the differences between the headline-belief consistent group and inconsistent group. *RPD* in headline-belief consistent group were larger than headline-belief inconsistent group when participants maintained their correct beliefs (stay-right) in FALSE claims. *RPD* in headline-belief inconsistent group were larger than headline-belief consistent group when participants changed to the correct belief (to-right) in both TRUE and FALSE claims, and when participants maintained their correct beliefs (stay-right) in TRUE claims. A paired-sample t-test was conducted to compare *RPD* in headline-belief consistent and inconsistent groups within each belief change and claim correctness combination group. There were significant differences in *RPD* in two of the belief change groups: stay-right for FALSE claims, $t(28) = 4.44, p < .05$, to-right in TRUE claims, $t(13) = -5, 25, p < .05$. A Mann-Whitney U test was conducted to determine whether there was a difference when the participants' beliefs stayed right and moved to the right in each evidence correctness or headline-belief-consistency groups and claim groups. Table 1 demonstrates the results. The Mann-Whitney U test and paired-samples t-test indicated that *RPD* was significantly larger when the participants corrected their beliefs (to-right) especially in the headline-belief consistent group for TRUE claims, and in the headline-belief inconsistent group for FALSE claims.

5 DISCUSSION

In this study, we investigated how the cognitive load is impacted in the fact-checking context and if it is related to users' belief change. We conducted a within-subject, lab-based, quasi-experiment, in which we manipulated the evidence correctness (correct evidence, incorrect evidence), the headline-belief-consistency (consistent, inconsistent), and measured participants' belief change (stay-right, stay-neutral, stay-wrong, to-right, to-wrong). We evaluated the cognitive load when participants read the news headlines by measuring pupil dilation on the headline AOIs.

We found that *RPD* is the highest when users read the news headlines in the top row, and that the *RPD* decreases on the lower positioned headline rows. This suggested that the cognitive load is higher when people are processing top headlines. This could be because of the position bias [Azzopardi 2021], where highly ranked results tend to attract more attention, reading and clicks. Since the news headlines with different stances were randomly assigned to the headline row positions, the position bias does not influence our hypothesis testing of other factors potentially impacting cognitive load.

Evidence correctness. We found that higher cognitive load is imposed when reading news headlines that are denying the claim. That is, higher cognitive load is required when users read incorrect evidence for TRUE claims and correct evidence for FALSE claims. This finding supports H1 when users are fact-checking a TRUE claim but provides no support for H1 when users are fact-checking a FALSE claim. Additionally, when users are checking TRUE claims, higher cognitive load is imposed regardless of the evidence correctness.

Headline-belief-consistency. When checking the relationship between headline stance and prior beliefs, the results indicate that cognitive load differs more between supportive and unsupportive headlines when the users perceived the claim correctness as true or false, instead of being neutral. The observation could be explained by that users would like to check all the information and a similar amount of cognitive load is imposed to read both supportive and unsupportive news since they do not have a prior tendency to claim correctness. Furthermore, we looked into the association between cognitive load and the news headlines when the users held non-neutral prior beliefs. We found evidence for the effect of headline-belief consistency on cognitive load. Reading headline-belief inconsistent news imposed higher cognitive load when checking TRUE claims, while reading headline-belief consistent news imposed higher cognitive load when checking FALSE claims. This finding supports H2 for TRUE claims but denies H2 for FALSE claims.

Belief change. We did not find a relationship between cognitive load and belief change, which denies H3. This result is not aligned with previous findings on discerning misinformation performance [Mirhoseini et al. 2022]. It is possible that our findings are due to the participants using the fact-checking system in the experiment and being more aware of the tasks to discern misinformation, hence they invest similar mental effort to process all news headlines shown in the interface and to judge the claim correctness. Additionally, the belief change results suggest that participants generally performed well on the fact-checking tasks. Among all the experiment trials, users' kept their correct beliefs (stay-right) or corrected them (to-right) in more than 90% trials. This indicates that the fact-checking system helped users to discern misinformation. In the analysis within the stay-right and to-right belief change groups, the results cross-validated the findings in evidence correctness and headline-belief-consistency - H1 and H2 are supported for TRUE claims and rejected for FALSE claims. Moreover, the cognitive load was higher when users' corrected their beliefs (to-right) compared to when they maintained their beliefs (stayed-right) when users were reading headline-belief-consistent news for TRUE claims and headline-belief-inconsistent for FALSE claims. This implies that checking news headlines when users' beliefs were corrected (to-right) imposed higher cognitive load than checking news headlines when they maintained their correct beliefs (stay-right).

In summary, when users were reading news headlines for TRUE claims, our proposed hypotheses H1 and H2 were supported, while when users were reading news headlines for FALSE news, H1 and H2 were not supported. There was not enough evidence in our study to support H3. We found that incorrect evidence and headline-belief inconsistency may not always impose higher cognitive load. Instead, the cognitive load level imposed by reading headlines appeared to be associated with the claim correctness. When checking TRUE claims, higher cognitive load was imposed when users read incorrect evidence or read headline-belief-inconsistent news. When checking FALSE claims, higher cognitive load was imposed when users read correct evidence or read headline-belief-consistent news. The findings plausibly indicated that people tended to engage more with the news they believed in when they were checking FALSE claims, while they engaged more with the news that countered their belief when they were checking TRUE claims. We also found that cognitive load did not significantly differ between belief change conditions, which suggests that the fact-checking tasks imposed similar level of cognitive load regardless how people's beliefs changed. However, the results indicate that a higher cognitive load was imposed when users corrected their beliefs and when users were reading headline-belief-consistent news for TRUE claims, or headline-belief-inconsistent for FALSE claims.

This research develops an understanding of cognitive load in discerning misinformation in realistic scenarios. Based on previous research on the association between cognitive load and reading and identifying misinformation [Mirhoseini et al. 2022], we studied how people process misinformation when they encounter several headlines (or news search results) at the same time while examining a single claim. Our findings suggest that cognitive load is imposed differently when checking true claims versus false claims. Previous research suggests that information system should encourage people to engage more cognitive effort (System 2), which could help them to identify misinformation more effectively [Rieger et al. 2021]. Meanwhile, we need to prevent the cognition to be overloaded which could drive the users back to utilize their heuristic [Whelan et al. 2020]. Our study implies that there are nuances in cognitive load when people are processing information with different claim correctness, evidence correctness, and headline-belief-consistency. As suggested in [Littrell et al. 2022], different kinds of misinformation could invoke different information behavior. In practical system design, we should not simply increase or decrease the cognitive load, but instead seek to calibrate the cognitive load with respect to the information context. We need to adopt a more nuanced approach to nudge people to discern misinformation, such as providing personalized labels or explanations to remind people to pay attention to the misinformation at the appropriate condition.

Our study has some limitations. We only observed the effect of evidence correctness, headline-stance-consistency, and belief change on the cognitive load when reading the news headlines. The cognitive load could also be impacted by users' familiarity and the knowledge level of the claim topic. Even though we have excluded the participants with expert topic familiarity of the content based on self-reported information, there's a possibility that people are not aware of their expertise in the topic. Since higher familiarity could impose lower the cognitive load [Jen-Hwa Hu et al. 2017], this limitation would impact the internal validity of our research. Additionally, the claims and news headlines were pre-selected to conduct the controlled within-subject experiment. Future work should include using sets of claims on different topics and investigating cognitive load in the context of naturally generated fact-checking tasks. Another limitation of this study is that we only measured pupil dilation when they were looking at the news headline AOIs. It would therefore be interesting to measure pupil dilation when they read the full news articles and compare the cognitive load variations between distinct news conditions. Lastly, the eye-tracking sequences in the experiment are relatively short. This renders them inapplicable to use the LHIPA technique to process and analyze the pupil dilation. Future work could improve the experimental design and allow for other pupillary response measurements (i.e., LHIPA) to reflect cognitive load with higher accuracy [Duchowski et al. 2020], or even other physiological measures, such as Electroencephalography (EEG) [Antonenko et al. 2010].

6 CONCLUSION

We presented results from a within-subject, lab-based, quasi-experiment with eye-tracking in which we examined how cognitive load is impacted by reading news headlines in a fact-checking context (i.e., by the evidence correctness and users' prior beliefs), and how it is related to people's belief change and their misinformation judgment. We found that incorrect evidence and headline-belief inconsistency imposed higher cognitive load when people were checking true claims, while correct evidence and headline-belief consistency imposed higher cognitive load when people were checking false claims. Additionally, cognitive load was not significantly different when people's beliefs changed. By developing the understanding of the cognition in discerning misinformation in a realistic scenario, the findings contribute to designing future information systems that support curbing of misinformation spread via appropriate technical and cognitive interventions.

ACKNOWLEDGMENTS

This research was completed under UT Austin IRB study 2017070049 and supported in part by Wipro, the Micron Foundation, the Knight Foundation, and by Good Systems[2] , a UT Austin Grand Challenge to develop responsible AI technologies. The statements made herein are solely the opinions of the authors and do not reflect the views of the sponsoring agencies.

REFERENCES

Pavlo Antonenko, Fred Paas, Roland Grabner, and Tamara Van Gog. 2010. Using electroencephalography to measure cognitive load. *Educational psychology review* 22, 4 (2010), 425–438.

Phoebe Arnold. 2020. *The challenges of online fact checking*. Technical Report December. Full Fact.

Leif Azzopardi. 2021. Cognitive biases in search: a review and reflection of cognitive biases in Information Retrieval. In *Proceedings of the 2021 conference on human information interaction and retrieval*. 27–37.

Bence Bago, David G Rand, and Gordon Pennycook. 2020. Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of experimental psychology: general* 149, 8 (2020), 1608.

Serena Chen, Kimberly Duckworth, and Shelly Chaiken. 1999. Motivated heuristic and systematic processing. *Psychological Inquiry* 10, 1 (1999), 44–49.

Anubrata Das, Houjiang Liu, Venelin Kovatchev, and Matthew Lease. 2023. The state of human-centered NLP technology for fact-checking. *Information Processing & Management* 60, 2 (2023), 103219. https://doi.org/10.1016/j.ipm.2022.103219

Anubrata Das, Kunjan Mehta, and Matthew Lease. 2019. CobWeb: a research prototype for exploring user bias in political fact-checking. *arXiv preprint arXiv:1907.03718* (2019).

Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H Eugene Stanley, and Walter Quattrociocchi. 2016. The spreading of misinformation online. *Proceedings of the National Academy of Sciences* 113, 3 (2016), 554–559.

Andrew T Duchowski, Krzysztof Krejtz, Nina A Gehrer, Tanya Bafna, and Per Bækgaard. 2020. The low/high index of pupillary activity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

Andrew T Duchowski, Krzysztof Krejtz, Izabela Krejtz, Cezary Biele, Anna Niedzielska, Peter Kiefer, Martin Raubal, and Ioannis Giannopoulos. 2018. The index of pupillary activity: Measuring cognitive load vis-à-vis task difficulty with pupil oscillation. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–13.

Ullrich KH Ecker, Stephan Lewandowsky, John Cook, Philipp Schmid, Lisa K Fazio, Nadia Brashier, Panayiota Kendeou, Emily K Vraga, and Michelle A Amazeen. 2022. The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology* 1, 1 (2022), 13–29.

D Graves. 2018. Understanding the promise and limits of automated fact-checking. (2018).

Lucas Graves and Michelle A Amazeen. 2019. Fact-checking as idea and practice in journalism. In *Oxford Research Encyclopedia of Communication*.

[2]http://goodsystems.utexas.edu/

Maurício Gruppi, Benjamin D. Horne, and Sibel Adalı. 2020. NELA-GT-2019: A Large Multi-Labelled News Dataset for The Study of Misinformation in News Articles. *arXiv:2003.08444 [cs]* (March 2020). http://arxiv.org/abs/2003.08444 arXiv: 2003.08444.

Zhijiang Guo, Michael Schlichtkrull, and Andreas Vlachos. 2022. A Survey on Automated Fact-Checking. *Transactions of the Association for Computational Linguistics* 10 (02 2022), 178–206. https://doi.org/10.1162/tacl_a_00454

Jacek Gwizdka. 2014. Characterizing relevance with eye-tracking measures. In *Proceedings of the 5th information interaction in context symposium*. 58–67.

Jacek Gwizdka. 2021. "Overloading" Cognitive (Work) Load: What Are We Really Measuring? *NeuroIS Retreat* (2021), 77–89.

Jacek Gwizdka, Rahilsadat Hosseini, Michael Cole, and Shouyi Wang. 2017. Temporal dynamics of eye-tracking and EEG during reading and relevance decisions. *Journal of the Association for Information Science and Technology* 68, 10 (2017), 2299–2312.

Christian Hansen, Casper Hansen, Jakob Grue Simonsen, Birger Larsen, Stephen Alstrup, and Christina Lioma. 2020. Factuality checking in news headlines with eye tracking. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2013–2016.

Sandra G Hart. 2006. NASA-task load index (NASA-TLX); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 50. Sage publications Sage CA: Los Angeles, CA, 904–908.

Eckhard H Hess and James M Polt. 1964. Pupil size in relation to mental activity during simple problem-solving. *Science* 143, 3611 (1964), 1190–1192.

Jennifer L Hochschild and Katherine Levine Einstein. 2015. *Do facts matter?: Information and misinformation in American politics*. Vol. 13. University of Oklahoma Press.

Gahangir Hossain and Mohammed Yeasin. 2014. Understanding effects of cognitive load from pupillary responses using hilbert analytic phase. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 375–380.

Paul Jen-Hwa Hu, Hu Han-fen, and Fang Xiao. 2017. Examining the Mediating Roles of Cognitive Load and Performance Outcomes in User Satisfaction with a Website: A Field Quasi-Experiment. *MIS Quarterly* 41, 3 (2017).

Shan Jiang, Simon Baumgartner, Abe Ittycheriah, and Cong Yu. 2020. Factoring fact-checks: Structured information extraction from fact-checking articles. In *Proceedings of The Web Conference 2020*. 1592–1603.

Daniel Kahneman. 2011. *Thinking, fast and slow*. Macmillan.

Daniel Kahneman and Jackson Beatty. 1966. Pupil diameter and load on memory. *Science* 154, 3756 (1966), 1583–1585.

Krzysztof Krejtz, Andrew T Duchowski, Anna Niedzielska, Cezary Biele, and Izabela Krejtz. 2018. Eye tracking cognitive load using pupil diameter and microsaccades with fixed gaze. *PloS one* 13, 9 (2018), e0203629.

Jan-Louis Kruger, Esté Hefer, and Gordon Matthew. 2013. Measuring the impact of subtitles on cognitive load: Eye tracking and dynamic audiovisual texts. In *Proceedings of the 2013 Conference on Eye Tracking South Africa*. 62–66.

Wagner Junior Ladeira, Marlon Dalmoro, Fernando de Oliveira Santini, and William Carvalho Jardim. 2022. Visual cognition of fake news: the effects of consumer brand engagement. *Journal of Marketing Communications* 28, 6 (2022), 681–701.

Shane Littrell, Ethan Andrew Meyers, and Jonathan Albert Fugelsang. 2022. Not all bullshit pondered is tossed: Reflection decreases receptivity to some types of misinformation but not others. (2022).

Sandra P Marshall. 2002. The index of cognitive activity: Measuring cognitive workload. In *Proceedings of the IEEE 7th conference on Human Factors and Power Plants*. IEEE, 7–7.

Mahdi Mirhoseini, Spencer Early, and Khaled Hassanein. 2022. All eyes on misinformation and social media consumption: A pupil dilation study. In *NeuroIS Retreat 2022*.

Sina Mohseni, Fan Yang, Shiva K Pentyala, Mengnan Du, Yi Liu, Nic Lupfer, Xia Hu, Shuiwang Ji, and Eric D Ragan. 2021. Machine Learning Explanations to Prevent Overtrust in Fake News Detection.. In *ICWSM*. 421–431.

Patricia L Moravec, Antino Kim, and Alan 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 (2020), 987–1006.

Preslav Nakov, D. Corney, Maram Hasanain, Firoj Alam, Tamer Elsayed, Alberto Barrón-Cedeño, Paolo Papotti, Shaden Shaar, and Giovanni Da San Martino. 2021. Automated Fact-Checking for Assisting Human Fact-Checkers. In *IJCAI*.

An T Nguyen, Aditya Kharosekar, Saumyaa Krishnan, Siddhesh Krishnan, Elizabeth Tate, Byron C Wallace, and Matthew Lease. 2018. Believe it or not: Designing a human-ai partnership for mixed-initiative fact-checking. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. 189–199.

Jeppe Nørregaard, Benjamin D Horne, and Sibel Adalı. 2019. NELA-GT-2018: A large multi-labelled news dataset for the study of misinformation in news articles. In *Proceedings of the international AAAI conference on web and social media*, Vol. 13. 630–638.

Robert D O'DONNELL. 1986. Workload assessment methodology. *Cognitive processes and performance* (1986).

Fred Paas, Juhani E Tuovinen, Huib Tabbers, and Pascal WM Van Gerven. 2016. Cognitive load measurement as a means to advance cognitive load theory. In *Educational psychologist*. Routledge, 63–71.

Oskar Palinko, Andrew L Kun, Alexander Shyrokov, and Peter Heeman. 2010. Estimating cognitive load using remote eye tracking in a driving simulator. In *Proceedings of the 2010 symposium on eye-tracking research & applications*. 141–144.

Gordon Pennycook and David G Rand. 2019. Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition* 188 (2019), 39–50.

Alisa Rieger, Tim Draws, Mariët Theune, and Nava Tintarev. 2021. This Item Might Reinforce Your Opinion: Obfuscation and Labeling of Search Results to Mitigate Confirmation Bias. In *Proceedings of the 32nd ACM Conference on Hypertext and Social Media*. 189–199.

Jon Roozenbeek, Claudia R Schneider, Sarah Dryhurst, John Kerr, Alexandra LJ Freeman, Gabriel Recchia, Anne Marthe Van Der Bles, and Sander Van Der Linden. 2020. Susceptibility to misinformation about COVID-19 around the world. *Royal Society open science* 7, 10 (2020), 201199.

Li Shi, Nilavra Bhattacharya, Anubrata Das, Matthew Lease, and Jacek Gwizdka. 2022. The Effects of Interactive AI Design on User Behavior: An Eye-tracking Study of Fact-checking COVID-19 Claims. In *Proceedings of the 7th ACM SIGIR Conference on Human Information, Interaction and Retrieval (CHIIR)*. https://doi.org/10.1145/3498366.3505786

Yu Shi, Natalie Ruiz, Ronnie Taib, Eric Choi, and Fang Chen. 2007. Galvanic skin response (GSR) as an index of cognitive load. In *CHI'07 extended abstracts on Human factors in computing systems*. 2651–2656.

Greg J Siegle, Naho Ichikawa, and Stuart Steinhauer. 2008. Blink before and after you think: Blinks occur prior to and following cognitive load indexed by pupillary responses. *Psychophysiology* 45, 5 (2008), 679–687.

Keith E Stanovich, Richard F West, and Maggie E Toplak. 2011. The complexity of developmental predictions from dual process models. *Developmental Review* 31, 2-3 (2011), 103–118.

Ömer Sümer, Efe Bozkir, Thomas Kübler, Sven Grüner, Sonja Utz, and Enkelejda Kasneci. 2021. FakeNewsPerception: An eye movement dataset on the perceived believability of news stories. *Data in brief* 35 (2021), 106909.

Nerea Urrestilla and David St-Onge. 2020. Measuring cognitive load: Heart-rate variability and pupillometry assessment. In *Companion Publication of the 2020 International Conference on Multimodal Interaction*. 405–410.

Yuexin Wang, Yining Guo, Jiajia Wang, Ziyuan Liu, and Xuemin Li. 2021. Pupillary response to moving stimuli of different speeds. *Journal of Eye Movement Research* 14, 1 (2021).

Eoin Whelan, AKM Najmul Islam, and Stoney Brooks. 2020. Applying the SOBC paradigm to explain how social media overload affects academic performance. *Computers & Education* 143 (2020), 103692.

Bo Xie, Daqing He, Tim Mercer, Youfa Wang, Dan Wu, Kenneth R Fleischmann, Yan Zhang, Linda H Yoder, Keri K Stephens, Michael Mackert, et al. 2020. Global health crises are also information crises: A call to action. *Journal of the Association for Information Science and Technology* 71, 12 (2020), 1419–1423.

John Zarocostas. 2020. How to fight an infodemic. *The lancet* 395, 10225 (2020), 676.