SM+S social media + society

Special Issue: Multidisciplinary approaches to mis- and disinformation studies

Unpacking Multimodal Fact-Checking: Features and Engagement of Fact-Checking Videos on Chinese TikTok (Douyin)

Social Media + Society
January-March 2023: 1–16

DOI: 10.1177/20563051221150406
journals.sagepub.com/home/sms
SAGE

Yingdan Lu[1] and Cuihua (Cindy) Shen[2]

Abstract
As fact-checking videos increasingly circulate on video-sharing platforms, more research is needed to understand the prevalent features of such videos and how they are associated with audience engagement. Drawing from the literature on fact-checking, communication, marketing, and computer science, we identified eight audiovisual features as well as seven persuasive strategies that are most relevant to fact-checking videos. Using a hybrid video analysis framework combining both automated and manual content analysis, we examined 4,309 fact-checking videos on Douyin, the Chinese version of TikTok. We found that fact-checking videos on Douyin tended to have higher brightness, less cool color dominance, and faster tempo than non-fact-checking videos from the same accounts and Douyin Trending videos, and frequently used persuasive strategies like clickbait and humor. Through feature clustering, we established three types of fact-checking videos on Douyin—long storytelling cartoons, short stimulating videos, and short authoritative videos. We found that several audiovisual features and persuasive strategies were associated with audience engagement, such as likes, comments, and reshares. This study sheds light on the common practices of fact-checking videos in Chinese cyberspace, extends the current image-as-data paradigm to fact-checking videos, and helps fact-checkers make evidence-based decisions on content creation.

Keywords
fact-checking, misinformation, multimodal, TikTok, Douyin, video, image-as-data

Introduction

Misinformation, defined as false or misleading information regardless of intent (Lazer et al., 2018), has become prevalent on social media. One effective strategy for countering misinformation is fact-checking, which typically identifies the misinformation and presents a corrective message (Pennycook & Rand, 2019; J. Zhang et al., 2021). Afforded by the low-cost production and high-reach connectivity of social media, fact-checking providers like Snopes and PolitiFact frequently posted nearly-real-time correction posts on social media to combat misinformation. Recent studies have shown that fact-checking efforts in politics, health, and science may reduce audience misperceptions, increase correction acceptance, or trigger corrective responses to misinformation (Chan et al., 2017; Hameleers et al., 2020; Margolin et al., 2018; Tully et al., 2020; J. Zhang et al., 2021).

However, most existing studies focus on textual, rather than multimodal, fact-checking content, despite the growing popularity of video-based social media platforms in the contemporary social media landscape (Auxier & Anderson, 2021) and the emergence of fact-checking videos produced by health professionals and scientists on platforms such as TikTok (Southerton, 2021; Wong, 2021). Prior research suggests superior impression and persuasiveness of videos (Clark & Paivio, 1991; Sundar, 2000) as well as the effectiveness of fact-checks in multimodal forms than that of text alone (Hameleers et al., 2020; Vraga et al., 2022). In addition, scholars have started to examine the format, types, and themes of TikTok videos on specific public health issues (Basch et al., 2021; Fowler et al., 2022; Li et al., 2021; Song

[1]Stanford University, USA
[2]University of California, Davis, USA

Corresponding Author:
Yingdan Lu, Department of Communication, Stanford University, Room 110, Building 120, 450 Jane Stanford Way, Stanford, CA 94305, USA.
Email: yingdan@stanford.edu

et al., 2021). However, research remains scant on (1) what specific audiovisual features of fact-checking content are prevalent and (2) how these features contribute to audience engagement on video-sharing social media platforms. Answers to these two questions are crucial for systematically analyzing existing fact-checking content and designing engaging and effective video campaigns to combat misinformation at scale.

This article attempts to fill this critical gap by reporting an exploratory examination of the audiovisual features and persuasive strategies of fact-checking videos on Douyin, the Chinese version of TikTok, which has over 600 million daily active users across mainland China (Bursztynsky, 2021). Drawing from the literature on fact-checking and audiovisual analysis, we systematically analyzed 4,309 videos from nine fact-checking accounts on Douyin. We uncovered diverse audiovisual features and persuasive strategies, which converged into three distinct types of fact-checking videos. Through regression analysis, we also explored the relationship between video features and user engagement.

Our findings make a number of contributions to fact-checking research. Theoretically, this study is among the first to unpack the "video" modality into specific audiovisual features and persuasive strategies specific to fact-checking videos, and examine their respective association—both independently and in conjunction with each other—with audience engagement. Methodologically, unlike previous research on TikTok videos which relies solely on human-based content analysis to study features or themes (Basch et al., 2021; Fowler et al., 2022; Li et al., 2021), this study proposes a hybrid video analysis framework that combines both automated and manual content analysis. It also provides a typology of fact-checking videos on Douyin by identifying three distinct types, creating a baseline for further studies across platforms and cultural contexts. Practically, the typology of fact-checking videos on Douyin revealed from this study allows fact-checkers to make evidence-based decisions on how to craft their content with engagement metrics in mind.

In the following sections, we first review fact-checking on short-video platforms and the contextual details of TikTok and Douyin. Then, we explicate the audiovisual features and persuasive strategies relevant to fact-checking videos and present four research questions. Then, we detail the data, analytical methods, and results. We conclude the article by discussing the implications of our findings, limitations, and future research directions.

Fact-Checking on Short-Video Platforms

Video-sharing platforms have become a hallmark of today's social media landscape (Auxier & Anderson, 2021). Unlike traditional social media platforms like Twitter and Facebook, short-video platforms afford a suite of video-editing features, which allow users to create, disseminate, and consume fact-checking content in richer information modalities with ease. As a result, many videos circulated on TikTok and Douyin are "quirky videos" that are "not too professionally or aesthetically produced" (Wang, 2020). These unique affordances of TikTok and Douyin significantly lowered the barriers to creating highly effective fact-checking videos for amateurs and professionals alike. They also enabled fact-checkers to debunk rumors in a more timely fashion. Taken together, short-video platforms could potentially amplify the accessibility and effectiveness of fact-checking content, making them a particularly important, yet understudied, context of fact-checking.

TikTok has been increasingly used for disseminating fact-based messages on public health issues, such as Covid-19 information and prevention (Basch et al., 2021; Li et al., 2021), sex education (Fowler et al., 2022) and pulmonary disease (Song et al., 2021). In addition to promoting accurate information, health professionals and scientists have been actively posting videos to correct scientific misunderstandings and tackle misinformation (Southerton, 2021; Wong, 2021). For example, Vietnamese officials started a TikTok dance video challenge to show proper ways of washing hands during the Covid-19 pandemic (Zulli & Zulli, 2022).

We situated our study on Douyin, the Chinese version of TikTok. While Douyin and TikTok were created with similar functionality and affordances, Douyin is under the control of the Chinese state and follows all Chinese Internet regulations (Lu & Pan, 2022). Fact-checking efforts on Douyin are part of the Chinese government's Internet control protocols. Since August 2018, the Cyberspace Administration of China (CAC) has been posting videos on its official rumor-debunking channel on Douyin (Xinhuanet, 2021). Since 2020, many local governments have followed suit.[1] In other words, fact-checking content on Douyin comes from a mixture of government-controlled channels (e.g., government agencies and state media) and independent fact-checkers (e.g., private companies and influencers).

Video Features and Fact-Checking

Videos contain multiple modalities (text, visual, and audio) and are more quickly processed, more easily recalled, and more persuasive than text (Clark & Paivio, 1991; Sundar, 2000; Wittenberg et al., 2020). As videos may cue more realism heuristics, enhance the perceived quality of the content, and elicit greater positive emotions and trust (Marmolin, 1992; Sundar & Limperos, 2013), fact-checking videos can be more powerful in correcting misperceptions. Studies comparing fact-checking in different formats found that fact-checking videos are more effective in reducing beliefs about misinformation than text-based fact-checking articles (Vraga et al., 2022; Young et al., 2018). However, studies on fact-checking videos have not parsed videos into specific and discrete features; therefore, it remains unclear exactly how and why fact-checking videos engage the audience and correct misconceptions. Next, we examined specific audiovisual features and persuasive strategies relevant to fact-checking videos.

Audiovisual Features

Computer science research on videos often starts with computing audiovisual features that are objective and independent of audience perceptions and can be computed by computer vision algorithms (Patiño-Escarcina & Costa, 2008). We explored the audiovisual features shown in past literature to be most relevant to audience attention and engagement with fact-checking videos.

Visual Features. We focused on five basic visual features—*brightness, entropy, warm colors, cool colors*, and the presence of *faces. Brightness*, as the basic pixel-level color feature, has been directly examined in the fact-checking literature. K. Chen et al. (2022) found that correction videos regarding COVID-19 conspiracies on YouTube used higher brightness than conspiracy videos. Though previous studies have shown how higher brightness may improve the visual working memory and emotional valence (Qian et al., 2018; Wilms & Oberfeld, 2018), it remains unclear whether brightness is directly associated with audience engagement in fact-checking videos. Image *entropy is* a measure of color complexity and is operationalized as the heterogeneity of pixels (Lu & Pan, 2022; Shen et al., 2022; Yang et al., 2019). Shen et al. (2022) found that the visual complexity of YouTube educational videos was positively associated with video comments but negatively predicted video views and votes (sum of video likes and dislikes). Different colors can also be associated with visual engagement at different levels. *Warm colors* like reds and yellows were associated with higher content engagement on social media than *cool colors* like blues and greens, according to prior research on images (Peng & Jemmott, 2018). Yet, it is unknown whether warm and cool colors have similar effects on fact-checking videos. Another important visual feature is the presence of *faces* that has gained more attention in misinformation literature due to the prevalence of face-swap deep fakes. Studies found that images with human faces on social media were associated with more engagement (Kizilcec et al., 2014; Li & Xie, 2020). But whether face presence has a similar effect in fact-checking videos remains unknown.

Auditory Features. We explored two auditory features, *tempo* and *loudness* in fact-checking videos as they might be associated with content engagement. Tempo, also known as the speed of the audio, "determines the emotional effect of rhythms" (Duerr, 1981, p. 183) and faster tempo could reduce feelings of depression, sadness, and other negative affective responses to the music (Kellaris & Rice, 1993). Loudness could be predictive of arousal and appreciation of the content (Oxley et al., 2008; Q. Zhang et al., 2020). Hwang et al. (2021) found that influencers on video-sharing platforms used low loudness in their sponsored videos to gain favorable consumer sentiment. Therefore, we explored the level of tempo and loudness among fact-checking videos and their association with audience engagement.

Temporal Features. The length of a video is another important feature. Prior research showed that shorter videos, such as online learning videos and government TikTok videos, were associated with higher engagement (Q. Chen et al., 2021; Guo et al., 2014). Together, the importance of audiovisual features and the lack of knowledge related to fact-checking videos motivated us to pursue the first research question:

> *RQ1: Among the audiovisual features relevant to audience engagement, which are prevalent in fact-checking videos on Douyin?*

Persuasive Strategies

Beyond audiovisual characteristics, videos often employ certain strategies to intentionally persuade the audiences, and such persuasive strategies may influence perception and engagement. For example, the humor strategy was shown to capture audience's attention, improve knowledge acquisition, and elicit emotional responses to videos (Vraga et al., 2019; Wang, 2020; Young et al., 2018). Research on TikTok and Douyin highlighted the popularity of humor strategies in video creation (Basch et al., 2021; Negreira-Rey et al., 2022) and how humor videos were positively associated with audience engagement (Basch et al., 2021; Li et al., 2021). These short-video platforms, through unique affordances like "lip-syncing" and video editing features, incentivized users to engage in meme-making and replication of popular content (Kaye et al., 2021; Zulli & Zulli, 2022). Previous research has investigated different representations of humor. For example, *cartoons*, as a common type of visual humor, can "make scientific subjects more accessible and engaging for a wider audience" (Farinella, 2018, p. 1). Studies on fact-checking videos found that cartoon-based corrections were effective in reducing misinformation beliefs (Vraga et al., 2019). Opgenhaffen (2022) found that fact-checking interventions by a cartoon character were evaluated as more credible and authentic than interventions by the mainstream media accounts on sociopolitical topics. Internet *memes*, normally integrating text and images from pop culture, are heavily employed in misinformation posts given the strong discursive power of "sophomoric humor" that makes fun out of immaturity (Applegate & Cohen, 2016; Smith, 2019; Tuters & Hagen, 2020). In addition to visual humor, Young and colleagues (2018) included *funny sound effects* as one type of video humor when they designed humorous fact-checking video stimuli.

Besides humor, other persuasive strategies are commonly used in fact-checking videos. *Logic-based* correction is to explain the logical fallacies of the false claim. It has been found effective in correcting misperceptions through infographic posts on Twitter and image-based posts on Instagram (Vraga et al., 2019, 2020), but has not yet been explored in video engagement. *Storytelling* strategy is also popular in fact-checking videos because embedding facts in stories was found to be more effective in triggering behavioral change than just presenting facts alone (de Wit et al., 2011). Scientists used storytelling strategies to communicate accurate scientific information, correct false beliefs, and achieve more effective scientific persuasion (Barriga et al., 2010; ElShafie, 2018; Sangalang et al., 2019; Shelby & Ernst, 2013). Furthermore, storytelling in social media advertisements was found to be an effective strategy to gain traffic and raise engagement (Shreedhar, 2021). Recent studies also demonstrated that *authoritative sources* such as government, media, and experts labeled in fact-checking content were effective not only in reducing misperceptions (van der Meer & Jin, 2020; Vraga & Bode, 2017; J. Zhang et al., 2021), but also in increasing audience engagement (Kim et al., 2022). As users on video-sharing platforms encounter video thumbnails in their search results, like they encounter headlines when reading text news, their content engagement may be influenced by the *clickbait thumbnail* of a video. Clickbait could create a curiosity gap, causing the audience to click the headline to close the gap (Loewenstein, 1994). Research on Chinese government social media posts found that clickbait was positively correlated with content engagement in terms of reads and likes (Lu & Pan, 2021). Thus, clickbait thumbnails may attract more engagement with fact-checking videos with similar functionalities. Therefore, we ask:

> *RQ2: Considering the persuasive strategies relevant to audience engagement, which are prevalent in fact-checking videos on Douyin?*

We summarized above how each individual audiovisual feature and persuasive strategy may influence fact-checking effectiveness and engagement. There is no systematic study on fact-checking videos that uncovers how multiple video features—both independently and in conjunction with each other—function in fact-checking efforts. Knowing the correlations and interactions of different features will contribute to a parsimonious typology of fact-checking efforts on video-sharing platforms. Therefore, we ask:

> *RQ3: What audiovisual features and persuasive strategies tend to be employed together in fact-checking videos?*

Audience Engagement of Fact-Checking Videos on Social Media

In addition to evaluating the prevalence of audiovisual features and persuasive strategies of fact-checking videos on Douyin, we aim to explore how these video features and their combinations are associated with audience engagement. The contemporary social media environment is increasingly fragmented. Various content creators are facing considerable competition for viewers' attention. Audience engagement, quantified by the number of reads, likes, comments, and/or reshares on social media, has become an important indicator of the content's reach and success in social media literature (Li & Xie, 2020; Lu & Pan, 2021; Park et al., 2016; Peng & Jemmott, 2018; Tan et al., 2014).

Specifically, regarding fact-checking videos, audience engagement can be seen as the "demand" metric of such content (Siwakoti et al., 2021). Fact-checkers assess "how involved or responsive the audience is" (Humphreys, 2016, p. 48) to these videos, so they can adapt and craft more engaging content. Engagement metrics including likes, comments, and reshares act as endorsement cues that further increase the visibility and effectiveness of the fact-checking content. As Kim et al. (2022) concluded, "it is ultimately through audience engagement that fact-checking can be widely accepted and disseminated in the current media environment" (p. 782). Further, videos with high audience engagement metrics could reach a larger audience through Douyin's recommendation algorithms, gaining even more visibility and engagement (Z. Chen et al., 2019). Therefore, we ask:

> *RQ4: Which video features are associated with audience engagement in fact-checking videos on social media?*

Methods

Data

We collected videos from nine fact-checking accounts on Douyin: *Douyin Rumor-debunking* (抖音辟谣), *Toutiao Rumor-debunking* (头条辟谣), *Chinese Internet Rumor-debunking Platform* (中国互联网联合辟谣平台), *Scientific Facts* (科学辟谣), *Xlab* (求真实验室), *Dingxiang Doctor* (丁香医生), *Southern Health* (南方健康), *Feidie Shuo* (飞碟说), and *Daddy Lab* (老爸评测). We selected these nine accounts by searching for account names on Douyin with the keyword "rumor-debunking" or "fact-checking," and in news reports that featured popular fact-checking accounts on Douyin.[2] We selected these accounts, because they regularly debunked rumors in health and science, with different levels of following and popularity to increase generalizability.[3] Two out of the nine accounts were government-controlled (*Chinese Internet Rumor-debunking Platform* and *Scientific Facts*), while the rest were independent fact-checkers. We collected all videos from account creation through 15 February 2021, resulting in a multimodal dataset of 4,309 videos and their metadata (i.e., text description of the video, online engagement metrics, and video thumbnail).

Table 1. Video Data by Accounts.

Account name	All videos	Fact-checking videos
Douyin Rumor-debunking	285	163
Toutiao Rumor-debunking	152	85
Chinese Internet Rumor-debunking Platform	155	47
Scientific Facts	663	154
Xlab	276	71
Dingxiang Doctor	636	76
Southern Health	723	94
Feidie Shuo	536	48
Daddy Lab	883	107

We operationalized fact-checking content as content that debunked misinformation, as debunking is "the central form of fact-checking" (J. Zhang et al., 2021, p. 2). During the same data collection period, these accounts also posted health or science-related videos without debunking rumors or misperceptions (e.g., videos on exploring water resources on the moon), which we termed as "non-fact-checking videos." We first extracted fact-checking videos using 16 rumor-related keywords.[4] Two research assistants then manually annotated videos as fact-checking videos only if they refuted incorrect information or presented a corrective message to challenge prior misinformation. A total of 845 videos were identified as fact-checking videos (intercoder agreement = 85%), and the distribution of the data can be seen in Table 1.

Hybrid Video Analysis Framework

We developed a hybrid video analysis framework that used (1) automated analysis to extract audiovisual features and (2) manual content analysis to annotate persuasive strategies.

Audiovisual Features. As videos comprise a sequence of visual frames, we computed frame-level visual features of sampled frames[5] by applying automated visual analysis using the OpenCV library in Python (Culjak et al., 2012). We then used the frame-level features to derive video-level features, including *brightness*, *entropy*, *warm color dominance*, *cool color dominance*, and *face presence*, as detailed below (K. Chen et al., 2022; Lu & Pan, 2022).

We calculated the perceptual lightness of pixels in the CIELAB color space to denote *brightness* (Baldevbhai & Anand, 2012). After calculating the frame-level brightness score, we took the average brightness score across all sampled frames to derive the video brightness on the converted scale of 0 to 255, where 0 stood for pure black (no light) and 255 stood for white (maximum brightness).[6] For *entropy*, we converted color pixels from RGB to grayscale, computed the share of pixels with each grayscale color, and derived the entropy score across the 256 grayscale colors using Shannon's entropy metric, with the minimum entropy at 0 (a monochromatic light) and the maximum entropy at 8 (picture with extremely complicated use of color). We then took the average entropy score across all sampled frames to derive the video entropy. For dominant colors of the video, we computed the dominant color of each frame by matching each pixel to one of the 17 base colors,[7] counted pixels affiliated with each color, and found the dominant color with the highest score. We then derived the proportion of frames dominated by warm colors (red, yellow, orange, maroon, olive) and cool colors (aqua, blue, green, navy, teal) across frames as the rate of *warm color dominance* and the rate of *cool color dominance*, respectively, in a video.

To measure the presence of human faces within a video, we applied the *face_recognition* Library in Python (Geitgey, 2020) to identify whether each frame of a video contains any human faces. We calculated the proportion of frames with faces across all sampled frames as the *face presence* score. If each sampled frame of a fact-checking video contained at least one human face, the general video-level score of face presence equaled 1. By contrast, if the video contained no human face at all, the face presence score equaled 0.

To extract audio features, we used the Librosa Library in Python (McFee et al., 2015), which has been increasingly applied by scholars in disinformation studies to extract and analyze the *tempo* and *loudness* of fact-checking videos (Agarwal, 2021; Nasar et al., 2020). Tempo was measured in beats per minute (bpm) in music literature (McAuley, 2010). Higher bpm represented a faster tempo. For example, music pieces with over 120 bpm were deemed fast-tempo stimuli, while music pieces with 60–76 bpm were considered slow-tempo stimuli in experimental designs (Kellaris & Rice, 1993; Liu et al., 2018). We used the "beat" module of Librosa to derive the global tempo of each video. Following prior research (Agarwal, 2021; Schindler et al., 2016), we operationalized the loudness of the signal by calculating the root-mean-square energy (RMSE) of the signal, which compared the waveforms and their equivalent energy to portray the average level of energy. We divided the music into signal frames and calculated the RMSE for each frame. Then we took the average of RMSE scores across all signal frames to derive the video-level loudness, where a higher RMSE score (ranging from 0 to 1) represented a louder video with a higher average level of energy. For *video length*, we extracted the temporal video length (in seconds) from the metadata of the videos.

Video Persuasive Strategies All persuasive strategies of fact-checking videos were coded by two research assistants based on a predefined coding scheme. The coders were trained with the initial codebook with variables drawn from the literature, coded on 100 videos (12% overlaps) randomly drawn from the data, and their intercoder reliability was calculated in each round of coding test. After three rounds of coding tests and refinements, the codebook was set with an intercoder agreement above 80% for all variables. Then two coders independently coded all videos based on the final coding scheme.

Humor Strategies. Following prior research, we included three visual and auditory humor strategies for human annotation: *cartoons*, *memes*, and *funny sounds*. We defined cartoons with a broader concept that includes cartoons, comics, graphics, illustrations, and cartoon animations (Farinella, 2018; Kemnitz, 1973). In the coding scheme, we first presented examples and then asked the coders to annotate to what extent each video used cartoons using a 4-point Likert-type scale (0=*never*; 1=*a few*; 2=*many*; 3=*almost always*). We included both Internet memes and animated GIFs as memes (Suryawanshi et al., 2020; Tuters & Hagen, 2020), and the coders also saw examples of memes for reference before annotating the use of memes using the same 4-point Likert-type scale. For funny sounds, we presented the coders with a list of funny sounds[8] in the training phase. Then we asked them to annotate the use of funny sounds using the same 4-point Likert-type scale.

Logic-Based Strategy. For the logic-based strategy, we annotated whether the video pointed out any logical fallacies, such as causation flaw, cherry-picking, and exaggeration (Bennett, 2012; Schmid & Betsch, 2019; Vraga et al., 2020), in the rumor.

Storytelling Strategy. After qualitatively analyzing 50 sampled fact-checking videos in our dataset, we summarized three basic forms of "storytelling" for the coding scheme: narrating contemporary or historical stories, inoculation through mini-plays, and doing in-lab experiments to debunk rumors. Coders annotated the presence of these storytelling strategies using a 4-point Likert-type scale (0=*never*; 1=*a few*; 2=*many*; 3=*almost always*).

Authoritative Source. We annotated if the video explicitly cited debunking information from any of the following sources: (a) departments, agencies, platforms, officials, or websites of any level of government in China (e.g., central government office, city mayor, Centers for Disease Control and Prevention); (b) news media (e.g., party newspapers, state-owned media enterprises, *The New York Times*); (c) fact-checking websites (e.g., Tencent Fake-news Debunker); and (d) research universities and scientific institutions (e.g., Tsinghua University, doctors in Beijing Hospital, scientific reports).

Clickbait Thumbnail. To classify whether a video used a clickbait thumbnail, we first extracted the text from all thumbnails through a human-assisted optical character recognition procedure.[9] Following existing clickbait literature in communication, linguistics, and computer science (Chakraborty et al., 2016; Lu & Pan, 2021; Wei & Wan, 2017), we checked thumbnails for any of the nine clickbait indicators: listicles, general nouns, pronouns, hyperbolic words, slang, fixed phrase patterns, and special punctuation marks (exclamation marks, ellipses, and question marks) by combining automated dictionary method and human validation. Example clickbait thumbnails included text like "Should girls not drink these three kinds of drinks? Here is the truth!" and "Three things you must know about losing weight."

Table 2. Descriptive Statistics of Audiovisual Features, Persuasive Strategies, and Video Engagement Indicators.

Variable type	Variable name	*M*	*SD*
Audiovisual features	Brightness	150.24	40.45
	Entropy	6.03	1.08
	Warm color dominance	0.04	0.14
	Cool color dominance	0.04	0.15
	Face presence	0.40	0.38
	Tempo	123.40	21.25
	Loudness	0.15	0.09
	Video length(s)	41.2	45.97
Persuasive strategies	Cartoon	1.10	1.05
	Meme	0.26	0.52
	Funny sound	0.38	0.65
	Storytelling	0.40	0.82
	Logic	0.21	0.41
	Authoritative source	0.43	0.49
	Clickbait thumbnail	0.76	0.43
Engagement metrics	Video likes	36,583	114,988
	Video comments	1,716	5,409
	Video reshares	4,122	16,677

SD: standard deviation.

Engagement Metrics

The numbers of likes, comments, and reshares of each video as of 15 February 2021, were used as engagement metrics in this study. As they (see Figure S2 in Online Appendix) were highly skewed, they were log-transformed before analysis.

Analysis

Descriptive statistics of each variable can be seen in Table 2. To answer RQ1, we calculated and presented the raw statistics of all audiovisual features. As prior research has not established benchmark values for interpreting audiovisual features, to provide a frame of reference, the audiovisual statistics from two comparison groups were also included: (1) Douyin Trending videos collected by Lu and Pan (2022), which represented algorithmically selected high-visibility videos,[10] as a global benchmark; and (2) non-fact-checking videos published by the same accounts during the same data collection period, as a local benchmark to control for account-specific feature preferences. To answer RQ2, we calculated the average use of each persuasive strategy among

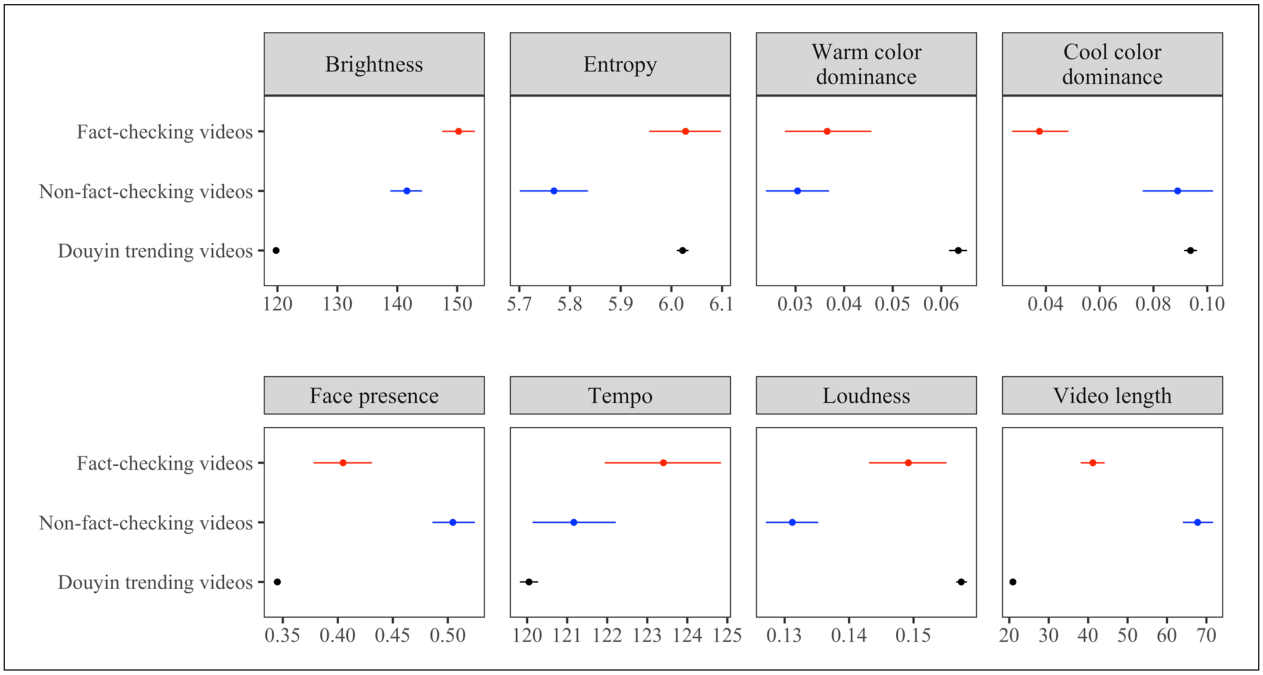

Figure 1. Audiovisual features of fact-checking videos versus non-fact-checking videos from same accounts versus Douyin Trending videos.
Note. Estimates of audiovisual features of fact-checking videos (red), non-fact-checking videos posted by the same accounts (blue), and Douyin Trending videos (black) with bootstrapped 95% confidence intervals (lines).

all fact-checking videos. As RQ3 looks for the co-presence of different features/strategies, we first calculated the correlation between each pair of features/strategies. We then built a feature vector for each video that included all audiovisual features and persuasive strategies and applied clustering analysis on the feature vectors to differentiate videos by distinct feature combinations. As both categorical and continuous variables were involved in the feature vectors, we applied K-prototype as the clustering algorithm for its usage in clustering data with mixed-type variables (Huang, 1997). To answer RQ4, we fitted linear mixed-effects regression models to estimate the association of video features with video likes, comments, and reshares, with individual accounts as random effects.

Results

RQ1 asks among audiovisual features relevant to audience engagement, what features are prevalent in fact-checking videos on Douyin. Figure 1 shows estimates of audiovisual features of fact-checking videos, compared with two benchmarks—non-fact-checking videos from the same accounts (a local benchmark) and Douyin trending videos (a global benchmark)—which contextualize our results. We found that, compared with both global and local benchmarks on Douyin, fact-checking videos exhibited significantly higher brightness ($M_{\text{fact}}=150.24$), lower cool color dominance ($M_{\text{fact}}=0.04$), and faster tempo ($M_{\text{fact}}=123.40$), while the trends for other features were more nuanced and less conclusive.

RQ2 asks about persuasive strategies relevant to audience engagement, which are prevalent in fact-checking videos on Douyin. As Table 2 shows, 76% of fact-checking videos used a clickbait thumbnail to attract attention. Humor strategies were prevalent among fact-checking videos: 63% of videos used at least one cartoon or comic. Memes ($M=0.26$) and funny sounds ($M=0.38$) were moderately used in fact-checking videos, as were storytelling ($M=0.40$) and authoritative sources ($M=0.43$). In comparison, the use of logic inoculation was low in Douyin fact-checking videos ($M=0.21$).

RQ3 asks what audiovisual features and persuasive strategies tend to be employed together in fact-checking videos. We found low to medium correlations among many audiovisual and persuasive features (Figure 2). For example, the use of cartoons had a negative correlation with entropy ($r=-0.34$, $p<.001$).

Following the elbow method, we derived three video clusters from K-prototype clustering analysis (See Table 3 and Figure S1 in Online Appendix). The first type of fact-checking videos was "long storytelling cartoons," with long video length ($M=209.28$), the prevalence of cartoons ($M=1.90$, Mode = "almost always"), and storytelling strategy ($M=1.48$, Mode = "many"). One example of this category was a video titled "*Is sugar-free milk tea a hoax?*" posted by the account *Feidie Shuo*. In this 3-min video, the creator used cartoons to present a complete story about the

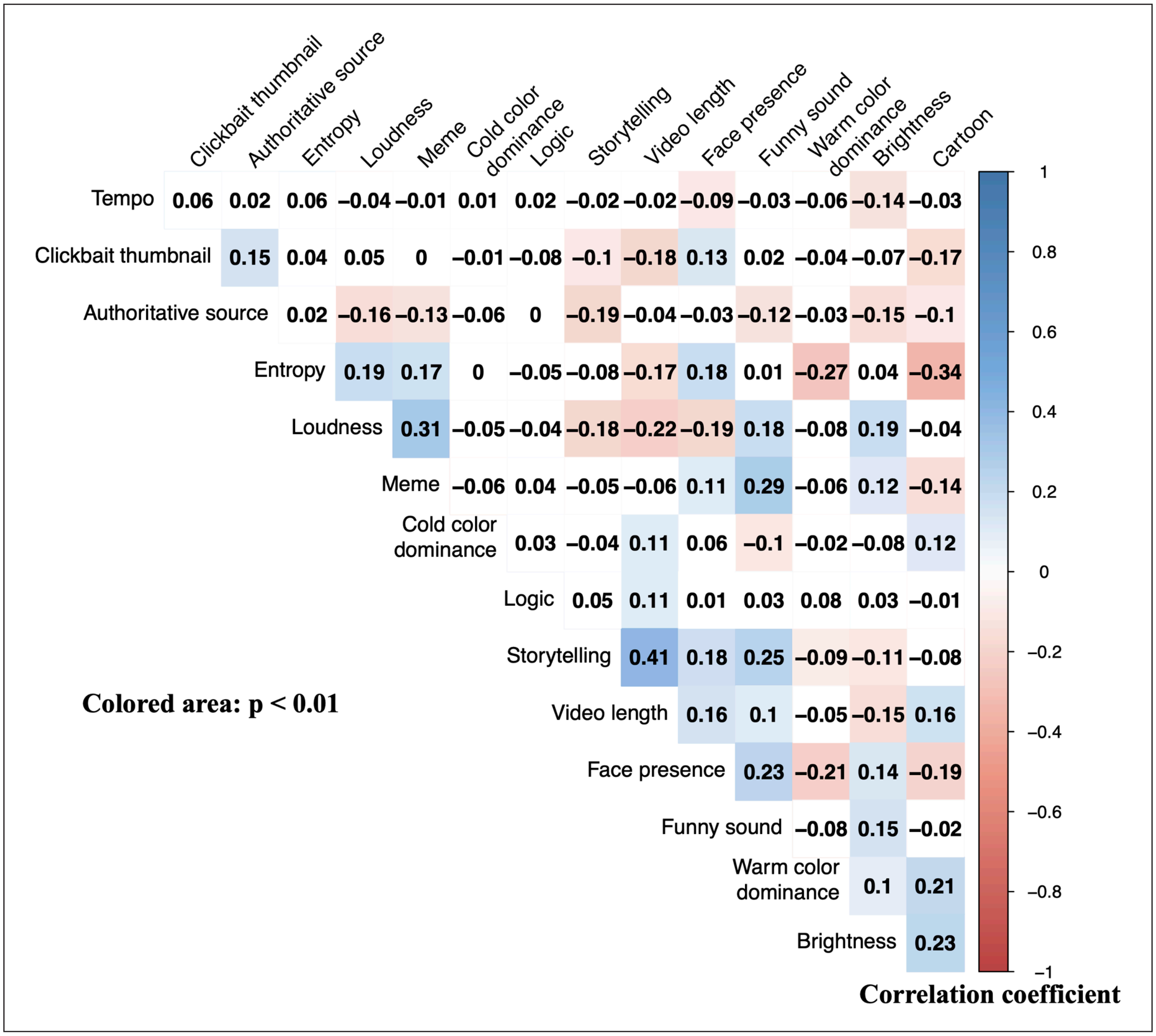

Figure 2. Correlations among audiovisual and persuasive features.
Note. The numbers are coefficients of correlation between each two variables, and the colored areas indicate coefficients that are statistically significant at a .01 level.

invention of zero-sugar beverages and its rise in popularity in China. It debunked the misinformation that zero-sugar beverages are 100% healthy and asked people to pay more attention to the ingredients when choosing zero-sugar beverages.

The second type of fact-checking videos could be described as "short stimulating videos," which were short in length (M=32.76) but high in brightness (M=175.70), entropy (M=6.09), and loudness (M=0.17). Most videos in this category used clickbait in their thumbnails (M=0.73), as shown in Panel B of Figure 3. The example video, "*Human organs detoxify on time? Believe it then you are fooled!*" posted by *Toutiao Rumor-debunking*, presented high-brightness frames that shifted quickly with loud background music and narration.

The third type of fact-checking videos are "short authoritative videos," which were shorter in duration (M=32.97). Sixty-three percent used authoritative sources like experts or government announcements to debunk rumors. For example, the 10-s video in Panel C of Figure 3, posted by *Scientific Facts*, quoted the *Chinese Dietary Guidelines* published by the government-led Chinese Nutrition Society to correct misinformation about sugar consumption.

RQ4 asks how video features are associated with video engagement, for which we fitted two sets of linear mixed-effects models predicting the number of likes, comments, and reshares, with the Douyin account as random effects. As video features were correlated with each other, we first used the three video categories identified previously as independent variables but found no significant associations between

Table 3. Video Features by Cluster.

Features	Cluster 1: Long Storytelling Cartoons	Cluster 2: Short Stimulating Videos	Cluster 3: Short Authoritative Videos
Number of videos	40	497	308
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Audiovisual features			
Brightness	121.50 (32.39)	175.70 (24.62)	112.90 (28.80)
Entropy	5.36 (1.27)	6.09 (1.05)	6.01 (1.07)
Warm color dominance	0.03 (0.04)	0.04 (0.16)	0.03 (0.11)
Cool color dominance	0.08 (0.15)	0.02 (0.11)	0.05 (0.20)
Face presence	0.43 (0.32)	0.42 (0.39)	0.38 (0.37)
Tempo	123.72 (17.87)	121.13 (19.78)	127.03 (23.39)
Loudness	0.10 (0.06)	0.17 (0.10)	0.12 (0.07)
Video length (s)	209.28 (73.40)	32.76 (20.79)	32.97 (23.76)
Persuasive strategies			
Cartoon	1.90 (1.15)	1.22 (1.06)	0.81 (0.92)
Meme	0.15 (0.43)	0.35 (0.57)	0.13 (0.39)
Funny sound	0.58 (0.59)	0.45 (0.69)	0.25 (0.58)
Storytelling	1.48 (1.04)	0.32 (0.74)	0.38 (0.82)
Logic	0.30 (0.46)	0.21 (0.41)	0.19 (0.39)
Authoritative source	0.45 (0.50)	0.30 (0.46)	0.63 (0.48)
Clickbait thumbnail	0.55 (0.50)	0.73 (0.44)	0.84 (0.37)

SD: standard deviation.

(A) Long storytelling cartoon

Notes: This 90s video used cartoons to present a complete story about the invention of zero-sugar beverages and its rise in popularity in China

(B) Short stimulating video

Notes: This 40s video had a video brightness at 210.3, loudness at 0.14, a clickbait thumbnail, and 26% frames dominated by warm colors.

(C) Short authoritative video

Notes: This 10s video applied evidence from the *Chinese Dietary Guidelines* to correct a misperception about sugar prevention.

Figure 3. Examples of fact-checking videos by category.

Table 4. Multilevel Model Explaining Video Engagement.

Variable name	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Video Likes	Video Comments	Video Reshares	Video Likes	Video Comments	Video Reshares
Video category						
Short stimulating videos	0.175	−0.128	−0.238			
	(0.248)	(0.248)	(0.301)			
Short authoritative videos	0.052	−0.348	−0.327			
	(0.256)	(0.256)	(0.311)			
Audiovisual features						
Brightness				0.001	0.003†	0.003†
				(0.002)	(0.002)	(0.002)
Entropy				−0.153**	−0.102†	−0.198**
				(0.053)	(0.053)	(0.065)
Warm color dominance				−0.090	−0.356	0.234
				(0.381)	(0.379)	(0.467)
Cool color dominance				−0.098	0.172	0.397
				(0.323)	(0.321)	(0.396)
Face presence				0.237	0.414†	0.335
				(0.224)	(0.223)	(0.275)
Tempo				−0.0003	0.002	–0.002
				(0.002)	(0.002)	(0.003)
Loudness				2.612***	1.681*	1.331
				(0.747)	(0.743)	(0.916)
Video length (s)				0.002†	0.005***	0.005**
				(0.001)	(0.001)	(0.002)
Persuasive strategies						
Cartoon				−0.061	−0.038	0.092
				(0.056)	(0.056)	(0.069)
Meme				0.333**	0.377***	0.287*
				(0.101)	(0.100)	(0.123)
Funny sound				0.100	0.107	−0.048
				(0.084)	(0.084)	(0.103)
Storytelling				−0.055	−0.067	−0.126
				(0.069)	(0.069)	(0.085)
Logic				0.131	0.081	0.229
				(0.117)	(0.117)	(0.144)
Authoritative source				−0.217†	−0.459***	−0.142
				(0.118)	(0.117)	(0.145)
Clickbait thumbnail				−0.270*	−0.229†	−0.161
				(0.125)	(0.124)	(0.153)
Constant	8.153***	5.285***	5.473***	8.650***	4.605***	5.574***
	(0.885)	(0.822)	(0.836)	(0.966)	(0.884)	(1.013)
ICC level	0.778	0.752	0.672	0.778	0.752	0.672
Log likelihood	−1,486.8	−1,486.6	−1,648.0	−1,483.7	−1,478.0	−1,651.0

Note. Number of observations = 845. Long storytelling cartoons were the comparison group for Models 1–3. Douyin accounts were included as random effects. Video likes, comments, and reshares were log-transformed due to high skewness.
†$p<.1$; *$p<.05$; **$p<.01$; ***$p<.001$.

video categories and audience engagement, as shown in Models 1–3 of Table 4. Then, we used specific video features as independent variables in Models 4–6. Loudness had positive associations with video likes ($\beta=2.612$, $p<.001$) and comments ($\beta=1.681$, $p<.05$). The extent of using memes had significant and positive associations with likes ($\beta=0.333$, $p<.01$), comments ($\beta=0.377$, $p<.001$), and reshares ($\beta=0.287$, $p<.05$). Video length was positively associated with video likes ($\beta=0.002$, $p<.1$), comments ($\beta=0.005$, $p<.001$), and reshares ($\beta=0.005$, $p<.01$). Brightness was also positively associated with video comments ($\beta=0.003$, $p<.1$) and reshares ($\beta=0.003$, $p<.1$). In contrast, entropy was negatively associated with video likes ($\beta=-0.153$, $p<.01$), comments ($\beta=-0.102$, $p<.1$), and reshares ($\beta=-0.198$, $p<.01$). Using authoritative evidence was negatively associated with video comments ($\beta=-0.459$,

$p < .001$) and likes (−0.217, $p < .1$), and so was clickbait on likes ($\beta = -0.270$, $p < .05$) and comments ($\beta = -0.229$, $p < .1$).

Discussion

As fact-checking videos increasingly circulate on video-sharing platforms, more research is needed to understand the prevalent features of such videos and how they are associated with audience engagement. Drawing from the literature in communication, marketing science, and computer science, we identified eight audiovisual features and seven persuasive strategies that are most relevant to understanding fact-checking videos. Using both automated and manual visual content analysis, we examined more than 4,000 fact-checking videos on Douyin, the mainland China version of the popular short-video platform TikTok. We found that fact-checking videos on Douyin tended to have higher brightness, less cool color dominance, and faster tempo than non-fact-checking videos from the same accounts and Douyin Trending videos, and frequently used persuasive strategies like clickbait and humor. As these video features were correlated with each other, we established through feature clustering three types of fact-checking videos on Douyin: long storytelling cartoons, short stimulating videos, and short authoritative videos. Finally, we found that several audiovisual features and persuasive strategies were associated with video engagement such as likes, comments, and reshares. Below, we discuss specific contributions in theory, practice, and methodology to the understanding of fact-checking videos.

First, our study contributes to the existing literature by identifying eight audiovisual features relevant to fact-checking videos: brightness, entropy, warm color dominance, cool color dominance, face presence, tempo, loudness, and video length. Our automated analysis of these audiovisual features offers a systematic first glimpse at their respective distribution and prevalence for fact-checking videos on the Douyin platform. To provide frames of reference, we analyzed the same audiovisual features for non-fact-checking videos from the same accounts and Douyin Trending videos. We found that fact-checking videos tend to have higher brightness, less cool color dominance, and faster tempo in comparison, but the trend was much less clear-cut for other features. We must note that both non-fact-checking videos and Douyin Trending videos were imperfect comparison groups, as the former came from an extremely homogeneous pool (same accounts) and the latter was algorithmically curated from the entire Douyin platform, so the comparison results should be interpreted with caution. Nonetheless, our exploratory findings provide a much-needed benchmark for these audiovisual features, which allows researchers to compare video analysis results in the future and design effective video stimuli for experimental research in the fact-checking domain.

Second, we identified five persuasive strategies relevant to fact-checking videos: humor, logic, storytelling, authoritative sources, and clickbait thumbnail. Humor strategies were the most prevalent in fact-checking videos, which is consistent with recent research on public health-related TikTok videos (Basch et al., 2021; Negreira-Rey et al., 2022). In comparison, the logic-based strategy was the least prevalent in our dataset, though it was common in previous fact-checking studies and contexts (Bennett, 2012; Schmid & Betsch, 2019; Vraga et al., 2020). Past findings reflected the playful nature and attention economy of short-video platforms where it is necessary to present serious content in an entertaining way to compete for users' attention (Abidin, 2021; X. Chen et al., 2021). In addition, this study unpacked humor into three dimensions (cartoon, meme, and funny sound) and found that they were associated with different levels of audience engagement. This extends the singular understanding of humor strategy and calls for future research to carefully investigate the effects of different humor strategies.

Third, as audiovisual features and persuasive strategies were correlated, we established through feature clustering three distinct types of Douyin fact-checking videos. Our attempt represents a first step toward a parsimonious typology of fact-checking videos that holistically examines meaningful clusters of features, rather than an atomistic analysis of solo features. Contextually, this typology also sheds light on the common practices of fact-checking videos in the Chinese cyberspace, where information control is of political importance and fact-checking services are sanctioned by authorities. As such, this typology provides a baseline for further validation and comparison across platforms and cultural contexts.

This study presented much-needed empirical data on the associations between specific video features—both independently and in conjunction with each other—and audience engagement of fact-checking videos. The results on some features were consistent with prior literature. For example, brightness, face presence, and memes were positively associated with at least one engagement metric. However, results on other features, such as loudness, video length, authoritative sources, and clickbait, were not aligned with prior research. These inconsistencies may be a result of our observational study design, as we were not able to isolate the effect of each individual feature while experimentally controlling the rest, when, in reality, many features were highly correlated and work in concert with each other. For example, clickbait was positively associated with the number of likes for textual posts published by government accounts on WeChat (Lu & Pan, 2021), but we found a negative association for fact-checking videos on Douyin. This discrepancy could potentially be explained by differences in modality (text vs video), theme (government content vs fact-checking content), or platform (WeChat vs Douyin). These findings again highlight the value of a parsimonious typology of fact-checking videos to reduce numerous individual features into meaningful clusters, as well as the need to conduct rigorous experimental research. Practically, our results provide guidance for fact-checkers and content creators on short-video platforms to maximize their impact. The typology of fact-checking videos helps content creators make evidence-based decisions on how to craft their content with engagement metrics in mind.

Methodologically, this study proposed a hybrid video analysis framework that combines automated and manual visual content analysis. Despite the advancement of computational video analysis (Nyhuis et al., 2021), most existing research on TikTok videos solely relies on human-based content analysis to extract features or themes (Basch et al., 2021; Fowler et al., 2022; Li et al., 2021). The hybrid video analysis framework allows us to measure audiovisual features using standardized metrics and at scale while preserving human insight for coding persuasive strategies. Through unsupervised clustering algorithms, we explored how discrete video features could be combined into meaningful categories. This feature-combination approach can be usefully applied to multimodal content in other research settings. As such, this study applies and extends the current image-as-data methodological advancement to fact-checking videos (Joo & Steinert-Threlkeld, 2018; Williams et al., 2020).

This study has three main limitations. First, our sample covered videos from nine fact-checking accounts on one specific platform, which limits its generalizability. Second, like other mainstream social media platforms, Douyin's engagement metrics (likes and reshares) are not immune to artificial manipulation tactics (Lu & Pan, 2022; Woolley, 2016), which might bias the results of this study. Finally, given this study's exploratory and cross-sectional nature, we were unable to establish causal relationships between video features and engagement.

We suggest two promising future research directions. One natural extension of our work is to experimentally test the causal relationships between the audiovisual and persuasive features of fact-checking videos—individually and in conjunction with each other—and video engagement. Furthermore, while the current study only investigated engagement, future studies should test the connection between video engagement and fact-checking effectiveness. The successful correction of misconceptions is the ultimate goal of fact-checking efforts, and it remains unclear how video features contribute to this goal, either directly or indirectly through engagement. Another related and highly promising future direction is to investigate the specific mechanisms through which video features may translate into better engagement and fact-checking effectiveness. While a theoretical framework in this area has yet developed and tested, we suggest that future researchers consider three distinct pathways through which video features may contribute to fact-checking success: (1) video features as persuasive arguments, (2) video features as realism heuristics, and (3) video features as attention determinants. Understanding the specific mechanisms would not only help build a much-needed theoretical framework of video and multimodal persuasion but also offer practical insights for platforms, fact-checking agencies, and content creators to develop successful corrective campaigns.

Acknowledgements

Our thanks to Mengmeng Guo and Huazhi Qin for superb research assistance; Jonathan Bright, Anna George, Alexandra Pavliuc, Viktoria Spaiser, and participants of the 3rd Multidisciplinary International Symposium on Disinformation in Open Online Media (MISDOOM), 72nd Annual Conference of the International Communication Association (ICA), and 8th International Conference on Computational Social Science (IC2S2) for many helpful comments and suggestions; and to the UC Davis Academic Senate Research Grant and the Stanford Program on Democracy and the Internet (PDI) Research Funding for research support.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Stanford University, Stanford Program on Democracy and the Internet (PDI) Research Funding; University of California > University of California, Davis, Academic Senate Research Grant.

ORCID iDs

Yingdan Lu https://orcid.org/0000-0002-9955-6070
Cuihua (Cindy) Shen https://orcid.org/0000-0003-1645-8211

Supplemental Material

Supplemental material for this article is available online.

Notes

1. See https://m.chinanews.com/wap/detail/chs/zwsp/9356091.shtml
2. See https://m.thepaper.cn/baijiahao_15680224 and https://www.163.com/news/article/FRKLDRSS00019OH3.html
3. Due to the pervasive control of social media in China, none of the social media accounts, including fact-checking accounts, are fully independent of government influence. To reflect the mixture of government and non-government fact-checkers on Douyin, we included two government-sponsored fact-checking accounts (*Chinese Internet Rumor-debunking Platform* and *Scientific Facts*) and seven independent accounts into our dataset. The themes of these accounts differed, with some accounts focusing on health-related issues and others tackling scientific misperceptions. None of the videos in our dataset concerned political issues. The popularity differed across accounts, with account followers ranging from 70,000 to 20 million followers.

4. The 16 keywords we used for filtering are: 谣 (rumor), 谣言 (rumor), 流言 (rumor), 真相 (truth), 传言 (rumor), 真的 (real), 假的 (fake), 伪 (fake), 不实 (fake), 错了 (wrong), 事实 (truth), 真新闻 (real news), 假新闻 (fake news), 不一定 (not really), 否认 (deny), 小道消息 (hearsay).
5. The videos we collected have 25 to 30 frames per second. For each video, we sampled and computed every sixth frame for analysis to reduce computational costs.
6. We used the VideoCapture function to unpack each video into frames, and the COLOR_BGR2LAB function to convert the color space from BGR to CIELAB for lightness score. As the videos are read in 8-bit format, the brightness ranges from 0 to 255 due to the conversion. More explanations of the coordinates in CIELAB can be seen at: https://www.colourphil.co.uk/lab_lch_colour_space.shtml.
7. We used the CSS2.1 base colors in this study. See https://www.w3.org/TR/CSS21/syndata.html#color-units. For pixels that do not have the exact values of the base colors, we matched them to one of the base colors by Euclidean distance. We used the WebColors library (see https://pypi.org/project/webcolors) to convert base color names to RGB values for pixel comparison.
8. See https://www.tukuppt.com/peiyueso/gaoxiao.html?plan=10030-6040-1451375&bd_vid=8560526680965312777.
9. We first used the high-accuracy OCR module in the Baidu AI platform (https://ai.baidu.com/tech/ocr) to derive the text output. Human coders then validated the outputs and corrected the erroneous ones.
10. The Douyin Trending Video Dataset was collected from 9,946 trending topics on Douyin Trending Page (抖音热搜榜) between 18 March 2020 and 17 June 2020. The Trending videos were selected by the Douyin algorithm based on trending topics. These videos are automatically played when a user clicks on a trending topic. As it is impossible to collect all videos from Douyin as a benchmark, we use these algorithmically selected and highly visible videos as a global benchmark. More details of this dataset can be found here: https://computationalcommunication.org/ccr/article/view/110.

References

Abidin, C. (2021). Mapping Internet celebrity on TikTok: Exploring attention economies and visibility labours. *Cultural Science Journal*, *12*(1), 77–103.

Agarwal, S. (2021). *Detecting synthetic faces by understanding real faces* [Doctoral dissertation, University of California, Berkeley].

Applegate, M., & Cohen, J. (2016). Communicating diagrammatically: Mimesis, visual language, & commodification as culture. *Cultural Politics*, *13*, 81–100.

Auxier, B., & Anderson, M. (2021, April 7). *Social media use in 2021*. Pew Research Center: Internet, Science & Tech. https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/

Baldevbhai, P. J., & Anand, R. S. (2012). Color image segmentation for medical images using L* a* b* color space. *IOSR Journal of Electronics and Communication Engineering*, *1*(2), 24–45.

Barriga, C. A., Shapiro, M. A., & Fernandez, M. L. (2010). Science information in fictional movies: Effects of context and gender. *Science Communication*, *32*(1), 3–24.

Basch, C. H., Fera, J., Pierce, I., & Basch, C. E. (2021). Promoting mask use on TikTok: Descriptive, cross-sectional study. *JMIR Public Health and Surveillance*, *7*(2), Article e26392.

Bennett, B. (2012). *Logically fallacious: The ultimate collection of over 300 logical fallacies* (Academic ed.). Ebookit.com.

Bursztynsky, J. (2021, September 27). TikTok says 1 billion people use the app each month. *CNBC*. https://www.cnbc.com/2021/09/27/tiktok-reaches-1-billion-monthly-users.html

Chakraborty, A., Paranjape, B., Kakarla, S., & Ganguly, N. (2016, August). Stop clickbait: Detecting and preventing clickbaits in online news media. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), San Francisco, CA, USA* (pp. 9–16). IEEE.

Chan, M. P. S., Jones, C. R., Jamieson, K. H., & Albarracín, D. (2017). Debunking: A meta-analysis of the psychological efficacy of messages countering misinformation. *Psychological Science*, *28*(11), 1531–1546.

Chen, K., Kim, S. J., Gao, Q., & Raschka, S. (2022). Visual framing of science conspiracy videos: Integrating machine learning with communication theories to study the use of color and brightness. *Computational Communication Research*, *4*(1), 98–134.

Chen, Q., Min, C., Zhang, W., Ma, X., & Evans, R. (2021). Factors driving citizen engagement with government TikTok accounts during the COVID-19 pandemic: Model development and analysis. *Journal of Medical Internet Research*, *23*(2), Article e21463.

Chen, X., Valdovinos Kaye, D. B., & Zeng, J. (2021). #PositiveEnergy Douyin: Constructing “playful patriotism” in a Chinese short-video application. *Chinese Journal of Communication*, *14*(1), 97–117.

Chen, Z., He, Q., Mao, Z., Chung, H. M., & Maharjan, S. (2019, May). A study on the characteristics of douyin short videos and implications for edge caching. In *Proceedings of the ACM Turing Celebration Conference-China* (pp. 1–6). Association for Computing Machinery, New York, NY, United States.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, *3*(3), 149–210.

Culjak, I., Abram, D., Pribanic, T., Dzapo, H., & Cifrek, M. (2012, May). A brief introduction to OpenCV. In *2012 proceedings of the 35th International Convention MIPRO*, Opatija, Croatia (pp. 1725–1730). IEEE.

de Wit, S., Barker, R. A., Dickinson, A. D., & Cools, R. (2011). Habitual versus goal-directed action control in Parkinson disease. *Journal of Cognitive Neuroscience*, *23*(5), 1218–1229.

Duerr, W. (1981). Rhythm in music: A formal scaffolding of time. In J. T. Fraser (Ed.), *The voices of time* (pp. 180–201). The University of Massachusetts Press.

ElShafie, S. J. (2018). Making science meaningful for broad audiences through stories. *Integrative and Comparative Biology*, *58*(6), 1213–1223.

Farinella, M. (2018). The potential of comics in science communication. *Journal of Science Communication*, *17*(1), Article Y01. https://doi.org/10.22323/2.17010401

Fowler, L. R., Schoen, L., Smith, H. S., & Morain, S. R. (2022). Sex education on TikTok: A content analysis of themes. *Health Promotion Practice*, *23*, 739–742. https://doi.org/10.1177/15248399211031536

Geitgey, A. (2020). *Ageitgey/face_recognition: The world's simplest facial recognition API for python and the command line*. https://github.com/ageitgey/face_recognition

Guo, P. J., Kim, J., & Rubin, R. (2014, March). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the First ACM Conference on Learning@ Scale Conference* (pp. 41–50).

Hameleers, M., Powell, T. E., Van Der Meer, T. G., & Bos, L. (2020). A picture paints a thousand lies? The effects and mechanisms of multimodal disinformation and rebuttals disseminated via social media. *Political Communication*, *37*(2), 281–301.

Huang, Z. (1997, February). Clustering large data sets with mixed numeric and categorical values. In *Proceedings of the 1st Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)* (pp. 21–34). Association for Computing Machinery, New York, NY, United States.

Humphreys, A. (2016). *Social media: Enduring principles*. Oxford University Press.

Hwang, S., Liu, X., & Srinivasan, K. (2021). *Voice analytics of online influencers—Soft selling in branded videos*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3773825

Joo, J., & Steinert-Threlkeld, Z. C. (2018). *Image as data: Automated visual content analysis for political science*. https://arxiv.org/pdf/1810.01544.pdf

Kaye, D. B. V., Chen, X., & Zeng, J. (2021). The co-evolution of two Chinese mobile short video apps: Parallel platformization of Douyin and TikTok. *Mobile Media & Communication*, *9*(2), 229–253.

Kellaris, J. J., & Rice, R. C. (1993). The influence of tempo, loudness, and gender of listener on responses to music. *Psychology & Marketing*, *10*(1), 15–29.

Kemnitz, T. M. (1973). The cartoon as a historical source. *Journal of Interdisciplinary History*, *4*(1), 81–93.

Kim, H. S., Suh, Y. J., Kim, E. M., Chong, E., Hong, H., Song, B., . . .Choi, J. S. (2022). Fact-checking and audience engagement: A study of content analysis and audience behavioral data of fact-checking coverage from news media. *Digital Journalism*, *10*(5), 781–800.

Kizilcec, R. F., Papadopoulos, K., & Sritanyaratana, L. (2014, April). Showing face in video instruction: Effects on information retention, visual attention, and affect. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 2095–2102). Association for Computing Machinery, New York, NY, United States.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., . . .Zittrain, J. L. (2018). The science of fake news. *Science*, *359*(6380), 1094–1096.

Li, Y., Guan, M., Hammond, P., & Berrey, L. E. (2021). Communicating COVID-19 information on TikTok: A content analysis of TikTok videos from official accounts featured in the COVID-19 information hub. *Health Education Research*, *36*(3), 261–271.

Li, Y., & Xie, Y. (2020). Is a picture worth a thousand words? An empirical study of image content and social media engagement. *Journal of Marketing Research*, *57*(1), 1–19.

Liu, Y., Liu, G., Wei, D., Li, Q., Yuan, G., Wu, S., Wang, G., & Zhao, X. (2018). Effects of musical tempo on musicians' and non-musicians' emotional experience when listening to music. *Frontiers in Psychology, 9*, Article 2118. https://doi.org/10.3389/fpsyg.2018.02118

Loewenstein, G. (1994). The psychology of curiosity: A review and reinterpretation. *Psychological Bulletin*, *116*(1), 75–98.

Lu, Y., & Pan, J. (2021). Capturing clicks: How the Chinese government uses clickbait to compete for visibility. *Political Communication*, *38*(1-2), 23–54.

Lu, Y., & Pan, J. (2022). The pervasive presence of Chinese Government content on Douyin trending videos. *Computational Communication Research*, *4*(1), 68–97.

Margolin, D. B., Hannak, A., & Weber, I. (2018). Political fact-checking on Twitter: When do corrections have an effect? *Political Communication*, *35*(2), 196–219.

Marmolin, H. (1992). Multimedia from the perspectives of psychology. In L. Kjelldahl (Ed.), *Multimedia* (pp. 39–52). Springer.

McAuley, J. D. (2010). Tempo and rhythm. In M. R. Jones, R. R. Fay, & A. N. Popper (Eds.), *Music perception* (pp. 165–199). Springer Science + Business Media. https://doi.org/10.1007/978-1-4419-6114-3_6

McFee, B., Raffel, C., Liang, D., Ellis, D. P., McVicar, M., Battenberg, E., & Nieto, O. (2015, July). Librosa: Audio and music signal analysis in Python. In *Proceedings of the 14th Python in Science Conference* (Vol. 8, pp. 18–25). SCIPY.

Nasar, B. F., Sajini, T., & Lason, E. R. (2020, December). Deepfake detection in media files: Audios, images and videos. In *2020 IEEE Recent Advances in Intelligent Computational Systems (RAICS)* (pp. 74–79). IEEE.

Negreira-Rey, M. C., Vázquez-Herrero, J., & López-García, X. (2022). Blurring boundaries between journalists and tiktokers: Journalistic role performance on TikTok. *Media and Communication*, *10*(1), 146–156.

Nyhuis, D., Ringwald, T., Rittmann, O., Gschwend, T., & Stiefelhagen, R. (2021). Automated video analysis for social science research. In U. Engel, A. Quan-Haase, S.X. Liu, & L. Lyberg (Eds.), *Handbook of computational social science* (pp. 386–398). Routledge.

Opgenhaffen, M. (2022). Fact-checking interventions on social media using cartoon figures: Lessons learned from "the Tooties." *Digital Journalism*, *10*(5), 888–911.

Oxley, D. R., Smith, K. B., Alford, J. R., Hibbing, M. V., Miller, J. L., Scalora, M., Hatemi, P. K., & Hibbing, J. R. (2008). Political attitudes vary with physiological traits. *Science*, *321*(5896), 1667–1670.

Park, H., Reber, B. H., & Chon, M. G. (2016). Tweeting as health communication: Health organizations' use of Twitter for health promotion and public engagement. *Journal of Health Communication*, *21*(2), 188–198.

Patiño-Escarcina, R. E., & Costa, J. A. F. (2008, August). The semantic clustering of images and its relation with low level color features. In *2008 IEEE International Conference on Semantic Computing* (pp. 74–79). IEEE, Santa Monica, CA, United States.

Peng, Y., & Jemmott, J. B., III (2018). Feast for the eyes: Effects of food perceptions and computer vision features on food photo popularity. *International Journal of Communication*, *12*, 313–336.

Pennycook, G., & Rand, D. G. (2019). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences*, *116*(7), 2521–2526.

Qian, J., Zhang, K., Wang, K., Li, J., & Lei, Q. (2018). Saturation and brightness modulate the effect of depth on visual working memory. *Journal of Vision*, *18*(9), 1–12.

Sangalang, A., Ophir, Y., & Cappella, J. N. (2019). The potential for narrative correctives to combat misinformation. *Journal of Communication*, *69*(3), 298–319.

Schindler, A., Gordea, S., & van Biessum, H. (2016, October). The Europeana sounds music information retrieval pilot. In M. Ioannides, E. Fink, A. Moropoulou, M. Hagedorn-Saupe, A. Fresa, G. Liestøl, V. Rajcic, P. Grussenmeyer (Eds.), *Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection*. Springer, Cham. https://doi.org/10.1007/978-3-319-48974-2_13

Schmid, P., & Betsch, C. (2019). Effective strategies for rebutting science denialism in public discussions. *Nature Human Behaviour*, *3*(9), 931–939.

Shelby, A., & Ernst, K. (2013). Story and science: How providers and parents can utilize storytelling to combat anti-vaccine misinformation. *Human Vaccines & Immunotherapeutics*, *9*(8), 1795–1801.

Shen, Z., Tan, S., & Pritchard, M. J. (2022). Understanding the effects of visual cueing on social media engagement with YouTube educational videos. *IEEE Transactions on Professional Communication*. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=9757141

Shreedhar, G. (2021). Evaluating the impact of storytelling in Facebook advertisements on wildlife conservation engagement: Lessons and challenges. *Conservation Science and Practice*, *3*(11), Article e534.

Siwakoti, S., Yadav, K., Bariletto, N., Zanotti, L., Erdogdu, U., & Shapiro, J. N. (2021). How COVID drove the evolution of fact-checking. *Harvard Kennedy School Misinformation Review*. https://misinforeview.hks.harvard.edu/article/how-covid-drove-the-evolution-of-fact-checking/

Smith, C. A. (2019). Weaponized iconoclasm in Internet memes featuring the expression “fake news.” *Discourse & Communication*, *13*(3), 303–319.

Song, S., Xue, X., Zhao, Y. C., Li, J., Zhu, Q., & Zhao, M. (2021). Short-video apps as a health information source for chronic obstructive pulmonary disease: Information quality assessment of TikTok videos. *Journal of Medical Internet Research*, *23*(12), Article e28318.

Southerton, C. (2021). Lip-syncing and saving lives: Healthcare workers on TikTok. *International Journal of Communication*, *15*, 3248–3268.

Sundar, S. S. (2000). Multimedia effects on processing and perception of online news: A study of picture, audio, and video downloads. *Journalism & Mass Communication Quarterly*, *77*(3), 480–499.

Sundar, S. S., & Limperos, A. M. (2013). Uses and grats 2.0: New gratifications for new media. *Journal of Broadcasting & Electronic Media*, *57*(4), 504–525.

Suryawanshi, S., Chakravarthi, B. R., Arcan, M., & Buitelaar, P. (2020, May). Multimodal meme dataset (MultiOFF) for identifying offensive content in image and text. In Proceedings of the Second Workshop on Trolling, Aggression and Cyberbullying (pp. 32–41). European Language Resources Association (ELRA), Marseille, France.

Tan, C., Lee, L., & Pang, B. (2014). *The effect of wording on message propagation: Topic- and author-controlled natural experiments on Twitter* [Paper presentation]. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Vol. 1: Long Papers). https://doi.org/10.3115/v1/p14-1017

Tully, M., Bode, L., & Vraga, E. K. (2020). Mobilizing users: Does exposure to misinformation and its correction affect users’ responses to a health misinformation post? *Social Media+ Society*, *6*(4), 2056305120978377.

Tuters, M., & Hagen, S. (2020). (((They))) rule: Memetic antagonism and nebulous othering on 4chan. *New Media & Society*, *22*(12), 2218–2237.

van der Meer, T. G., & Jin, Y. (2020). Seeking formula for misinformation treatment in public health crises: The effects of corrective information type and source. *Health Communication*, *35*(5), 560–575.

Vraga, E. K., & Bode, L. (2017). Using expert sources to correct health misinformation in social media. *Science Communication*, *39*(5), 621–645.

Vraga, E. K., Bode, L., & Tully, M. (2022). The effects of a news literacy video and real-time corrections to video misinformation related to sunscreen and skin cancer. *Health Communication*, *37*, 1622–1630.

Vraga, E. K., Kim, S. C., & Cook, J. (2019). Testing logic-based and humor-based corrections for science, health, and political misinformation on social media. *Journal of Broadcasting & Electronic Media*, *63*(3), 393–414.

Vraga, E. K., Kim, S. C., Cook, J., & Bode, L. (2020). Testing the effectiveness of correction placement and type on Instagram. *The International Journal of Press/Politics*, *25*(4), 632–652.

Wang, Y. (2020). Humor and camera view on mobile short-form video apps influence user experience and technology-adoption intent, an example of TikTok (DouYin). *Computers in Human Behavior, 110*, Article 106373.

Wei, W., & Wan, X. (2017). *Learning to identify ambiguous and misleading news headlines* [Conference session]. Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence. https://doi.org/10.24963/ijcai.2017/583

Williams, N. W., Casas, A., & Wilkerson, J. D. (2020). *Images as data for social science research: An introduction to convolutional neural nets for image classification*. Cambridge University Press.

Wilms, L., & Oberfeld, D. (2018). Color and emotion: Effects of hue, saturation, and brightness. *Psychological Research*, *82*(5), 896–914.

Wittenberg, C., Zong, J., & Rand, D. (2020). *The (minimal) persuasive advantage of political video over text* (Working Paper). https://psyarxiv.com/r5yun/download?format=pdf

Wong, B. (2021, October 12). Meet the medical experts debunking COVID misinformation on TikTok. *Yahoo! Sports*. https://sports.yahoo.com/meet-medical-experts-debunking-covid-181044900.html

Woolley, S. C. (2016). Automating power: Social bot interference in global politics. *First Monday*, *21*(4). https://firstmonday.org/article/view/6161/5300

Xinhuanet. (2021). 中国互联网联合辟谣平台上线三周年_新华网 [Chinese internet rumor-debunking platform coming to its third anniversary]. http://www.xinhuanet.com/politics/ldzt/pyptszn/index.htm

Yang, X., Ram, N., Robinson, T., & Reeves, B. (2019). Using screenshots to predict task switching on smartphones. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–6). Association for Computing Machinery, New York, NY, United States.

Young, D. G., Jamieson, K. H., Poulsen, S., & Goldring, A. (2018). Fact-checking effectiveness as a function of format and tone:

Evaluating FactCheck.org and FlackCheck.org. *Journalism & Mass Communication Quarterly*, *95*(1), 49–75. https://doi.org/10.1177/1077699017710453

Zhang, J., Featherstone, J. D., Calabrese, C., & Wojcieszak, M. (2021). Effects of fact-checking social media vaccine misinformation on attitudes toward vaccines. *Preventive Medicine, 145*, Article 106408.

Zhang, Q., Wang, W., & Chen, Y. (2020). Frontiers: In-consumption social listening with moment-to-moment unstructured data: The case of movie appreciation and live comments. *Marketing Science*, *39*(2), 285–295.

Zulli, D., & Zulli, D. J. (2022). Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform. *New Media & Society*, *24*, 1872–1890. https://doi.org/10.1177/1461444820983603

Author Biographies

Yingdan Lu is a PhD candidate in the Department of Communication at Stanford University. Her research applies data analytic and computational methods to large-scale multimodal datasets to explore how authoritarian governments use digital media to maintain their rule and how individuals experience digital technology in different media environments. For more information, see her website: https://yingdanlu.com/.

Cuihua (Cindy) Shen (PhD, University of Southern California) is a professor at the Department of Communication and co-director of the Computational Communication Research Lab at University of California, Davis. Her research interests include online communities, misinformation, and computational social science research methods.