The Influence of Content Modality on Perceptions of Online Misinformation

Suwani Gunasekara
School of Computing and Information Systems
University of Melbourne
Melbourne, Victoria, Australia
gunasekara@student.unimelb.edu.au

Saumya Pareek
School of Computing and Information Systems
University of Melbourne
Melbourne, Victoria, Australia
spareek@student.unimelb.edu.au

Ryan M. Kelly
School of Computing Technologies
RMIT University
Melbourne, VIC, Australia
ryan.kelly@rmit.edu.au

Jorge Goncalves
School of Computing and Information Systems
University of Melbourne
Melbourne, Australia
jorge.goncalves@unimelb.edu.au

Abstract

Social media has become a primary information source, with platforms evolving from text-based to multi-modal environments that include images and videos. While richer media modalities enhance user engagement, they also increase the spread and perceived credibility of misinformation. Most interventions to counter misinformation on social media are text-based, which may lack the persuasive power of richer modalities. This study explores whether the effectiveness of misinformation correction varies by modality, and if certain modalities of misinformation are better countered by a specific correction modality. We conducted a survey-based experiment where participants rated the credibility of misinformation tweets before and after exposure to corrections, across all combinations of text, images and video modalities. Our findings suggest that corrections are most effective when their modality richness matches that of the original misinformation. We discuss factors affecting the perceived credibility of corrections and offer strategies to optimise misinformation correction.

CCS Concepts

• **Human-centered computing → Empirical studies in HCI**.

Keywords

Misinformation, Content Modality, Corrections, Social Media

ACM Reference Format:
Suwani Gunasekara, Saumya Pareek, Ryan M. Kelly, and Jorge Goncalves. 2025. The Influence of Content Modality on Perceptions of Online Misinformation. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3706598.3713098

1 Introduction

Social media platforms have become the predominant source of information for many individuals [54]. These platforms have evolved from being text-based to incorporate multiple modes of communication. For instance, *X* (formerly Twitter) began as a micro-blogging platform that allowed users to share short textual messages called "tweets". Tweets have now evolved to include multiple modalities beyond text, such as images, videos, and hyperlinks [90]. Undoubtedly, this diversification in media modality enriches user interaction and engagement on these online platforms. However, *richer media* can also facilitate the spread of misinformation and enhance the perceived credibility of misinformation [2, 77, 83].

Despite this, most interventions to counter misinformation on social media remain text-based [30]. However, text-based interventions may not be as engaging or persuasive as those involving richer media modalities, given their distinct lack of visual and auditory elements. Research has shown that videos, with their combination of visuals, sound, vivid details, and resemblance to real-life experiences, can be more compelling than text alone [74]. However, there is little research on how different modalities influence the effectiveness of misinformation correction. This raises key questions: Does the effectiveness of corrections to misinformation depend on the modality used? Is a particular modality of misinformation more effectively countered by a specific correction modality?

In this short paper, we address these questions by examining the impact of different misinformation and correction modalities. Specifically, we investigate whether certain modalities of misinformation are more effectively countered by particular correction modalities. We conduct our study within the context of X, due to its status as a primary tool for journalists and news seekers [52], and because it is used by millions of people for daily news consumption [38]. Despite its utility, X has been criticised for spreading misinformation, particularly under recent leadership changes [7, 37, 51]. By focusing on tweets in this study, we aim to understand how different corrective modalities can mitigate the spread of misinformation and enhance the ability of users to discern truth from misinformation. We conducted a survey-based experiment where participants viewed and rated the credibility of misinformation tweets before and after viewing correction tweets, both presented in three modalities: TEXT, IMAGE, and VIDEO. Our findings indicate that video corrections are most effective when delivered in the same video format as the misinformation. However, when addressing misinformation presented in a text format, providing

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713098

corrections in text form is more effective. We discuss these findings, provide recommendations for designing more effective corrections, and outline potential directions for future research.

2 Related Work

2.1 Media Richness Theory and Modes of Misinformation

Media Richness Theory (MRT) posits that communication media vary in their effectiveness at conveying information effectively [21]. Specifically, media capable of transmitting non-verbal cues enhance audience comprehension of the message more effectively than those that cannot. According to MRT, media types that provide a wider range of cues are classified as *rich media*, while those lacking these characteristics are classified as *lean media*. For instance, videos incorporate numerous visual and auditory cues, including body language, facial expressions, and tone of voice, which are absent in images or text. Consequently, rich media like videos can enhance the believability of a message, due to their ability to simulate real-life scenarios with greater clarity and detail [74].

MRT also offers insights into how misinformation spreads through social media. Importantly, the formats of rich media have diversified over time. For instance, when using images, information is now communicated not only through traditional photographs but also via memes and data visualisations [49, 62]. Furthermore, videos have expanded beyond conventional long segments to include short video clips that provide concise and captivating content, appealing to a wide range of viewers. However, the video modality also poses a greater risk for the dissemination of misinformation compared to other modalities. Research has shown that altered or manipulated video news segments are more readily believed and shared on social media, compared to their text or audio counterparts [75]. Given that modality plays a crucial role in the spread of misinformation, exploring how this very element can be effectively leveraged to counteract misinformation is key to developing more effective strategies for mitigating its spread.

2.2 Misinformation Interventions

Misinformation refers to the dissemination of false information, regardless of intent [5, 40]. Misinformation is spreading rapidly on social media platforms [3, 16], leading to negative emotions, disrupted platform use, and the facilitation of harmful online activities [12, 17]. Prior work in HCI has explored various approaches to combat online misinformation. This work includes designing interventions to warn users about harmful content through symbols [29] and labels [31, 35, 46, 69], removing [13, 69] or down-ranking misinformation content [27], and equipping users with tools to evaluate [8, 9, 15, 41, 43, 50, 82] and filter out [23] misinformation.

Another approach to combatting misinformation involves presenting people with factual content that challenges widespread false narratives. These narratives can be presented using text, image and/or video content. Given that the richness levels of these modalities vary, the believability of the message conveyed through each also differs. Text-based corrections are commonly used to counteract misinformation due to their ease of production and distribution. Some studies have indicated that text-based interventions can be effective in reducing misconceptions and misunderstandings among individuals [10, 81]. However, studies have also shown that text-based interventions can be ineffective [19, 76]. This mixed evidence may be attributed to the fact that text-based interventions often lack the persuasive appeal of richer media formats, such as video or audio, due to the absence of visual and auditory elements [66]. For instance, a study by Young et al. [91] found that fact-checking videos are more effective than long-form fact-checking articles in correcting beliefs and reducing misperceptions. Similarly, research indicates that images can significantly influence the credibility of information, as they provide visual cues that can make the content more believable [25].

Moreover, the effectiveness of correcting misinformation is influenced not only by the modality of the corrective information, but also by the modality in which the original misinformation was presented. The alignment between these modalities can impact how well the correction is received and internalised. Prior research has shown that when addressing misperceptions conveyed through text, presenting corrections in the same text modality can be an effective strategy [10, 33, 67, 85]. However, text-based corrections may not always be effective when addressing misinformation presented in other modalities. For instance, using text corrections to counter video misinformation has been found to be less effective than providing no corrective information at all [84].

While there has been some research on the effectiveness of different modalities in correcting misinformation, there remains a gap in examining the interaction of all major modalities simultaneously—an issue this study seeks to address. Previous research has also produced inconsistent results [10, 25, 32, 81, 89], and most studies focus on one or two modalities in isolation, making it difficult to draw broader conclusions. Furthermore, the impact of correction modality has been underexplored in prior HCI research. Studies have focused on aspects such as identifying the optimal timing for providing corrections [34], the impact of when a correction is provided by users [6, 56], optimizing the re-sharing of debunking messages [61], and examining the role and impact of fact-checkers and fact-checking in correcting misinformation [1, 18, 26, 45, 47, 68, 70]. Our work extends the HCI literature by investigating the relationship between content modality and correction modality, and how this might be considered in intervention design.

3 Methodology

3.1 Experimental Design

Our experiment followed a 3 (misinformation modality: **TextM**, **ImageM**, or **VideoM**) × 3 (correction modality: **TextC**, **ImageC**, or **VideoC**) factorial design. In our survey-based study, each participant encountered five pairs of tweets. The first tweet in each pair contained misinformation, and the second tweet contained a corresponding correction (the tweets were shown separately and in sequence, described further below). To minimise potential bias, identifiable information in the tweets, such as profile images and names, were blurred. To reduce potential order effects, tweet pairs were presented in a randomised order. The misinformation and correction tweets we used were adopted verbatim from Snopes' fact-checked tweets. Snopes is a reputable online fact-checking website, which offers detailed explanations on the accuracy of tweets by presenting evidence from sources with relevant expertise or

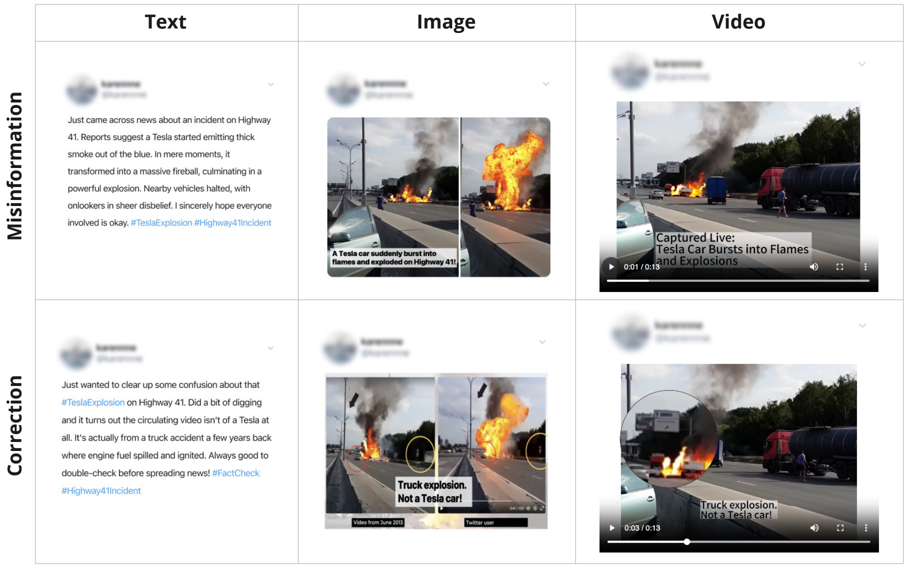

Figure 1: Examples of misinformation and corrections tweets used in the experiment.

knowledge [71]. We ensured a diverse selection of topics, including disasters, technology, and military. Examples of the used stimuli are provided in Figure 1.

3.2 Data Collection and Participants

The study was deployed on Prolific[1], following approval from the Human Ethics Committee at our university. We recruited participants who were fluent English speakers and had a platform approval rating of over 98%. Since the highest number of X users are from the United States [22], we pre-screened for participants living in the US. We used G*Power [28] to determine the required sample size. The minimum recommended sample size was 118 participants, considering an α = 0.05, and a power of 0.8. We conservatively recruited a total of 126 participants (63 men, 63 women, with an average age of 39.1) to cover all the misinformation and correction modality combinations in a balanced manner. Participants spent a median time of 13 minutes on the survey and received US$3 for their participation.

3.3 Measures

3.3.1 Dependent Variable. We aimed to understand the impact of different correction modalities on misinformation presented in three formats: **Text**, **Image**, and **Video**. To assess this, we measured the perceived credibility of the misinformation tweet both before and after participants were shown the correction tweet using a 7 point Likert scale (1 = Not Credible At All and 7 = Completely Credible). The dependent variable, **`Credibility Shift`**, was calculated by taking the difference between the credibility ratings post-correction (**`Final Credibility`**) and pre-correction (**`Initial Credibility`**).

[1]https://www.prolific.com/

3.3.2 Covariates/Dispositional measures. We measured several covariates to control for individual differences:

- **`Usage:`** Participants rated their frequency of using X for news updates. This was measured using a five-point Likert scale, ranging from 1 = Never to 5 = Always.
- **`Familiarity:`** Building on previous research that examined shifts in participants' beliefs [42, 60, 86–88], participants' prior knowledge of the news story was quantified by their indicated familiarity with the information presented in each misinformation tweet using a seven-point Likert scale (1 = Not Knowledgeable at all to 7 = Fully Knowledgeable).
- **`BSR:`** Following past literature in misinformation [64, 65, 78] we measured participants' receptivity to pseudo-profound statements using the Bullshit Receptivity scale (BSR) [63]. Participants rated the profoundness of 11 statements on a five-point Likert scale (1 = Not at all Profound to 5 = Very Profound). Our instructions told participants that "profound" refers to 'of deep meaning; of great and broadly inclusive significance' [63]. We hypothesise that participants with higher BSR scores will be more likely to believe misinformation.
- **`AOT:`** Following past literature in misinformation [53, 78], we measured the tendency to be open towards opinions or positions different from one's own using the Actively Open-Minded Thinking scale (AOT) [72]. Participants rated 13 statements on a six-point Likert scale (1 = Disagree Strongly to 6 = Agree Strongly). We hypothesise that participants with higher AOT scores will be more receptive to the corrections than those with lower AOT scores.

3.4 Procedure

Figure 2 illustrates the experimental procedure. Following a brief demographic questionnaire, participants rated their frequency of

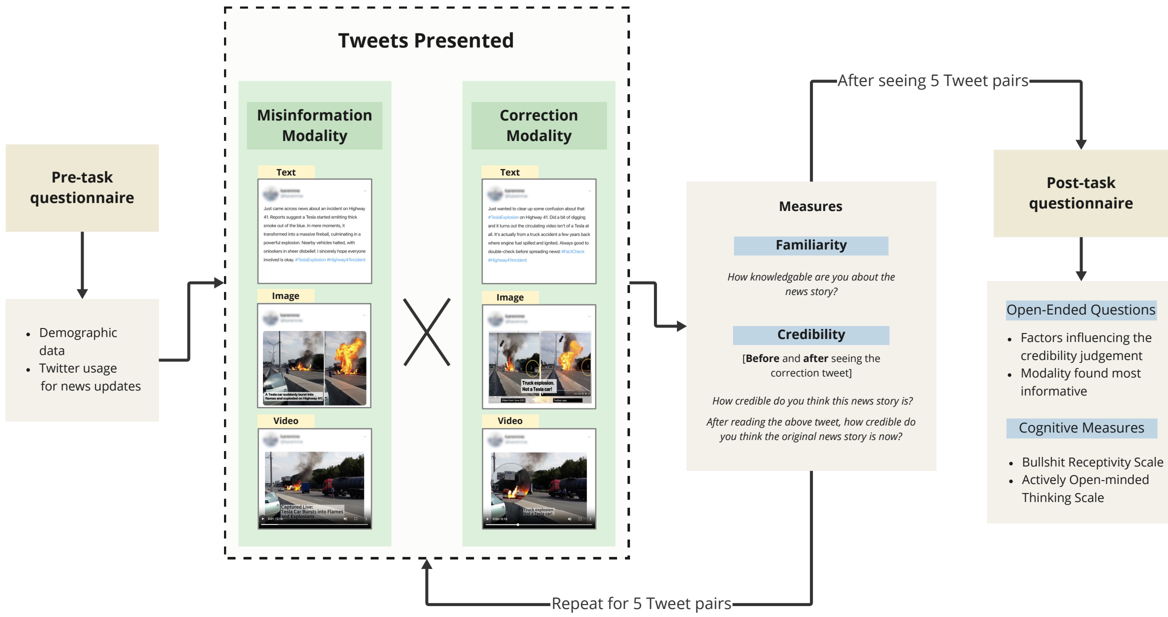

Figure 2: Experimental procedure used in the study.

X usage for news updates. Participants were then randomly assigned to one of three conditions, which determined the modality of misinformation tweets they encountered: either TEXT, IMAGE, or VIDEO. The misinformation modality was a between-subjects factor, meaning each participant saw all misinformation tweets presented in the same modality. In contrast, the correction modality (TEXT, IMAGE, or VIDEO) was a within-subjects factor, allowing participants to see corrections in all three modalities over the course of the experiment.

For each pair of tweets, participants first viewed the misinformation tweet and rated their familiarity with and perceived credibility of the tweet. Subsequently, they were shown the corresponding correction tweet (which was either text, image or video), and were asked to rate the perceived credibility of the original misinformation tweet again (see Figure 2). After viewing five pairs of tweets, each participant answered an open-ended question about factors that may have influenced their credibility judgements, and a second question about which modality they found the most informative. Participants then completed the survey by filling out the BSR and AOT questionnaires.

Two Instructional Manipulation Checks (IMC) were included in the online survey to ensure participant attentiveness [59]. The checks instructed participants to "Please write the word 'Purple' to answer this question" and respond to "Which of the following is a vegetable?" by selecting the correct answer from a list of options. These checks were programmed to appear randomly while rating the tweets. All participants successfully passed both IMCs. The used stimuli and the dataset are made available to the research community [2].

[2]https://osf.io/hvxyu/?view_only=56a4d5e5db394558add6b05628c7a7f4

4 Results

4.1 Quantitative Findings

4.1.1 Initial Credibility of Misinformation. Since values of **`Initial Credibility`** of the three modalities are independent of each other, are ordinal values as they are likert scale data and do not follow a normal distribution (Shapiro-Wilk Test returned $p < 0.05$ for all three modalities), and has a similar shape (Levene's Test for Equality of Variances returned $p > 0.05$), a Kruskal-Wallis test was performed [73]. The test revealed a statistically significant difference in **`Initial Credibility`** across the different misinformation modalities ($\chi^2 = 8.667$, $p < 0.05$), as illustrated in Figure 3. This suggests that the perceived initial credibility of misinformation varied depending on whether the content was presented in text, image, or video format. Since the Levene's test indicated that the data have equal variance ($p > 0.05$), we conducted the pairwise Wilcoxon Rank Sum Test, for post-hoc analysis. The test confirmed a statistically significant difference between VIDEO and TEXT modalities ($p < 0.05$), showing that misinformation presented in video format was perceived to be significantly more credible than the same content presented in text.

4.1.2 Credibility Shift. To investigate whether correction modality significantly affects the perceived credibility of misinformation, we employed a Generalised Linear Mixed Model (GLMM) using the R package `lme4`. We treated the change in misinformation credibility before and after the correction was displayed (**`Credibility Shift`**) as our outcome variable. For ease of interpretation, we multiplied the credibility shift value by -1, such that a positive estimate reflects a reduction in the perceived credibility of the misinformation following the correction (a desirable outcome), and a negative estimate indicates an increase in perceived credibility. We incorporated participant IDs and Tweet IDs as random effects in our model to

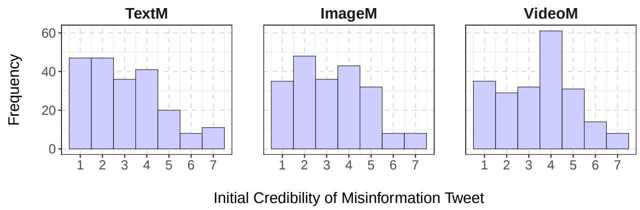

Figure 3: Distribution of Initial credibility ratings across different modality types. A rating of 1 indicates low perceived credibility, while a rating of 7 indicates high perceived credibility.

Table 1: Summary statistics of the GLMM model

Predictor	β (Std.Error)	p value
(Intercept)	1.825(2.058)	0.377
Age	-0.003(0.008)	0.691
meanBSR	0.030(0.113)	0.78
meanAOT	0.801(0.515)	0.123
Familiarity	0.149(0.054)	**0.005**
TextM	-1.001(0.332)	**0.003**
VideoM	0.130(0.331)	0.695
TextC	-0.166(0.279)	0.551
VideoC	-0.396(0.279)	0.156
TextM:TextC	0.732(0.392)	0.063
VideoM:TextC	0.157(0.394)	0.690
TextM:VideoC	0.571(0.394)	0.147
VideoM:VideoC	0.955(0.394)	**0.015**

account for individual differences and any variations among the five misinformation Tweets. The predictor variables in the GLMM model had no multicollinearity present, the random effects were normally distributed and independent of the response variable. Additionally, we conducted post-hoc analyses using pairwise comparisons to further examine the differences between correction modalities across various misinformation modalities. We present these pairwise comparisons along with the estimated marginal means (EMMs) for each condition to illustrate the effectiveness of different misinformation-correction pairs.

We observed a significant main effect of **`Familiarity`** *(β = 0.149, SE = 0.054, $p < 0.01$)* on **`Credibility Shift`**, indicating that participants who were more familiar with the news story showed a greater reduction in the perceived credibility of the misinformation after viewing the correction. In addition, we observed a significant main effect of **`Misinformation Modality`** *(β = -1.001, SE = 0.332, $p < 0.01$)* on **`Credibility Shift`**. Pairwise comparisons revealed a statistically significant difference between **Text** misinformation *(M = 0.73, SD = 0.201)* and **Video** misinformation *(M = 1.80, SD = 0.199)* on credibility shift *(β = -1.067, SE = 0.247, $p < 0.001$)*. This indicates that the mean credibility shift for video misinformation was significantly higher than that for text misinformation, suggesting that video-based misinformation is more amenable to correction compared to text-based misinformation.

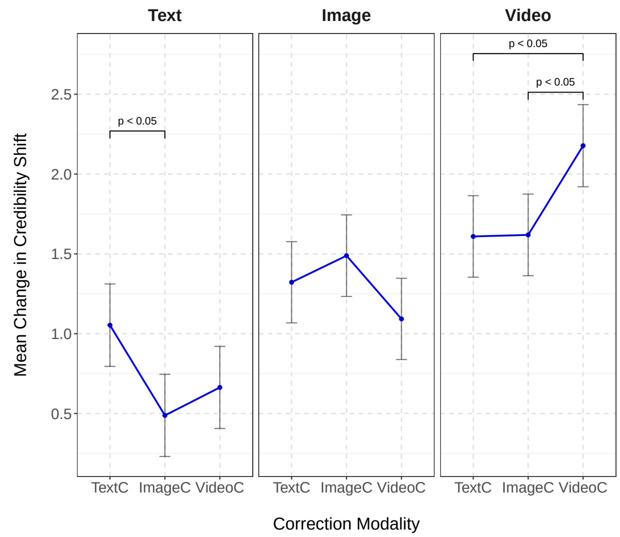

Figure 4: Pairwise comparison of the effectiveness of correction modalities (TextC, ImageC, VideoC) for each misinformation modality (TextM, ImageM, VideoM). Error bars denote Standard Error (SE).

Lastly, our findings revealed an interaction effect between the **`Misinformation Modality`** and **`Correction Modality`** *(β = 0.955, SE = 0.393, $p < 0.05$)*. The resultant *EMMs* from the post-hoc analysis are illustrated in Figure 4. We find the following:

- **For misinformation presented in Video format**, overall, Video corrections were the most effective modality for reducing the credibility of this misinformation, outperforming both Image and Text corrections.
  - We observed a significant difference between Video *(EMM = 2.178, SE = 0.257)* and Image *(EMM = 1.619, SE = 0.256)* corrections *(β = -0.559, SE = 0.279, p = 0.045)*, suggesting that for misinformation presented in video format, video corrections were more effective than image corrections in reducing perceived credibility.
  - Similarly, there was a significant difference between Video and Text *(EMM = 1.609, SE = 0.255)* corrections *(β = -0.568, SE = 0.279, p = 0.042)*, indicating that video corrections were also more effective than text corrections for video-based misinformation.
- **For misinformation presented in Image format**, no significant differences were found between the correction modalities in reducing the credibility of the misinformation. Specifically, pairwise comparisons between Image and Text corrections *(p = 0.551)*, Image and Video corrections *(p = 0.156)*, and Text and Video corrections *(p = 0.411)* all showed no significant effects on credibility shift.
- **For misinformation presented in Text format**, Text corrections were significantly more effective than Image corrections, but no significant differences were observed between the other correction modalities.

– We observed a significant difference between IMAGE *(EMM = 0.488, SE = 0.257)* and TEXT *(EMM = 1.322, SE = 0.254)* corrections *(β = -0.565, SE = 0.277, p = 0.042)*, suggesting that text corrections were more effective than image corrections in reducing the perceived credibility of text-based misinformation.
– However, no significant difference was observed between IMAGE and VIDEO corrections *(p = 0.530)*, nor between TEXT and VIDEO corrections *(p = 0.162)*.

4.2 Qualitative Findings

We analysed the qualitative responses to our open-ended survey questions to understand issues that affected participants' credibility judgements. The analysis aimed to examine how the modality of the misinformation content and their corrections affected participants' credibility judgements, and to explore factors beyond tweet modality that impacted these judgements. Following Braun and Clarke [11]'s inductive thematic analysis approach, two authors first familiarised themselves with the responses by reading through all responses multiple times. Initial codes were then generated, capturing key features of the data. These codes were collated into potential themes, and relevant participant responses (or parts thereof) were assigned to these themes. The themes were iteratively refined and revised.

4.2.1 The Influence of Correction Modality on Content Credibility. The majority of participants perceived **VIDEO** corrections as the most effective modality against misinformation, expressing several reasons for this preference. Participants highlighted how videos provided more context and clarity, and appeared credible because they provided *"more proof of what really happened"* (P43), allowing them to make an informed credibility judgement: *"[Video] was the most informative because I could see the correction in action and judge for myself."* (P58).

Videos were described as more attention-grabbing, and provided richer detail compared to text or images, which increased their persuasiveness: *"Video was the most informative because it had the most details, and [gave] much information while having your attention, compared to text or just images."* (P25). Participants found videos to be cognitively easier to parse: *"Videos were the most helpful, because they were easier to understand compared to text or images."* (P81). Some participants also expressed more trust in video-based corrections because they found them to be less susceptible to manipulation, and thus more credible: *"I found videos the most informative. [They] take significantly more effort to 'Photoshop' for the sake of creating a new narrative, so I find [them] easier to trust."* (P57). While some participants recognised that videos could be manipulated, they still preferred videos over other modalities: *"I think being shown something is always better than being told, although information can be faked in many different ways."* (P21).

Although less preferred than videos, many participants nevertheless found **TEXT** corrections to be effective in debunking misinformation. Those who expressed preference for text corrections highlighted that text seemed the most informative, as *"none of the images and videos provided enough for me to form conclusions from them alone."* (P7). Additionally, participants appreciated the concise and efficient nature of text-based corrections, which *"simply refute any claims and state facts, stripping away ambiguity and leaving little room for interpretation."* (P26). For certain cases, text was seen as more reliable and detailed compared to short-form videos that typically circulate on social media platforms: *"I think both text and video helped, but sometimes long text is more informative than a quick video."* (P94). Lastly, personal preference for information consumption played a role, as some participants found text corrections more suitable to their reading habits: *"I take information better if it's written, so I found the text corrections to be more informative."* (P96).

Despite theoretically being a richer modality than text, **IMAGE** corrections were preferred far less than video and text-based corrections. The few participants who did express a preference for images mentioned that they found them easier to read compared to videos: *"I liked when the corrections were in picture format instead of video because it was easier to read it all. Sometimes videos can move at too quick of a pace."* (P112). Some participants also appreciated the visual aspect of image-based corrections, much like participants stated for videos: *"[I prefer] pictures and videos because you have a visual of what is being reported."* (P67).

4.2.2 Factors Other than Modality that Influenced Content Credibility. We also undertook an exploratory analysis to identify factors other than correction modality that participants identified as also impacting their perception of content credibility:

- **Effect of Platform:** Participants' general scepticism towards social media, particularly X, impacted their credibility judgements: *"I tend to mistrust a lot of what I see on social media, especially Twitter after it was bought out by Musk."* (P46), and: *"I'm generally sceptical by nature, but doubly so when it comes to social media."* (P109).
- **Content Source:** Although we removed sources from our Tweet stimuli to reduce potentially confounding effects, some participants nevertheless highlighted that in general, it is important to know the source of the tweet when judging credibility: *"Normally I look at who is tweeting and if it is a credible source. Since [no source] was present, I had no way of knowing who was tweeting these things. Anyone can make up anything."* (P78).

5 Discussion

5.1 Rich Media is not Always the Best Approach to Combat Online Misinformation

Our findings indicate that misinformation presented in the **VIDEO** modality is perceived as more credible than misinformation presented in other formats, aligning with existing research on the effectiveness of richer media in convincing and persuading end-users [36]. However, when it comes to correcting misinformation, our results suggest that **VIDEO corrections were only significantly more effective for misinformation also presented as VIDEO.** This finding underscores the importance of matching the correction modality to the misinformation modality. The high sensory engagement provided by videos may require corrections that are equally engaging and rich in order to effectively counteract the initial credibility fostered by video misinformation. In contrast, video corrections did not provide a significant advantage when correcting misinformation presented in **Text** or **Image**. This suggests that while rich media *can* enhance the perceived credibility of misinformation, it does not universally enhance the effectiveness of corrections.

Conversely, when misinformation is presented through lean media with fewer audio-visual cues, corrections delivered via similarly lean modalities can still be effective. This is evidenced in our results, where **Text** corrections were more effective than **Image** corrections for misinformation presented in **Text** form. Previous research has established that text corrections are effective in addressing text-based misinformation [10, 33, 67, 85]. However, these studies did not compare text corrections with corrections delivered through other media, leaving the relative effectiveness of different correction modalities unexplored. Our results provide important evidence that, **when misinformation is presented in Text form, lean corrections may outperform richer media corrections.**

Our study advances this understanding by demonstrating that **Text** corrections are just as effective as other modalities in correcting text-based misinformation. Interestingly, several participants expressed a preference against richer modalities, including **Image** and **Video**, citing concerns over how easily content in these modalities can be manipulated. Notably, similar concerns were not raised about the **Text** modality, despite it being arguably the easiest to manipulate. With advancement in generative artificial intelligence (AI), creating realistic-looking visuals now requires only a simple text prompt. Our qualitative findings suggest that people may become increasingly sceptical towards richer modalities, which may, in turn, make them less convincing. Specifically, as advancements in text-to-video models make generating high-quality videos even easier, the potential for misinformation to spread through video content will likely increase, but public scepticism towards information (both fake and factual) conveyed through video may also grow. End-users may increasingly doubt the authenticity of video content due to how easily it can be fabricated, making video-based misinformation potentially more prevalent, and video-based corrections potentially less convincing. Future research should investigate how the effectiveness of richer media corrections evolves over time, especially in the era of generative AI.

In addition, modality of the correction should not be considered in isolation. Our findings suggest that like-for-like correction of misinformation (in terms of modality) may be most effective when counteracting text and video. However, other design factors—particularly the length of the correction relative to the original item of misinformation—may interact with the effect of modality. While prior research indicates that short videos may be more effective than lengthy articles for correcting misinformation [91], Tandoc et al. [79] argue longer content can be perceived as more credible due to its detailed nature and potential for more counter-arguments. Hence, the relationship between the length of misinformation and correction remains unclear. For instance, can a brief video effectively address more extensive misinformation narratives? Our study shows that when comparing short-forms of misinformation and correction, it is effective to match the correction modality to misinformation modality. Future research should explore whether this holds for different lengths of misinformation and correction modalities.

5.2 Towards Effectively Combating Online Misinformation

Our analysis revealed that factors such as familiarity with the news story and its source (source of information and/or platform) also play a crucial role in credibility assessment. Participants with prior knowledge of a particular news topic experienced a significant decrease in their perceived credibility of misinformation after seeing the correction tweet. This is in line with the finding that that individuals who have higher news literacy are less likely to endorse misinformation [14, 20]. As highlighted by Chan [14], a higher overall level of news literacy within a society enhances the collective ability to identify and reject misinformation, thereby reducing its impact and dissemination. We also observed that some participants' views were affected by their distrust of the X platform, which our stimuli were designed to mirror. The distrust towards social media content may stem from the widespread prevalence of misleading or fabricated information frequently encountered on these platforms [4]. As Zhang et al. [93] highlighted, when users harbour an innate distrust towards a platform, they are likely to view fact-checking interventions with scepticism as well. While some literature suggests that users on social media platforms pay less attention to the source of news [24, 44, 55, 58], we found that the presence of a source—or the absence of it—was a key indicator of distrust.

Our findings on effective correction modalities have several implications for HCI research and design. First, current misinformation correction interventions are largely text-based, including those implemented on video-sharing platforms. For instance, YouTube uses text-based information panels to provide additional context for topics which are often subject to misinformation [92], while X itself includes text-based 'community notes' that can be used to correct misinformation. Based on our findings, we recommend that correction interventions should be designed to match the modality of misinformation to be most effective. Furthermore, emphasis of the source when providing corrections is essential. When providing corrections platforms could design tools which helps the users to highlight the source, such as interactive overlays that display detailed information about the origin, credibility, and verification process of the source with a simple click or hover action. Finally, warning messages indicating whether specific information is true or false have already been implemented by various platforms [39, 57]. Taking this a step further, platforms should increasingly incorporate AI detection algorithms to identify and label potential AI-generated content with disclaimers like 'Generated by AI' or "Likely Manipulated," helping users make informed judgments about the authenticity and credibility of the content they encounter, akin to TikTok's Auto-Label feature [80], particularly when dealing with videos as they are typically seen as more persuasive.

5.3 Limitations and Future Work

Our study has several limitations. First, our experiment was limited to a single platform (X), chosen due to its prominent role as a key resource for journalists and news consumers [52]. Future work should conduct further cross-platform studies to assess the effects of misinformation and correction modalities independent of platform-specific factors. Second, participants may have been previously exposed to the news content in the tweets, although we addressed this potential confound by capturing their familiarity with the news and accounting for it in our statistical models. Third, we tested only three distinct modalities; **Text**, **Image**, and **Video**. However, each modality encompasses a range of properties that influence its level of richness. For example, the **Image** modality is not only comprised of traditional photographs, but also includes infographics and data visualisations, which may be more compelling in counteracting misinformation. Similarly, the effectiveness of the **Video** modality may vary based on factors such as brightness, clarity, sound [48]. Future research should investigate how these additional properties of modalities influence the effectiveness of misinformation correction. This can help determine the appropriate level of richness required to effectively debunk various types of misinformation, such as political, pseudo-scientific, and medical misinformation.

6 Conclusion

In this short paper, we examined how various modalities for correcting misinformation impact their effectiveness, taking into account the format in which the misinformation was originally presented. Our study provides much-needed insight towards understanding the effectiveness of richer modalities for misinformation correction. Specifically, we found that correcting misinformation delivered through a rich modality is most effective when using another rich modality. However, for lean modalities like **Text**, simply using the same modality is often enough to correct misinformation. They key contribution of this research to HCI is that is essential to choose the appropriate modality for presenting misinformation correction and to align it with the modality of the original misinformation since our results show that corrections do not exist in a vacuum where videos would always be preferred. By considering modality, alongside other important credibility factors such as source and platform, corrections can be better designed to counter misinformation and enhance their impact.

References

[1] Jennifer Allen, Cameron Martel, and David G Rand. 2022. Birds of a feather don't fact-check each other: Partisanship and the evaluation of news in Twitter's Birdwatch crowdsourced fact-checking program. In *Proceedings of the 2022 CHI conference on human factors in computing systems*. 1–19.

[2] Susie Allen. 2023. How Social Media Rewards Misinformation. https://insights.som.yale.edu/insights/how-social-media-rewards-misinformation.

[3] Malik Almaliki. 2019. Online misinformation spread: A systematic literature map. In *Proceedings of the 2019 3rd international conference on information system and data mining*. 171–178.

[4] J. Baxter Oliphant Amy Mitchell, Mark Jurkowitz and Elisa Shearer. 2020. Americans Who Mainly Get Their News on Social Media Are Less Engaged, Less Knowledgeable.

[5] Scott Appling, Amy Bruckman, and Munmun De Choudhury. 2022. Reactions to Fact Checking. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–17.

[6] Ahmer Arif, John J. Robinson, Stephanie A. Stanek, Elodie S. Fichet, Paul Townsend, Zena Worku, and Kate Starbird. 2017. A Closer Look at the Self-Correcting Crowd: Examining Corrections in Online Rumors. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '17)*. ACM. doi:10.1145/2998181.2998294

[7] Nora Benavidezo. 2023. Musk, Twitter and Misinformation in a Time of War. https://www.freepress.net/blog/musk-twitter-x-disinformation-war.

[8] Md Momen Bhuiyan, Michael Horning, Sang Won Lee, and Tanushree Mitra. 2021. NudgeCred: Supporting News Credibility Assessment on Social Media Through Nudges. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (Oct. 2021), 1–30. doi:10.1145/3479571

[9] Md Momen Bhuiyan, Kexin Zhang, Kelsey Vick, Michael A. Horning, and Tanushree Mitra. 2018. FeedReflect: A Tool for Nudging Users to Assess News Credibility on Twitter. In *Companion of the 2018 ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW '18)*. ACM, 205–208. doi:10.1145/3272973.3274056

[10] Leticia Bode and Emily K Vraga. 2015. In related news, that was wrong: The correction of misinformation through related stories functionality in social media. *Journal of Communication* 65, 4 (2015), 619–638.

[11] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2 (Jan. 2006), 77–101. doi:10.1191/1478088706qp063oa

[12] Ceren Budak, Divyakant Agrawal, and Amr El Abbadi. 2011. Limiting the spread of misinformation in social networks. In *Proceedings of the 20th international conference on World wide web*. 665–674.

[13] Cody Buntain, Richard Bonneau, Jonathan Nagler, and Joshua A. Tucker. 2021. YouTube Recommendations and Effects on Sharing Across Online Social Platforms. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 1–26. doi:10.1145/3449085

[14] Michael Chan. 2022. News literacy, fake news recognition, and authentication behaviors after exposure to fake news on social media. *new media & society* (2022), 14614448221127675.

[15] Guangyu Chen, Paolo Ciuccarelli, and Sara Colombo. 2022. VisualBubble: Exploring how reflection-oriented user experiences affect users' awareness of their exposure to misinformation on social media. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–7.

[16] Sijing Chen, Lu Xiao, and Akit Kumar. 2023. Spread of misinformation on social media: What contributes to it and how to combat it. *Computers in Human Behavior* 141 (2023), 107643.

[17] Xinran Chen and Sei-Ching Joanna Sin. 2013. 'Misinformation? What of it?'Motivations and individual differences in misinformation sharing on social media. *Proceedings of the American Society for Information Science and Technology* 50, 1 (2013), 1–4.

[18] Yuwei Chuai, Haoye Tian, Nicolas Pröllochs, and Gabriele Lenzini. 2024. Did the Roll-Out of Community Notes Reduce Engagement With Misinformation on X/Twitter? *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (Nov. 2024), 1–52. doi:10.1145/3686967

[19] Saoirse Connor Desai and Stian Reimers. 2023. Does explaining the origins of misinformation improve the effectiveness of a given correction? *Memory & Cognition* 51, 2 (2023), 422–436.

[20] Stephanie Craft, Seth Ashley, and Adam Maksl. 2017. News media literacy and conspiracy theory endorsement. *Communication and the Public* 2, 4 (2017), 388–401.

[21] Richard L Daft and Robert H Lengel. 1986. Organizational information requirements, media richness and structural design. *Management science* 32, 5 (1986), 554–571.

[22] Statista Research Department. 2024. X/Twitter: Global audience 2024 | Statista — statista.com. https://www.statista.com/statistics/242606/number-of-active-twitter-users-in-selected-countries/.

[23] Nicholas Diakopoulos, Munmun De Choudhury, and Mor Naaman. 2012. Finding and assessing social media information sources in the context of journalism. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*. ACM, 2451–2460. doi:10.1145/2207676.2208409

[24] Nicholas Dias, Gordon Pennycook, and David G Rand. 2020. Emphasizing publishers does not effectively reduce susceptibility to misinformation on social media. (2020).

[25] Graham N Dixon, Brooke Weberling McKeever, Avery E Holton, Christopher Clarke, and Gina Eosco. 2015. The power of a picture: Overcoming scientific misinformation by communicating weight-of-evidence information with visual exemplars. *Journal of Communication* 65, 4 (2015), 639–659.

[26] Chiara Patricia Drolsbach and Nicolas Pröllochs. 2023. Diffusion of Community Fact-Checked Misinformation on Twitter. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (Sept. 2023), 1–22. doi:10.1145/3610058

[27] Ziv Epstein, Gordon Pennycook, and David Rand. 2020. Will the Crowd Game the Algorithm?: Using Layperson Judgments to Combat Misinformation on Social Media by Downranking Distrusted Sources. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. ACM. doi:10.1145/3313831.3376232

[28] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods* 39, 2 (2007), 175–191.

[29] Kathrin Figl, Samuel Kießling, and Ulrich Remus. 2023. Do symbol and device matter? The effects of symbol choice of fake news flags and device on human interaction with fake news on social media platforms. *Computers in Human Behavior* 144 (2023), 107704.

[30] Dilrukshi Gamage, James Stomber, Farnaz Jahanbakhsh, Bill Skeet, and Gautam Kishore Shahi. 2022. Designing credibility tools to combat mis/disinformation: A human-centered approach. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–4.

[31] Mingkun Gao, Ziang Xiao, Karrie Karahalios, and Wai-Tat Fu. 2018. To label or not to label: The effect of stance and credibility labels on readers' selection and perception of news articles. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–16.
[32] R Kelly Garrett, Erik C Nisbet, and Emily K Lynch. 2013. Undermining the corrective effects of media-based political fact checking? The role of contextual cues and naïve theory. *Journal of Communication* 63, 4 (2013), 617–637.
[33] R Kelly Garrett and Brian E Weeks. 2013. The promise and peril of real-time corrections to political misperceptions. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 1047–1058.
[34] R. Kelly Garrett and Brian E. Weeks. 2013. The promise and peril of real-time corrections to political misperceptions. In *Proceedings of the 2013 conference on Computer supported cooperative work (CSCW '13)*. ACM, 1047–1058. doi:10.1145/2441776.2441895
[35] Chen Guo, Nan Zheng, and Chengqi Guo. 2023. Seeing is not believing: a nuanced view of misinformation warning efficacy on video-sharing social media platforms. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–35.
[36] Michael Hameleers, Thomas E Powell, Toni GLA Van Der Meer, and Lieke Bos. 2020. A picture paints a thousand lies? The effects and mechanisms of multimodal disinformation and rebuttals disseminated via social media. *Political communication* 37, 2 (2020), 281–301.
[37] Dr Miah Hammond-Errey. 2023. Elon Musk's Twitter is becoming a sewer of disinformation. https://www.ussc.edu.au/elon-musks-twitter-is-becoming-a-sewer-of-disinformation.
[38] Andrew Hutchinson. 2022. New Study Shows Twitter is the Most Used Social Media Platform Among Journalists. https://www.socialmediatoday.com/news/new-study-shows-twitter-is-the-most-used-social-media-platform-among-journa/626245/.
[39] Instagram. 2019. Combatting Misinformation on Instagram | Instagram — about.instagram.com. https://about.instagram.com/blog/announcements/combatting-misinformation-on-instagram.
[40] Caroline Jack. 2017. Lexicon of lies: Terms for problematic information. *Data & Society* 3, 22 (2017), 1094–1096.
[41] Farnaz Jahanbakhsh and David R Karger. 2024. A Browser Extension for in-place Signaling and Assessment of Misinformation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–21.
[42] Farnaz Jahanbakhsh, Yannis Katsis, Dakuo Wang, Lucian Popa, and Michael Muller. 2023. Exploring the use of personalized AI for identifying misinformation on social media. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–27.
[43] Farnaz Jahanbakhsh, Amy X. Zhang, Adam J. Berinsky, Gordon Pennycook, David G. Rand, and David R. Karger. 2021. Exploring Lightweight Interventions at Posting Time to Reduce the Sharing of Misinformation on Social Media. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 1–42. doi:10.1145/3449092
[44] Maurice Jakesch, Moran Koren, Anna Evtushenko, and Mor Naaman. 2019. The role of source and expressive responding in political news evaluation. In *Computation and journalism symposium*.
[45] Prerna Juneja and Tanushree Mitra. 2022. Human and Technological Infrastructures of Fact-checking. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (Nov. 2022), 1–36. doi:10.1145/3555143
[46] Jan Kirchner and Christian Reuter. 2020. Countering fake news: A comparison of possible solutions regarding user acceptance and effectiveness. *Proceedings of the ACM on Human-computer Interaction* 4, CSCW2 (2020), 1–27.
[47] Travis Kriplean, Caitlin Bonnar, Alan Borning, Bo Kinney, and Brian Gill. 2014. Integrating on-demand fact-checking with public dialogue. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW'14)*. ACM, 1188–1199. doi:10.1145/2531602.2531677
[48] Yingdan Lu and Cuihua Shen. 2023. Unpacking multimodal fact-checking: Features and engagement of fact-checking videos on Chinese TikTok (Douyin). *Social Media+ Society* 9, 1 (2023), 20563051221150406.
[49] Michael P Lynch. 2022. Memes, misinformation, and political meaning. *The Southern Journal of Philosophy* 60, 1 (2022), 38–56.
[50] Ying Ma, Zhanna Sarsenbayeva, Jarrod Knibbe, and Jorge Goncalves. 2025. Exploring the effects of location information on perceptions of news credibility and sharing intention. *International Journal of Human-Computer Studies* 193 (2025), 103378. doi:10.1016/j.ijhcs.2024.103378
[51] Farhad Manjoo. 2023. Twitter Is Broken. Thanks, Elon. https://www.nytimes.com/2023/04/14/opinion/elon-musk-twitter.html.
[52] Jeffrey Gottfried Mark Jurkowitz. 2022. Twitter is the go-to social media site for U.S. journalists, but not for the public - Editor and Publisher — editorandpublisher.com. https://www.editorandpublisher.com/stories/twitter-is-the-go-to-social-media-site-for-us-journalists-but-not-for-the-public,232619.
[53] Cameron Martel, Mohsen Mosleh, and David G Rand. 2021. You're definitely wrong, maybe: Correction style has minimal effect on corrections of misinformation online. *Media and Communication* 9, 1 (2021), 120–133.
[54] Nicole Martin. 2021. How Social Media Has Changed How We Consume News. https://www.forbes.com/sites/nicolemartin1/2018/11/30/how-social-media-has-changed-how-we-consume-news/.
[55] Sarah McGrew, Joel Breakstone, Teresa Ortega, Mark Smith, and Sam Wineburg. 2018. Can students evaluate online sources? Learning from assessments of civic online reasoning. *Theory & research in social education* 46, 2 (2018), 165–193.
[56] Mohsen Mosleh, Cameron Martel, Dean Eckles, and David Rand. 2021. Perverse Downstream Consequences of Debunking: Being Corrected by Another User for Posting False Political News Increases Subsequent Sharing of Low Quality, Partisan, and Toxic Content in a Twitter Field Experiment. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM. doi:10.1145/3411764.3445642
[57] Adam Mosseri. 2016. Addressing Hoaxes and Fake News | Meta — about.fb.com. https://about.fb.com/news/2016/12/news-feed-fyi-addressing-hoaxes-and-fake-news/.
[58] Thomas Nygren, Fredrik Brounéus, and Göran Svensson. 2019. Diversity and Credibility in Young People's News Feeds: A Foundation for Teaching and Learning Citizenship in a Digital Era. *Journal of Social Science Education* 18, 2 (2019), 87–109.
[59] Daniel M Oppenheimer, Tom Meyvis, and Nicolas Davidenko. 2009. Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of experimental social psychology* 45, 4 (2009), 867–872.
[60] Saumya Pareek and Jorge Goncalves. 2024. Peer-supplied credibility labels as an online misinformation intervention. *International Journal of Human-Computer Studies* 188 (2024), 103276.
[61] Irene V. Pasquetto, Eaman Jahani, Shubham Atreja, and Matthew Baum. 2022. Social Debunking of Misinformation on WhatsApp: The Case for Strong and In-group Ties. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW1 (March 2022), 1–35. doi:10.1145/3512964
[62] Yilang Peng, Yingdan Lu, and Cuihua Shen. 2023. An agenda for studying credibility perceptions of visual misinformation. *Political Communication* 40, 2 (2023), 225–237.
[63] Gordon Pennycook, James Allan Cheyne, Nathaniel Barr, Derek J Koehler, and Jonathan A Fugelsang. 2015. On the reception and detection of pseudo-profound bullshit. *Judgment and Decision making* 10, 6 (2015), 549–563.
[64] Gordon Pennycook and David G Rand. 2020. Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of personality* 88, 2 (2020), 185–200.
[65] Stefan Pfattheicher and Simon Schindler. 2016. Misperceiving bullshit as profound is associated with favorable views of Cruz, Rubio, Trump and conservatism. *PloS one* 11, 4 (2016), e0153419.
[66] Patrick R Rich and Maria S Zaragoza. 2016. The continued influence of implied and explicitly stated misinformation in news reports. *Journal of experimental psychology: learning, memory, and cognition* 42, 1 (2016), 62.
[67] Patrick R Rich and Maria S Zaragoza. 2020. Correcting misinformation in news stories: An investigation of correction timing and correction durability. *Journal of Applied Research in Memory and Cognition* 9, 3 (2020), 310–322.
[68] Mohammed Saeed, Nicolas Traub, Maelle Nicolas, Gianluca Demartini, and Paolo Papotti. 2022. Crowdsourced Fact-Checking at Twitter: How Does the Crowd Compare With Experts?. In *Proceedings of the 31st ACM International Conference on Information & Knowledge Management (CIKM '22)*. ACM, 1736–1746. doi:10.1145/3511808.3557279
[69] Emily Saltz, Claire R Leibowicz, and Claire Wardle. 2021. Encounters with visual misinformation and labels across platforms: An interview and diary study to inform ecosystem approaches to misinformation interventions. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–6.
[70] Connie Moon Sehat, Ryan Li, Peipei Nie, Tarunima Prabhakar, and Amy X. Zhang. 2024. Misinformation as a Harm: Structured Approaches for Fact-Checking Prioritization. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (April 2024), 1–36. doi:10.1145/3641010
[71] Snopes. [n. d.]. About Us | Snopes.com — snopes.com. https://www.snopes.com/about/.
[72] Keith E. Stanovich and Maggie E. Toplak. 2023. Actively Open-Minded Thinking and Its Measurement. *Journal of Intelligence* 11, 2 (2023). doi:10.3390/jintelligence11020027
[73] Laerd Statistics. [n. d.]. Kruskal-Wallis H Test in SPSS Statistics | Procedure, output and interpretation of the output using a relevant example. — statistics.laerd.com. https://statistics.laerd.com/spss-tutorials/kruskal-wallis-h-test-using-spss-statistics.php.
[74] S. Shyam Sundar. 2008. The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility. *Digital Media, Youth, and Credibility* (2008), 73–100. doi:10.1162/dmal.9780262562324.073
[75] S Shyam Sundar, Maria D Molina, and Eugene Cho. 2021. Seeing Is Believing: Is Video Modality More Powerful in Spreading Fake News via Online Messaging Apps? *Journal of Computer-Mediated Communication* 26, 6 (Aug. 2021), 301–319. doi:10.1093/jcmc/zmab010
[76] Briony Swire, Adam J Berinsky, Stephan Lewandowsky, and Ullrich KH Ecker. 2017. Processing political misinformation: Comprehending the Trump phenomenon. *Royal Society open science* 4, 3 (2017), 160802.

[77] Sadiq Muhammed T and Saji K. Mathew. 2022. The disaster of misinformation: a review of research in social media. *International Journal of Data Science and Analytics* 13, 4 (Feb. 2022), 271–285. doi:10.1007/s41060-022-00311-6

[78] Yuko Tanaka, Miwa Inuzuka, Hiromi Arai, Yoichi Takahashi, Minao Kukita, and Kentaro Inui. 2023. Who does not benefit from fact-checking websites? A psychological characteristic predicts the selective avoidance of clicking uncongenial facts. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.

[79] Edson C. Tandoc, James Chong Boi Lee, Sangwon Lee, and Pei Jun Quek. 2022. Does Length Matter? The Impact of Fact-Check Length in Reducing COVID-19 Vaccine Misinformation. *Mass Communication and Society* 27, 4 (Dec. 2022), 679–709. doi:10.1080/15205436.2022.2155195

[80] TikTok. [n. d.]. About AI-generated content | TikTok Help Center — support.tiktok.com. https://support.tiktok.com/en/using-tiktok/creating-videos/ai-generated-content.

[81] Yu-Chia Tseng, Nanyi Bi, Yung-Ju Chang, and Chien Wen (Tina) Yuan. 2022. Investigating Correction Effects of Different Modalities for Misinformation about COVID-19. In *Companion Computer Supported Cooperative Work and Social Computing*. ACM. doi:10.1145/3500868.3559455

[82] Jacqueline Urakami, Yeongdae Kim, Hiroki Oura, and Katie Seaborn. 2022. Finding Strategies Against Misinformation in Social Media: A Qualitative Study. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI '22)*. ACM, 1–7. doi:10.1145/3491101.3519661

[83] 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 (Jan. 2016), 554–559. doi:10.1073/pnas.1517441113

[84] Emily K Vraga, Leticia Bode, and Melissa Tully. 2022. The effects of a news literacy video and real-time corrections to video misinformation related to sunscreen and skin cancer. *Health communication* 37, 13 (2022), 1622–1630.

[85] Nathan Walter and Sheila T Murphy. 2018. How to unring the bell: A meta-analytic approach to correction of misinformation. *Communication monographs* 85, 3 (2018), 423–441.

[86] Senuri Wijenayake, Danula Hettiachchi, Simo Hosio, Vassilis Kostakos, and Jorge Goncalves. 2021. Effect of Conformity on Perceived Trustworthiness of News in Social Media. *IEEE Internet Computing* 25, 1 (2021), 12–19. doi:10.1109/MIC.2020.3032410

[87] Senuri Wijenayake, Niels van Berkel, Vassilis Kostakos, and Jorge Goncalves. 2019. Measuring the effects of gender on online social conformity. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–24.

[88] Senuri Wijenayake, Niels van Berkel, Vassilis Kostakos, and Jorge Goncalves. 2020. Quantifying the Effect of Social Presence on Online Social Conformity. *Proc. ACM Hum.-Comput. Interact.* 4, CSCW1, Article 55 (May 2020), 22 pages. doi:10.1145/3392863

[89] Chloe Wittenberg, Ben M. Tappin, Adam J. Berinsky, and David G. Rand. 2021. The (minimal) persuasive advantage of political video over text. *Proceedings of the National Academy of Sciences* 118, 47 (nov 2021). doi:10.1073/pnas.2114388118

[90] Amir Hossein Yazdavar, Mohammad Saeid Mahdavinejad, Goonmeet Bajaj, William Romine, Amit Sheth, Amir Hassan Monadjemi, Krishnaprasad Thirunarayan, John M. Meddar, Annie Myers, Jyotishman Pathak, and Pascal Hitzler. 2020. Multimodal mental health analysis in social media. *PLOS ONE* 15, 4 (April 2020), e0226248. doi:10.1371/journal.pone.0226248

[91] Dannagal G Young, Kathleen Hall Jamieson, Shannon Poulsen, and Abigail Goldring. 2018. Fact-checking effectiveness as a function of format and tone: Evaluating FactCheck. org and FlackCheck. org. *Journalism & Mass Communication Quarterly* 95, 1 (2018), 49–75.

[92] YouTube. [n. d.]. YouTube Misinformation Policies - How YouTube Works — youtube.com. https://www.youtube.com/howyoutubeworks/our-commitments/fighting-misinformation/#providing-context.

[93] Yixuan Zhang, Yimeng Wang, Nutchanon Yongsatianchot, Joseph D Gaggiano, Nurul M Suhaimi, Anne Okrah, Miso Kim, Jacqueline Griffin, and Andrea G Parker. 2024. Profiling the Dynamics of Trust & Distrust in Social Media: A Survey Study. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–24.

A Summary of Model Results

Table 2: Pairwise comparison of misinformation modality

Contrast	Estimate	Std.Error	*p* value
ImageM - TextM	0.566	0.243	0.055
ImageM - VideoM	-0.501	0.243	0.099
TextM - VideoM	-1.067	0.247	**0.000**

Table 3: Pairwise comparison of correction modalities based on misinformation modality. 'C' indicates "Correction"

Pairwise comparison	β (Std.Error)	*p* value
Modality = Text		
ImageC - TextC	-0.565 (0.277)	**0.042**
ImageC - VideoC	-0.175 (0.279)	0.530
TextC - VideoC	0.390 (0.278)	0.162
Modality = Image		
ImageC - TextC	0.166 (0.279)	0.551
ImageC - VideoC	0.396 (0.279)	0.156
TextC - VideoC	0.229 (0.279)	0.411
Modality = Video		
ImageC - TextC	0.009 (0.279)	0.973
ImageC - VideoC	-0.559 (0.279)	**0.045**
TextC - VideoC	-0.568 (0.279)	**0.04**2

Table 4: Estimated Marginal Means (EMMeans) of Credibility Shift after seeing corrections.

Misinformation Modality	Correction Modality	EMM	SE
TextM	TextC	1.053	0.258
ImageM	TextC	1.322	0.254
VideoM	TextC	1.609	0.255
TextM	ImageC	0.488	0.257
ImageM	ImageC	1.489	0.255
VideoM	ImageC	1.619	0.256
TextM	VideoC	0.663	0.257
ImageM	VideoC	1.093	0.254
VideoM	VideoC	2.178	0.257

Received 12 September 2024; revised 10 December 2024; accepted 16 January 2025