Perception and deception: Exploring individual responses to deepfakes across different modalities

Saifuddin Ahmed [*], Hui Wen Chua

Nanyang Technological University, Singapore

A R T I C L E   I N F O

A B S T R A C T

This study is one of the first to investigate the relationship between modalities and individuals' tendencies to believe and share different forms of deepfakes (also deep fakes). Using an online survey experiment conducted in the US, participants were randomly assigned to one of three disinformation conditions: video deepfakes, audio deepfakes, and cheap fakes to test the effect of single modality against multimodality and how it affects individuals' perceived claim accuracy and sharing intentions. In addition, the impact of cognitive ability on perceived claim accuracy and sharing intentions between conditions are also examined. The results suggest that individuals are likelier to perceive video deepfakes as more accurate than cheap fakes, but not audio deepfakes. Yet, individuals are more likely to share video deepfakes than cheap and audio deepfakes. We also found that individuals with high cognitive ability are less likely to perceive deepfakes as accurate or share them across formats. The findings emphasize that deepfakes are not monolithic, and associated modalities should be considered when studying user engagement with deepfakes.

1. Introduction

Numerous studies confirm that social media increases citizens' exposure to news both actively and incidentally [1], improving factual political knowledge [2], facilitating political expression [3], and ultimately paving the way for broader civic participation [4]. However, in recent years, growing misinformation on social media has threatened democratic ideals. Misinformation exists in disparate forms ranging from audio clips only, fake news to cheapfakes. However, the deleterious effects of social media on democracy extend beyond the common forms of misinformation. For example, the recent rise of deepfakes and their inimical threat is a harbinger of a possible "information apocalypse", where citizens may find it hard to discern truth from falsehood [5].

Deepfakes are a relatively new form of misinformation harboring the malicious intent to deceive individuals using cutting-edge artificial intelligence (AI) and nefarious actors to mimic real-life events. This also means that one could be portrayed in a deepfake doing or saying something they did not do without their consent, as exemplified by the many instances of deepfake pornography [6]. Deepfakes threaten organizations as they risk privacy and security [7]. Well-made deepfakes can erode public trust in government [8] and public figures [5] and disturb political stability [9,10]. Exposure to deepfakes can also raise skepticism about news content [11,12] and distrust of news [13].

The elusive dangers of deepfakes have gained much scholarly attention [5,6,8,11,12,14]. However, limited studies directly

---

* Corresponding author. University of California-Davis) is an Assistant Professor at the Wee Kim Wee School of Communication and Information, Nanyang Technological University, Nanyang Technological University, Singapore 637718.

E-mail address: sahmed@ntu.edu.sg (S. Ahmed).

compare differences in individuals' vulnerabilities to different forms of deepfakes, including video deepfakes, video cheapfakes, and audio deepfakes. Cheapfakes are a relatively new audio-visual (AV) manipulation involving cheap, accessible software or none at all [15].

Additionally, we are unaware of which modality plays a more significant role in influencing one's perceived claim accuracy or sharing intentions of deepfakes. Hence, to contribute to the existing literature, this study investigates if individuals are more vulnerable and are likely to share video deepfakes against other forms of deepfakes (e.g., audio deepfakes and cheapfakes). Further, given the importance of cognitive ability in misinformation discernment [16,17], we also explore whether cognitive ability influences the vulnerabilities to different forms of deepfakes (i.e., video deepfakes, video cheapfakes, and audio deepfakes).

1.1. Modality, perceived claim accuracy, and sharing intention of deepfakes

Each modality contains unique characteristics, and individuals encode modality-specific content when processing information [18]. Therefore, combining multiple modalities can provide complementary information and more robust inferences, as demonstrated by its vivid applications in AV speech recognition [19], emotion recognition [20], and language and vision tasks [21]. Hence, in a multimodal content like deepfake, many modalities such as facial cues [22], hand gestures [23], body posture [24], and speech cues [25], and tone of the voice [26] can interact to influence one's perceived claim accuracy.

Past research also lends support as human information processing is proven to be limited by attentional resources [27]. This means that humans selectively choose a certain quantity of sensory input to process while the other sensory inputs are neglected. Therefore, the amount of attention paid affects the amount of information absorbed and processed by an individual. Given that an individual has a limited capacity to comprehend informational messages, attempting to process "too much" information may exceed one's limit, failing to comprehend the information [28]. As information processing is critical in the decision-making process, this accentuates the need to compare different modalities as the amount of attention paid to each modality differs and may significantly influence how they process the information.

Even though deepfakes leverage the advantage of boosted believability through their extended span of modalities. Scholars have found that dissonance between the visual characteristics (e.g., loss of lip-syncing, unnatural facial expressions) tends to capture users' attention and act as an indicator of content veracity [29]. Blatant mismatches are frequent in cheapfakes that cannot achieve seamless audio and visual sync. Research has shown that dissonant AV elements induce a higher user cognitive load and are less likely to perceive the two modalities as a single stimulus, resulting in lower believability [30].

Many studies have explored the realms of video modality and strived to provide a plausible explanation for why video modality is more persuasive and shared. However, more evidence is needed for audio clips. The power of audio lies in its characteristics - speech rate, pitch, volume, and tone, all of which can influence the speaker's perceptions. For example, fast speakers are typically seen as more competent, truthful, and persuasive [31]. Conversely, signs of stress in a speaker's voice may be perceived as less competent and persuasive [32] - which may explain why some users may mistake real audio clips as fake as people may be stressed while speaking in real life. On the other hand, extreme levels of emotion in vocal tone may backfire as individuals tend to perceive that as a lack of realism [33,34].

Nevertheless, we lack empirical evidence for a direct comparison between these multimodal (video deepfakes and cheapfakes) and single-modality misinformation (deepfake audio clips) to determine if having multiple modalities is more influential in users' decision to believe in the media. Therefore, lending support from past studies on the importance of sensory modalities in the decision-making process, we hypothesize.

H1. Individuals are more likely to perceive the fabricated claims in video deepfakes as accurate than in cheapfakes and audio deepfakes.

Beyond the perceived accuracy of different forms of deepfakes, we also enquire about individual intentions to share the content. Scholars found that individuals are more likely to perceive fake news as credible when presented in a video format than audio and text due to the positive credibility effect [35]. This effect is further strengthened when individuals' issue involvement is low as they will pay less attention and process the information superficially. Likewise, relatively less informed individuals will be more susceptible to believing fake news. As people perceive misinformation as accurate, they tend to believe more in it and become more willing to share it with others online [14,36], accelerating the dissemination of fake news online. The cyclic nature of this spread of online falsehood is extremely minacious to the heavy consumers of online information. While some have concluded that sensory modalities of fake news are noteworthy in influencing one's perceived accuracy and sharing intentions on fake news [35], no studies have paid attention to its role in the sharing intention of deepfakes. This lends credibility to our research direction, and hence, we aim to answer the following research question.

RQ1. Are individuals more likely to share video deepfakes over cheap fakes and audio deepfakes?

1.2. The contingent role of cognitive ability

Cognitive ability is general intelligence and includes the capacity to "reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly, and learn from experience" [37]. Past research has posited that most individuals would not apply their cognitive skills when evaluating information unless needed [38]. The limited capacity model posits that individuals focus on only a few salient features due to limited cognitive space rather than processing all available features in a message when encoding new information [28]. The same applies when people evaluate a message's credibility [39]. Hence, it is essential to consider cognitive ability in

our study as it affects the users' evaluation of misinformation on social media and how much they believe or do not believe in it, directly impacting their sharing intentions.

The 'realism heuristic' in deepfakes [6] poses a new challenge to social media users, further blurring the line between reality and falsity. Recent research has observed that users follow a 'seeing-is-believing' heuristic [40], where most users mistake deepfakes for authentic videos and have overconfidence in their detection abilities [41]. This, when paired with how most online users lack critical thinking and conscious awareness when reading online news as they consume in a state of automaticity [42], magnifies the risk of deepfakes.

However, for a subset of users, seeing is only sometimes believing. Individuals with higher cognitive ability are better at discerning misinformation from real news [16], and those who believe in misinformation rarely doubt its content [17]. Cognitive ability has also been found to be positively associated with effective information processing [43], better decision-making [44], more healthy skepticism toward unfounded beliefs [45], and better capabilities in risk assessment pertinent to their decision of where and whether they should place their trust [46]. In terms of sharing intentions, low-cognitive individuals are more likely to share deepfakes [36,41]. This is sufficient to assume that cognitive ability plays a core part in buffering against manipulative deepfakes. Given that individuals with higher cognitive ability display healthier skepticism towards unreliable content, we propose our second hypothesis.

H2. Higher cognitive ability individuals are less likely to a) believe and b) share video deepfakes, cheapfakes, and audio deepfakes.

2. Method

2.1. Sample and measures

Using a quota sampling strategy, we recruited a sample of US residents through Qualtrics (N = 309). We then matched the sample frame to population parameters focusing on age and gender for greater representativeness of our findings. The study was approved by the Institutional Review Board at Nanyang Technological University (#635). Informed consent was undertaken from all participants, and the study complies with all the regulations and confirmation.

In the first step, the participants answered questions about their demographic characteristics, political motivations, and cognitive ability. Next, they were randomly assigned to one of the three deepfake conditions: a) video deepfake (n = 105), b) audio deepfake (n = 102), or c) cheap fake (n = 102). Then, the participants watched the video or listened to the audio and answered the questions
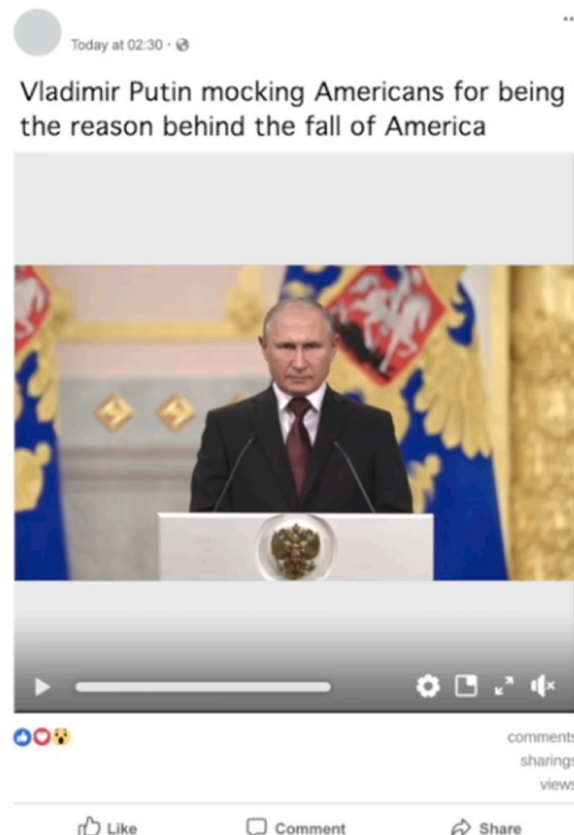

Fig. 1. A screenshot of the video deepfake.

regarding their perceived accuracy of the claims and sharing intentions. The participants were only informed about the nature of the stimulus at the end.

We use a real-world deepfake video of Vladimir Putin (created by RepresentUS) to increase the validity of our findings. The original deepfake was edited to create a cheapfake (video distorted) and audio-deepfake (video layer removed). A screenshot of the video deepfake is included in Fig. 1.

We also explored if there are characteristic differences between the three experimental conditions. ANOVA and Chi-square tests suggest that the three conditions do not differ across their demographic characteristics, political motivations, or cognitive ability. The results are presented in Appendix A.

The perceived claim accuracy of misinformation was measured by asking respondents to rate their level of agreement (1 = not at all accurate to 5 = extremely accurate) for the claim present in the stimulus "Vladimir Putin mocked Americans for being the reason behind the fall of America?".

Sharing intention was measured by asking respondents how likely (1 = extremely likely to 5 = not at all) they are to share the video/audio on their social media. The response options were reversed, so a higher value represents greater sharing intentions.

Cognitive ability was measured by the word sum test, where participants were required to match the source word to a closely associated word from a target list of five words. The test includes ten questions. The test shares a high variance with general intelligence and is frequently used to measure the cognitive ability of individuals [47,48]. The correct responses to the ten items were summed to create a scale of cognitive ability (M = 4.99, SD = 2.47, α = 0.74).

We also use demographic characteristics and partisanship as controls. Demographics included age (M = 46.05, SD = 18.24), gender (53% female), education (Median = bachelor's degree), income (Median = $5000 to $6999), and race (73% White). Political motivations include participants' political interest (M = 3.26, SD = 1.21, 1 = not at all interested to 5 = extremely interested) and partisanship (M = 3.58, SD = 2.18; 1 = strong democrat to 7 = strong republican).

3. Results

We compared the differences between conditions for perceived claim accuracy - there was a significant effect of conditions on perceived claim accuracy [F (2, 298) = 3.37, p < .05]. Furthermore, post hoc tests showed that individuals were more likely to perceive the deepfake (M = 3.06, SE = 0.13) to be more accurate than the cheap fake (M = 2.55, SE = 0.13, p < .01) but not the audio-deepfake (M = 2.78, SE = 0.14, p = .14). Please refer to Fig. 2 for the means plot. In addition, we also found the main effect of political interest (F = 14.77, p < .01) and cognitive ability (F = 12.33, p < .01), which are later explored through regression models.

Next, similar to the test of perceived accuracy, we ran another model testing the differences between the conditions for sharing intention. There was a significant effect of conditions on sharing intention [F (2, 298) = 4.34, p < .05]. Post hoc tests suggest that those who watched the video deepfake (M = 2.51, SE = 0.13) were more likely to share it than audio deepfakes (M = 2.07, SE = 0.13, p < .05) and cheapfakes (M = 2.04, SE = 0.12, p < .001). See Fig. 2 for the means plot. We also observed the main effects of cognitive ability (F = 25.62, p < .001).

We further tested the role of cognitive ability in perceived claim accuracy and sharing intention through regression models (see
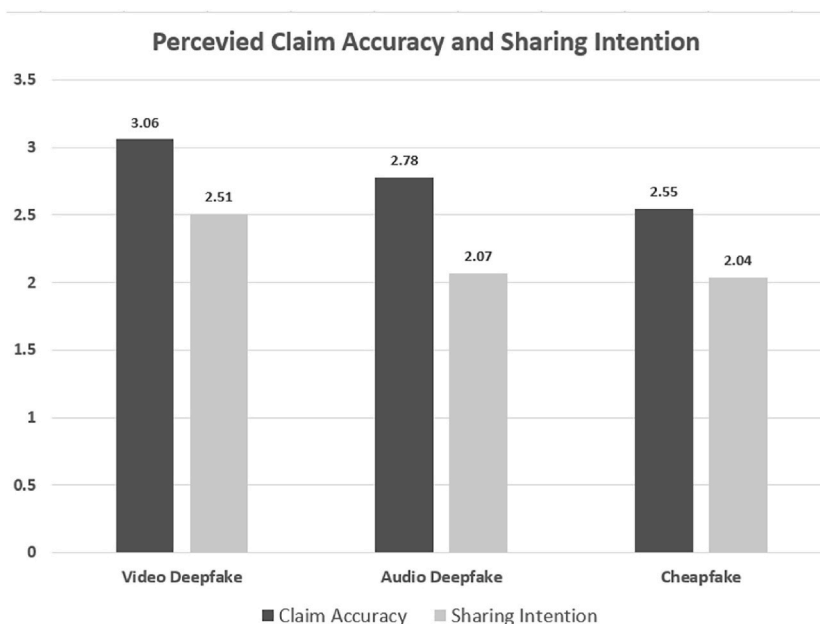

Fig. 2. The means plot for perceived claim accuracy and sharing intention across conditions.

Table 1

Regression examining perceived claim accuracy and sharing intention.

	Perceived Claim Accuracy	Sharing Intention
Predictor variable	β	β
Age	−.087	−.166*
Gender (males = 0)[d]	.060	−.081
Education	.062	.033
Income	.057	.018
Race (White = 1)[d]	.015	−.059
Political interest	.162**	.002
Partisanship	.026	−.095
Cognitive ability	−.154***	−.243***
Audio-deepfake[d] (ref = video deepfake)	−.096	−.120*
Cheap fake[d] (ref = video deepfake)	−.165**	−.115*
Perceived claim accuracy	–	.261***
Total R^2	9.5	27.3

Note: *p < .05; **p < .01; ***p < .001; d = dummy.

Table 1). Two models were constructed with perceived claim accuracy and sharing intention as dependent variables. The results presented in Table 1 suggest that those with high cognitive ability are less likely to perceive the claims to be accurate (claim: β = −0.154, p < .001) and share them on social media (sharing: β = −0.243, p < .001). In addition, we also found that perceived claim accuracy is positively associated with sharing intention (β = 0.261, p < .001), suggesting that individuals are more likely to share misinformation across AV formats if they believe it to be true.

Finally, we ran two-way ANCOVAs to test the interaction between cognitive ability and different conditions. However, no significant effects were found for either perceived accuracy [F (2, 296 = 1.21, p = .30) or sharing intentions [F (2, 296 = 0.75, p = .48). The means plot is included in Appendix B.

4. Discussion

Our analysis highlights the significance of source and consumer characteristics in user differences in engagement with deepfakes. The findings are discussed in detail below.

First, individuals are more likely to perceive video deepfakes as more accurate than cheapfakes but not audio deepfakes. The results emphasize and reinforce the deceptive notion of deepfakes, highlighting the importance of realism heuristics in AV content. The realism heuristics stemming from the well-made nature of deepfakes may influence the fluency of individuals when they process the information. For example, familiarity elicits a 'truthiness effect,' which states that people are more likely to accept information as accurate if perceived as familiar [49,50]. This effect may intensify if the portrayed figure in the deepfake is well-known, such as the one in our deepfake condition - Vladimir Putin. As a stronger sense of familiarity is positively associated with higher fluency, people are more likely to perceive it as accurate regardless of its veracity.

Out of the three deceptive conditions, only the video deepfake and video cheapfakes are multimodal mediums, while the audio deepfake comprises a single modality (audio). Therefore, our findings suggest that the dissonance between AV elements in cheapfakes may be more significant in influencing the perceived claim accuracy than the number and type of modality involved. Past studies have stated the picture superiority effect [51] and that misleading visuals are more duplicitous than illusive verbal content to generate false perceptions based on 'realism heuristics' [52].

Subsequently, we expected that people would perceive video deepfakes and cheapfakes more accurately than audio deepfakes; however, we found that people were more likely to perceive video deepfakes and audio deepfakes (single modality) as accurate but not cheapfakes. Our result challenges the popular claim of ocular centrism, the epistemological prioritization of sight above all other senses [53]. It posits that the richness of presented content does not always translate to a more significant role in influencing perceived accuracy. Recent research also confirms that individuals not only fall for video deepfakes [12,14,36] but also cannot reliably recognize audio deepfakes [54].

An explanation can be offered through the uncanny valley theory, which conceptualizes human affinity towards simulated humans regarding their fidelity to reality. As one encounters a simulated human of a certain fidelity, we pick up cues that indicate its belongingness to a human category and cues that indicate belonging to some non-human category [55]. Then, the trend in human affinity for the simulation is expected to follow a curvilinear trend with depression at "average" fidelity. Therefore, a well-made video deepfake and audio deepfake clip may render cues that affirm its humanness; in comparison, the dissonance in AV elements in cheapfake may have cued viewers into its lack of humanness. Hence, individuals have experienced more cognitive dissonance when viewing the video cheapfake than the video deepfake and audio deepfake, resulting in antipathy for the simulated human and triggering the inauthenticity alarm bells within readers to signify to them that it is fake.

Next, regarding sharing intentions, we found that individuals are likelier to share video deepfakes than video cheapfakes and audio deepfakes. There are two possible explanations for this. Firstly, past research has shown the compelling appeal of multimodal information with psychological mechanisms such as source vividness or realism [52]. The Heuristic-Systematic Model (HSM) of information processing posits that individuals rely on simple heuristics, such as the vividness of information when processing via the peripheral route [56] and have the ability to influence behavioral intentions [57]. Second, false news diffuses faster, possibly due to its

novelty [58]. Traditionally, novelty attracts attention [59]. Therefore, individuals likely prefer to share video deepfake and cheapfakes because they find them richer in information than audio deepfakes, thus drawing more attention. This canbe supported by the fact that the appeal of rich media also lies in its novelty (Osei-Appiah, 2006).

Nevertheless, comparing the two sets of results, we find that, although people are equally likely to perceive video deepfake and audio deepfake to be more accurate than video cheapfake, this trend is different from the results for sharing intentions. While results from one study are far from establishing broader trends, it may signify the imminent need to consider other factors apart from perceived claim accuracy when studying sharing intentions, as belief is not a necessary precursor to sharing [17]. Hence, our result exemplifies that a deepfake is not monolithic and should be studied as a combination of its characteristics.

Lastly, we found that individuals with higher cognitive ability are less likely to perceive deepfakes as accurate or share them. This aligns with existing literature as those with high cognitive ability are often associated with better decision-making processes [43] due to better elaborative processing and truth discernment [16]. High cognitive individuals spend more time on elaborative processes, and while thinking if a message contains persuasive intent, resistance may occur, and heightened suspicion can negatively influence behavioral intentions [60]. The results confirm that cognitive ability can safeguard against all forms of deepfakes. However, our study focuses on a political deepfake, which may have affected how people assess its quality and accuracy. Hence, this finding may only be generalizable to some genres of deepfakes.

Our findings also apply to real-world scenarios and may be helpful for social media companies and policymakers. While most of the societal attention has been on video deepfakes, the results suggest that the risks associated with audio deepfakes are also significant. This is critical, especially when recent empirical evidence points to the vulnerability of individuals to correctly identify the falsehoods in audio deepfakes [54]. In addition, the findings also highlight the susceptibility of low-cognitive individuals to malicious deepfakes. As such, more focused intervention strategies should be adopted to prevent the citizenry from the threats of deepfakes.

5. Limitations

Finally, we recognize some limitations of the study. First, we chose to use a political deepfake in our study. Hence, the findings may not be generalizable to all disinformation contexts. Second, the study is based on a panel of US citizens participating in online surveys. Therefore, the findings may not generalize to the broader US context or other countries. Nevertheless, given that deepfake exposure and engagement happen primarily within the section of the online population, our findings can inform us of the significant vulnerability of this population. More representative studies are necessary for generalizations.

6. Conclusion

To summarize, this is one of the rare studies to explore the role of sensory modalities in influencing perceived claim accuracy and sharing intentions across different forms of deepfakes, and our results provide empirical evidence that scholars should not use a single lens when studying deepfakes because citizen engagement can differ across forms. The findings have applications for technology policy, where social media companies can focus on sensory modalities for fact-checking and alerting users to potentially manipulated content. In governance, policymakers can draft guidelines to safeguard the more vulnerable section of society (e.g., low cognitive users) against misinformation by including informational cues on social media posts to buffer against misinformation.

Author contribution statement

Saifuddin Ahmed conceived and designed the experiments; performed the experiments; analyzed and interpreted the data; and wrote the paper.

Chua Huiwen analyzed and interpreted the data; and wrote the paper.

Data availability statement

Data will be made available on request.

Supplementary content related to this article has been published online at [URL].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e20383.

References

[1] Y. Kim, H.-T. Chen, Y. Wang, Living in the smartphone age: examining the conditional indirect effects of mobile phone use on political participation, J. Broadcast. Electron. Media 60 (4) (2016) 694–713, https://doi.org/10.1080/08838151.2016.1203318.

[2] M.A. Beam, M.J. Hutchens, J.D. Hmielowski, Clicking vs. sharing: the relationship between online news behaviors and political knowledge, Comput. Hum. Behav. 59 (2016) 215–220, https://doi.org/10.1016/j.chb.2016.02.013.

[3] M. Yamamoto, M.J. Kushin, F. Dalisay, Social media and mobiles as political mobilization forces for young adults: examining the moderating role of online political expression in political participation, New Media Soc. 17 (6) (2015) 880–898.

[4] H. Gil de Zúñiga, N. Jung, S. Valenzuela, Social media use for news and individuals' social capital, civic engagement and political participation, J. Computer-Mediated Commun. 17 (3) (2012) 319–336.

[5] M. Westerlund, The emergence of deepfake technology, A Review. Technology Innovation Management Review 9 (11) (2019) 40–53, https://doi.org/10.22215/timreview/128.

[6] J.T. Hancock, J.N. Bailenson, The Social Impact of Deepfakes | Cyberpsychology, Behavior, and Social Networking, 2021. https://www.liebertpub.com/doi/full/10.1089/cyber.2021.29208.jth.

[7] H. Chi, U. Maduakor, R. Alo, E. Williams, Integrating deepfake detection into cybersecurity curriculum, in: Proceedings of the Future Technologies Conference, Springer, Cham, 2020, pp. 588–598.

[8] B. Chesney, D. Citron, Deep fakes: a looming challenge for privacy, democracy, and national security, Calif. Law Rev. 107 (2019) 1753.

[9] E. Pashentsev, The malicious use of deepfakes against psychological security and political stability, in: The Palgrave Handbook of Malicious Use of AI and Psychological Security, Springer International Publishing, Cham, 2023, pp. 47–80.

[10] M. Sareen, Threats and challenges by DeepFake technology, in: DeepFakes, CRC Press, 2022, pp. 99–113.

[11] S. Ahmed, Navigating the maze: deepfakes, cognitive ability, and social media news skepticism, New Media Soc. 25 (5) (2023) 1108–1129.

[12] C. Vaccari, A. Chadwick, Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news, Social Media+ Society 6 (1) (2020), 2056305120903408.

[13] E. Temir, Deepfake: new era in the age of disinformation & end of reliable journalism, Selçuk İletişim 13 (2) (2020) 1009–1024.

[14] S. Ahmed, Fooled by the fakes: cognitive differences in perceived claim accuracy and sharing the intention of non-political deepfakes, Pers. Indiv. Differ. 182 (2021), 111074.

[15] Paris, B., & Donovan, J. (n.d.). Deepfakes and Cheapfakes. Available at: https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1-1.pdf.

[16] G. Pennycook, D.G. Rand, Lazy, not biased: susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning, Cognition 188 (2018) 39–50, https://doi.org/10.1016/j.cognition.2018.06.011.

[17] G. Pennycook, D.G. Rand, Fighting misinformation on social media using crowdsourced judgments of news source quality, Proc. Natl. Acad. Sci. USA 116 (7) (2019) 2521–2526.

[18] H.R. Unnava, R.E. Burnkrant, S. Erevelles, Effects of presentation order and communication modality on recall and attitude, J. Consum. Res. 21 (3) (1994) 481–490.

[19] M. Gurban, J.-P. Thiran, T. Drugman, T. Dutoit, October 20). Dynamic Modality Weighting for Multi-Stream Hmms Inaudio-Visual Speech Recognition | Proceedings of the 10th International Conference on Multimodal Interfaces, 2008. https://dl.acm.org/doi/abs/10.1145/1452392.1452442?casa_token=xiikZknpb84AAAAA:g66fwO_Aobg07TkOFf-XK9wy27G7nRjboJNyADLVfRnSxndHqUMH_lTEgj6oKjzl_CUKbunwCKqo.

[20] T. Mittal, U. Bhattacharya, R. Chandra, A. Bera, D. Manocha, M3er: multiplicative multimodal emotion recognition using facial, textual, and speech cues, Proc. AAAI Conf. Artif. Intell. 34 (No. 02) (2020, April) 1359–1367.

[21] J.P. Bigham, C. Jayant, H. Ji, G. Little, A. Miller, R.C. Miller, R. Miller, A. Tatarowicz, B. White, S. White, T. Yeh, VizWiz | Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology, 2010.

[22] M. Swerts, E. Krahmer, Facial expression and prosodic prominence: effects of modality and facial area, J. Phonetics 36 (2) (2008) 219–238, https://doi.org/10.1016/j.wocn.2007.05.001.

[23] Kjeldsen, F. C. M. (n.d.). Visual interpretation of hand gestures as a practical interface modality [Ph.D., Columbia University]. Retrieved June 29, 2022, from https://www.proquest.com/docview/304345725/abstract/F357CAAE24EC4EBFPQ/1.

[24] C. Navarretta, Individuality in communicative bodily behaviours, in: A. Esposito, A.M. Esposito, A. Vinciarelli, R. Hoffmann, V.C. Müller (Eds.), Cognitive Behavioural Systems, Springer, 2012, pp. 417–423, https://doi.org/10.1007/978-3-642-34584-5_37.

[25] M. Kamachi, H. Hill, K. Lander, E. Vatikiotis-Bateson, Putting the face to the voice': matching identity across modality, Curr. Biol. 13 (19) (2003) 1709–1714.

[26] G. Pourtois, M. Dhar, Integrating Face and Voice in Person Perception, 2013.

[27] B. Wahn, P. König, Is attentional resource allocation across sensory modalities task-dependent? Adv. Cognit. Psychol. 13 (1) (2017) 83–96, https://doi.org/10.5709/acp-0209-.

[28] A. Lang, The limited capacity model of mediated message processing, J. Commun. 50 (1) (2000) 46–70.

[29] P. Gupta, A. Dhall, R. Subramanian, K. Chugh, Not Made for Each Other– Audio-Visual Dissonance-Based Deepfake Detection and Localization, 2021. http://dspace.iitrpr.ac.in:8080/xmlui/handle/123456789/1978.

[30] T. Grimes, Mild auditory-visual dissonance in television news may exceed viewer attentional capacity, Hum. Commun. Res. 18 (2) (1991) 268–298, https://doi.org/10.1111/j.1468-2958.1991.tb00546.x.

[31] A. Oleszkiewicz, M. Karwowski, K. Pisanski, P. Sorokowski, B. Sobrado, A. Sorokowska, Who uses emoticons? Data from 86702 Facebook users, Pers. Indiv. Differ. 119 (2017) 289–295, https://doi.org/10.1016/j.paid.2017.07.034.

[32] W. Apple, L.A. Streeter, R.M. Krauss, Effects of pitch and speech rate on personal attributions, J. Pers. Soc. Psychol. 37 (5) (1979) 715–727, https://doi.org/10.1037/0022-3514.37.5.715.

[33] G. Carlo, M.V. Mestre, M.M. McGinley, P. Samper, A. Tur, D. Sandman, The interplay of emotional instability, empathy, and coping on prosocial and aggressive behaviors, Pers. Indiv. Differ. 53 (5) (2012) 675–680, https://doi.org/10.1016/j.paid.2012.05.022.

[34] S.A. Shields, The politics of emotion in everyday life, Appropriate" Emotion and Claims on Identity (2005). https://journals.sagepub.com/doi/abs/10.1037/1089-2680.9.1.3.

[35] S.S. Sundar, M.D. Molina, E. Cho, Seeing is believing: is video modality more powerful in spreading fake news via online messaging apps? J. Computer-Mediated Commun. 26 (6) (2021) 301–319, https://doi.org/10.1093/jcmc/zmab010.

[36] S. Ahmed, S.W.T. Ng, A.W.T. Bee, Understanding the role of fear of missing out and deficient self-regulation in sharing of deepfakes on social media: evidence from eight countries, Front. Psychol. 14 (2023), 1127507.

[37] R. Plomin, Genetics and general cognitive ability, Nature 402 (6761) (1999) C25–C29, https://doi.org/10.1038/35011520.

[38] S.T. Fiske, S.E. Taylor, Social Cognition: from Brains to Culture, SAGE, 2013.

[39] B.J. Fogg, Prominence-interpretation theory: explaining how people assess credibility online, CHI '03 Extended Abstracts on Human Factors in Computing Systems (2003) 722–723, https://doi.org/10.1145/765891.765951.

[40] N.C. Köbis, B. Doleželová, I. Soraperra, Fooled twice: people cannot detect deepfakes but think they can, iScience 24 (11) (2021), 103364, https://doi.org/10.1016/j.isci.2021.103364.

[41] S. Ahmed, Examining public perception and cognitive biases in the presumed influence of deepfakes threat: empirical evidence of third person perception from three studies, Asian J. Commun. 33 (3) (2023) 308–331.

[42] E.M. Powers, How Students Access, Filter and Evaluate Digital News: Choices that Shape what They Consume and the Implications for News Literacy Education—ProQuest, 2014. https://www.proquest.com/docview/1561141738?pq-origsite=gscholar&fromopenview=true.

[43] M. Lodge, R. Hamill, A partisan schema for political information processing, Am. Polit. Sci. Rev. 80 (2) (1986) 505–519.

[44] C. Gonzalez, R.P. Thomas, P. Vanyukov, The relationships between cognitive ability and dynamic decision making, Intelligence 33 (2) (2005) 169–186, https://doi.org/10.1016/j.intell.2004.10.002.

[45] T. Ståhl, J.-W. van Prooijen, Epistemic rationality: skepticism toward unfounded beliefs requires sufficient cognitive ability and motivation to be rational, Pers. Indiv. Differ. 122 (2018) 155–163, https://doi.org/10.1016/j.paid.2017.10.026.

[46] S. Zmerli, T.W.G. Van der Meer, Handbook on political trust, Edward Elgar Publishing, 2017.

[47] M.J. Brandt, J.T. Crawford, Answering unresolved questions about the relationship between cognitive ability and prejudice, Soc. Psychol. Personal. Sci. 7 (8) (2016) 884–892.

[48] Y. Ganzach, Y. Hanoch, B.L. Choma, Attitudes toward presidential candidates in the 2012 and 2016 American elections: cognitive ability and support for Trump, Soc. Psychol. Personal. Sci. 10 (7) (2019) 924–934.

[49] A.J. Berinsky, Rumors and health care reform: experiments in political misinformation, Br. J. Polit. Sci. 47 (2) (2017) 241–262.

[50] E.J. Newman, M. Garry, C. Unkelbach, D.M. Bernstein, D.S. Lindsay, R.A. Nash, Truthiness and falseness of trivia claims depend on judgmental contexts, J. Exp. Psychol. Learn. Mem. Cognit. 41 (5) (2015) 1337–1348, https://doi.org/10.1037/xlm0000099.

[51] D.L. Nelson, V.S. Reed, J.R. Walling, Pictorial superiority effect, J. Exp. Psychol. Hum. Learn. Mem. 2 (5) (1976) 523–528, https://doi.org/10.1037/0278-7393.2.5.523.

[52] S.S. Sundar, The MAIN Model: A Heuristic Approach to Understanding Technology Effects on Credibility, MacArthur Foundation Digital Media and Learning Initiative, Cambridge, MA, 2008, pp. 73–100.

[53] K. Geddes, Ocularcentrism and deepfakes: should seeing be believing? Fordham Intellect. Prop. Media Entertain. Law J. 31 (4) (2020) 1042.

[54] K.T. Mai, S. Bray, T. Davies, L.D. Griffin, Warning: humans cannot reliably detect speech deepfakes, PLoS One 18 (8) (2023), e0285333.

[55] A.E. Ferrey, T.J. Burleigh, M.J. Fenske, Stimulus-category competition, inhibition, and affective devaluation: a novel account of the uncanny valley, Front. Psychol. 6 (2015).

[56] A.H. Eagly, S. Chaiken, The Psychology of Attitudes, Harcourt Brace Jovanovich College Publishers, 1993.

[57] T.E. Powell, H.G. Boomgaarden, K. De Swert, C.H. de Vreese, A clearer picture: the contribution of visuals and text to framing effects, J. Commun. 65 (6) (2015) 997–1017, https://doi.org/10.1111/jcom.12184.

[58] S. Vosoughi, D. Roy, S. Aral, The spread of true and false news online, Science 359 (6380) (2018) 1146–1151.

[59] L. Itti, P. Baldi, Bayesian surprise attracts human attention, Vis. Res. 49 (10) (2009) 1295–1306, https://doi.org/10.1016/j.visres.2008.09.007.

[60] N.J. Evans, D. Park, Rethinking the persuasion knowledge model: schematic antecedents and associative outcomes of persuasion knowledge activation for covert advertising, J. Curr. Issues Res. Advert. 36 (2) (2015) 157–176.