

# *A Study of Public Awareness and Perceptions for Enhancing Deepfake Detection Technologies*

Jericka Guy  
Faculty - Dr. Jean Muhammad  
Computer Science Department  
Hampton University  
Hampton, VA

**Abstract - Deepfake technology presents a significant challenge to cybersecurity. These highly sophisticated AI-generated manipulations can compromise sensitive information and erode public trust, privacy, and security. This has led to broader societal impacts, including decreased trust and confidence in digital communications. This paper will discuss public knowledge, understanding, and perception of AI-generated deepfakes, which was obtained through an online survey that measured people's ability to identify video, audio, and images of deepfakes. The findings will highlight the public's knowledge and perception of deepfakes, the risks that deepfake media presents, and the vulnerabilities to detection and prevention. This awareness will lead to stronger defense strategies and enhanced cybersecurity measures that will ultimately enhance deepfake detection technology and strengthen overall cybersecurity measures that will effectively mitigate exploitation risks and safeguard personal and organizational interests.**

**Keywords:** *Cybersecurity, AI-Generated Media, Deepfake Technology, Social Engineering, Risk Mitigation*

## **I. Introduction**

A deepfake is a type of media that uses artificial intelligence (AI) to create realistic-looking content. Deepfake technology manipulates existing media to produce content that convincingly mimics real people and events. As deepfake technology advances, it is becoming a prominent issue in cybersecurity. The artificial content produced by manipulating existing videos, images, and audio now poses severe threats to privacy and security. The rise of social engineering and advanced AI-based software has enabled even those with little technical skills to be able to generate artificial content at an exponential rate. This growth has introduced challenges not just in detecting these fakes media but also in addressing their broader societal impacts. One example of deepfakes being misused was the recent release of images showing Taylor Swift supporting a presidential candidate. The photos were discovered to be fake, but the incident illustrated how fake media can deceive large audiences and demonstrate the vulnerabilities they exploit to manipulate public opinion and potentially cause harm.

Deepfakes present several significant issues across cybersecurity. Deepfaked media has the potential to severely influence public opinion and spread deceptive information, leading to a lack of public trust in digital

content. In addition, deepfake technology is often misused to create non-consensual explicit content, contributing to harassment and damaging the reputations of individuals.

Deepfakes can also be used as a tool for audio and visual impersonation. This allows attackers to easily deceive individuals and exploit vulnerabilities within systems thus increasing the risks of identity thefts, financial fraud, and social engineering attacks. To keep up with the sophistication of deepfaked media, improvements must be made in technology to detect them. It is also important to educate the public on recognizing falsified media. By addressing these challenges, cybersecurity professionals can better defend against the evolving threats posed by deepfake technology.

This paper will analyze the role of deepfakes similar to what is used in cyberattacks, focusing on the psychological manipulation and technical vulnerabilities they exploit. The research will examine the public's ability to identify deepfakes, using an online survey to gather data on the recognition of manipulated video, audio, and images. By understanding how easily individuals can be deceived, the study will highlight the risks that deepfaked media present. The findings will shed light on the current state of public awareness regarding deepfakes. By comparing responses across various demographics, the research aimed to identify gaps in knowledge of technology, AI and deepfakes and individuals' susceptibility to deepfake-related media postings. Ultimately, this research seeks to inform strategies for enhancing deepfake detection technology, thereby strengthening overall cybersecurity measures that will effectively mitigate exploitation risks and safeguard personal and organizational interests.

## II. Methodology

This research will use a combination of a literature review and an online survey to collect data and gather results directly related to deepfake detection and awareness. Each stage of the methodology is explained as follows.

### A. Literature Review

Completing a literature review of articles that explored the role of deepfakes in cyberattacks and the psychological manipulation and technical vulnerabilities that they exploit, it can be concluded that deepfakes emerge as a significant complex threat in the digital age. The possible consequences of their use could lead to widespread distrust in digital content and demonstrate a need for strengthening overall cybersecurity measures with strategies to enhance deepfake detection technology.

Several studies provided an overview of deepfake technology, including its creation, process, and potential applications (Chadha et al., 2021; Mirsky & Lee, 2021; Seow et al., 2022; Tolosana et al., 2020). These studies highlight AI technology as the primary technique for creating deepfakes and for allowing the manipulation of audio and video to develop realistic forgeries. Kwok & Koh (2020) also offer a social construction of the technology and the perspective that emphasizes the societal implications of deepfakes beyond malicious use.

Deepfakes present a unique and evolving threat to cybersecurity. Cybercriminals are able to create realistic videos or audio messages that trick users into downloading malicious content or opening compromised links. Whittaker (2019) reported instances when cybercriminals have used deepfakes to impersonate CEOs and bypass security measures, thereby defrauding companies and manipulating employees in order to obtain

large sums of money. These attackers often use deepfakes that come from a trusted source with authentication processes that rely on facial recognition, voice recognition, or behavioral patterns for security. Deepfakes can simulate these biometrics, creating security loopholes in access control systems (Firc et al 2023).

Deepfakes enhance the effectiveness of spear phishing attacks by personalizing messages with convincing, AI-generated videos or voice recordings of trusted individuals (Westerlund, 2019). Vaccari & Chadwick's (2020) research suggests that video and audio-based deepfakes can be particularly persuasive, as they mimic the authenticity of face-to-face communication. This makes them a potent tool for social engineering attacks. These social engineering attacks leverage human psychology to manipulate individuals into revealing sensitive information or performing actions that compromise security. Schmitt & Flechais (2023) discuss the rise of deepfakes in social engineering and phishing scams. Firc et al. (2023) highlight their use in compromising facial and speaker recognition systems. Doss et al. (2023) points out the potential for deepfakes to disrupt scientific knowledge dissemination by manipulating research findings.

While deepfakes present significant threats, the development of detection mechanisms offers some hope in combating these attacks. Gupta et al. (2023) reviews the various advanced machine-learning techniques used to identify deepfake media. These detection methods focus on identifying inconsistencies in visual or auditory features that are difficult to replicate with current AI models. However, recent developments in machine learning make deepfakes highly believable, and very difficult to differentiate between what is real and what is fake. Not only humans but also machines struggle to

identify deepfakes. This is supported by Westerlund (2019) who noted that AI-generated voice deepfakes can defeat voice authentication systems, highlighting a growing vulnerability in cybersecurity frameworks.

Deepfake attacks exploit a combination of psychological and technological vulnerabilities. Jacobson (2024) emphasizes the impact on social trust, while Gordon (2024) focuses on the difficulty of detecting audio deepfakes. Naitali et al. (2023) provide a comprehensive overview of deepfake attacks, highlighting generation techniques, detection methods, challenges, and research directions.

As deepfakes blur the line between real and fake, legal frameworks must evolve to ensure accountability and protect against misuse. Van der Sloot and Wagenveld (2022) discuss the regulatory challenges that deepfakes present, noting that existing laws often struggle to address the use of synthetic media in harmful ways. The U.S. Government Accountability Office (2020, 2024) has also weighed in on the risks associated with deepfakes, particularly regarding their potential use in fraud, espionage, and other criminal activities. These reports suggest that a combination of legal, technical, and educational approaches is necessary to combat the growing threat of deepfakes.

The literature reviewed emphasizes a need for heightened awareness and a stronger understanding of deepfakes, particularly regarding their cybersecurity risks and potential for psychological manipulation and technical vulnerabilities. As deepfakes continue to evolve and pose threats to trust in digital media, assessing the public's ability to identify these deceptive technologies becomes essential. The findings from these studies display the need for this research that

will focus on gauging participants' recognition of deepfakes, their perceived risks, and their experience with such media. By exploring these findings, this project aims to contribute valuable insights that can inform future strategies for deepfake detection, public education, and policy development.

### B. Survey

A survey-based research design was implemented to investigate the impact of deepfakes on media recognition and awareness. The survey was structured into two parts: media recognition and awareness questions. For media recognition, participants were presented with pairs of videos, audio clips, and images and asked to identify which of the media files were deepfake. This section evaluated the participants' ability to recognize manipulated content. The second part of the survey aimed to measure participants' understanding of deepfakes and consisted of questions that assessed the participants' awareness of deepfakes, their confidence in identifying deepfakes, and any concerns they had regarding the potential risks posed by deepfakes.

### C. Population

To ensure diversity in the results, the target audience for this survey included individuals of all ages, genders, and levels of technological familiarity. For ease of access, the survey is administered using Google Forms. The testing format included various question types, such as multiple-choice, fill-in the blank and scoring using the Likert scale.

Data collection occurred over three weeks. During this time, participants completed the survey online, and the responses were compiled using Google Forms. Quantitative analysis focused on determining the

percentage of participants who accurately identified the deepfakes video, audio clips, and images. Qualitative analysis focused on examining patterns and themes in participants' responses to the awareness and confidence questions. Overall, this research approach yielded comprehensive insights into the media recognition and participants' understanding of deepfakes.

## III. Results

This section will report the results from the research methodology outlined in Section II.

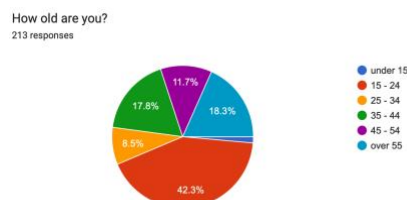
### A. Demographic Data

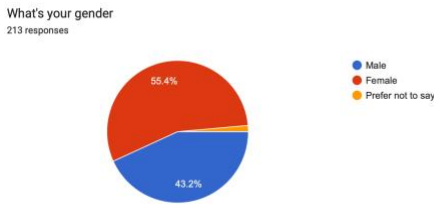
The demographic data was collected from 213 participants, resulting in a nearly even split of genders. As seen in Figure 2, the majority identified as female, 118 participants (55.4%), and 92 participants; (43.2%) identified as male. A few participants chose not to disclose their gender (3 participants; (1.4%).

As seen in Figure 1, most participants were in the 15–24 age group (90 participants; 42.3%) with a significant number over 55 years old (39 participants; 18.3%). The smallest subgroup was represented by three participants under the age of 15, 1.4%. 18 participants were aged 25-34 (8.5%); 38 participants aged 35-44 (17.8%); and 25 participants were aged 45-54, (11.7%).

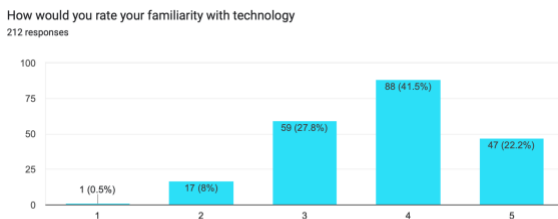
**Figure 1**

#### *Results of Participants Age*



**Figure 2***Results of Participants' Gender*

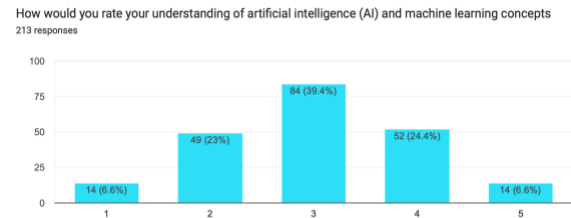
To achieve additional data on the participants' background, they were asked to use a Likert scale of 1 to 5, to rate their familiarity with technology (Figure 3), artificial intelligence (AI) (Figure 4), and deepfake technology (Figure 5).

**Figure 3***Familiarity with technology*

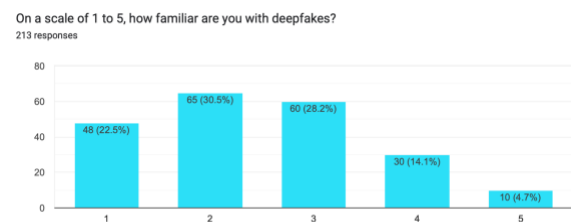
The data revealed varying levels of knowledge. 47 participants (22.2%) rated their technology familiarity as 5; 88 participants (41.5%) rated their technology familiarity as 4; 59 participants (27.8%) rated their familiarity as 3; 17 participants (8%) rated their technology familiarity as 2, and 1 participant (0.5%) rated their technology familiarity as 1.

In rating their understanding of AI and machine learning concepts, 14 participants (6.6%) rated their understanding as either a 1 or 5; 52 participants (24.4%) rated their understanding as 4; 84 participants (39.4%) rated their understanding as 3; and 49

participants (23%) rated their understanding as 2.

**Figure 4***Familiarity with AI and Machine Learning Concepts*

Ratings of their familiarity with deepfakes has 10 participants (4.7%) whose familiarity ratings are 5; 30 participants (14.1%) whose familiarity ratings are 4; 60 participants (28.2%) with familiarity ratings are 3; 65 participants (30.5%) familiarity ratings are 2, and 48 participants (22.5%) familiarity ratings are 1.

**Figure 5***Familiarity with Deepfake Technology**B. Quantitative Results*

The deepfake identification tests included two photo tests, two video tests, and an audio test. The participants were given two tests with two photos to compare, two tests with three videos to compare, and one audio test with two audio clips to compare. Participants were asked to assess the information provided in each test and identify whether the

information was real or manipulated (deepfake).

**Photo Sample Test 1:** The first sample involved a photo of celebrity actor Tom Cruise (photo B) and a deepfake photo made to look like him (photo A). Participants had to choose the deepfake using one of 4 options. Was Photo A deepfake; was Photo B deepfake; were both Photo A & B deepfakes; or whether neither one was deepfake?

As seen in Figure 7, ninety-three participants (43.7%) chose Photo A as a deepfake, 74 participants (34.7%) chose Photo B as a deepfake, 22 participants (10.3%) chose both Photo A and B as deepfakes, and 24 participants (11.3%) chose neither.

**Figure 6**

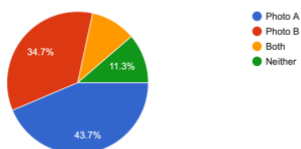
*Photo Test #1 - Photo A (left) Photo B (right)*



**Figure 7**

*Results from Photo Test #1*

Please view the two images above. Which one do you think is a deepfake?  
213 responses



**Photo Sample Test 2:** The second photo sample question was similar to the first and featured two deepfakes of celebrity actor Tom Cruise. Again, participants had to

choose the deepfake using one of 4 options. Was Photo A deepfake; was Photo B deepfake; were both Photo A & B deepfakes; or whether neither one was deepfake?

As seen in figure 9, seventy-three participants (34.3%) chose Photo A as deepfake, 52 participants (24.4%) chose Photo B as deepfake; 33 participants (15.5%) chose both Photo A & B as deepfake and 55 participants (25.8%) chose neither as deepfakes.

**Figure 8**

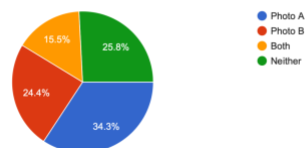
*Photo Test #2 - Photo A (left) Photo B (right)*



**Figure 9**

*Results from Photo Test #2*

Please view the two images above. Which one do you think is a deepfake?  
213 responses



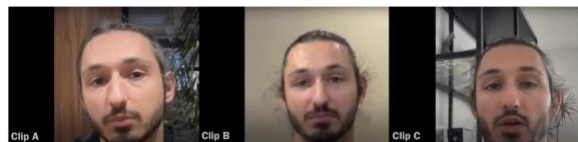
**Video Samples:** The video samples used video clip tests. There were three clips to choose from, and the participants had to identify which were the deepfakes. There were 8 choices. Whether Clip A, Clip B, Clip C, Clips A&B, Clips B&C, Clips A&C, all 3 Clips and none of the clips.

See Figure 11: **Video Sample Test 1.** The results: 20 participants chose Clip A (9.4%); 36 participants chose Clip B (17%); 25 participants chose Clip C (11.8%); 31 chose

Clips A&B (14.6%), 19 chose Clips B&C (9%); 23 participants chose Clips A&C (10.8%); 36 choose all 3 Clips (17%), and 22 participants chose none of the clips (10.4%).

### Figure 10

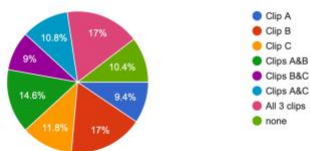
*Video Test #1 - Clip A (left) Clip B (middle) Clip C (right)*



### Figure 11

*Results from Video Test #1*

Please watch the video above. Which clip do you think is a deepfake?  
212 responses



See Figure 13: **Video Sample Test 2**. The 2nd video sample was similar to the first and used a test with three video clips. Again, there were three clips to choose from, and the participants had to identify which were the deepfakes. There were 8 choices. Whether Clip A, Clip B, Clip C, Clips A&B, Clips B&C, Clips A&C, all 3 Clips and none of the clips. The results: 34 participants chose Clip A (16%); 22 participants chose Clip B (10.4%); 57 participants chose Clip C (26.9%); 22 chose Clips A&B (10.4%), 18 chose Clips B&C (8.5%); 16 participants chose Clips A&C (7.5%); 23 choose all 3 Clips (10.8%), and 20 participants chose none of the clips (9.4%). (See figure 13).

### Figure 12

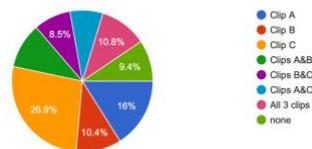
*Video Test #2 - Clip A (left) Clip B (middle) Clip C (right)*



### Figure 13

*Results from Video Test #2*

Please watch the video above. Which clip do you think is a deepfake?  
212 responses



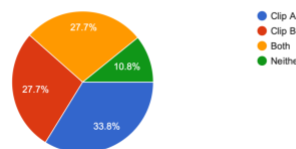
See Figure 14: **Audio Sample Test**. The audio sample tests involved two audio clips: Clip B depicts the voice of former President Obama, and Clip A depicts a deepfake. Participants had to choose the deepfake using one of 4 options. Was Clip A deepfake, was Clip B deepfake, were both Clips A & B deepfakes, or whether neither was deepfake?

Seventy-two participants (33.8%) chose Clip A as deepfake; 59 participants (27.7%) chose Clip B as deepfake; 59 participants (27.7%) chose both Clips A & B as deepfake and 23 participants (10.8%) chose neither as deepfakes.

### Figure 14

*Results from Audio Test*

Please listen to the audio above. Which clip do you think is a deepfake?  
213 responses



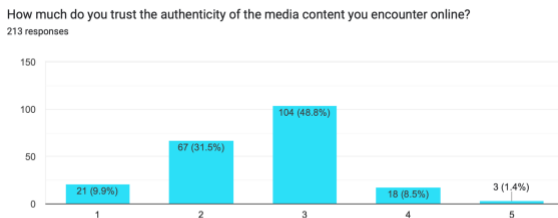
### C. Qualitative Results

Participants were asked what they believed were the greatest risks from deepfakes and their confidence level in identifying them. Using a Likert scale from 1 to 5, they reported on whether they trust the authenticity of the media content they encounter online, how concerned they are about the potential misuse of deepfakes in the media, and their confidence level in identifying deepfakes.

In Figure 15: Trust in Online Media Content, three participants (1.4%) rated the media content as 5; 18 participants (8.5%) rated the content as 4; 104 participants (48.8%) rated the content as 3; 67 participants (31.5%) rated the content as 2; and 21 participants (9.9%) rated the content as 1.

**Figure 15**

#### *Trust in Online Media Content*



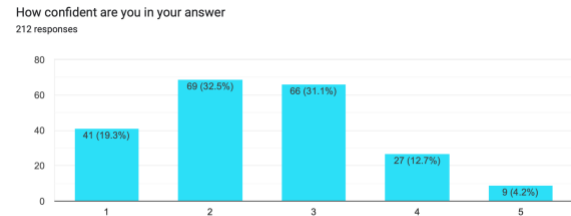
In reporting their confidence level to identify deepfakes, nine participants (4.2%) rated their confidence as 5; 27 participants (12.7%) rated their confidence as 4; 66 participants (31.1%) rated their confidence as 3; 69 participants (32.5%) rated their confidence as 2; and 41 participants (19.3%) rated their confidence as 1. (See Figure 16).

As seen in Figure 17, ninety-four participants (44.3%) rated their concern about media content as 5; 70 participants (33%) rated their concern as 4; 33 participants (15.6%) rated their concern as 3; 10 participants (4.7%)

rated their concern as 2; and 5 participants (2.4%) rated their concern as 1.

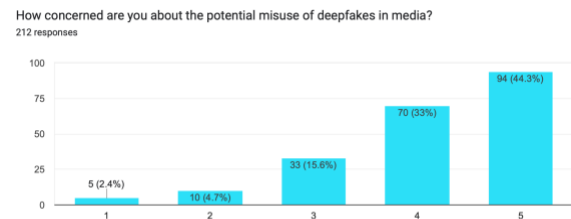
**Figure 16**

#### *Confidence in Video Test #2 Answer*



**Figure 17**

#### *Concern Over Deepfake Misuse in Media*



## IV. Data Analysis

This section will give an analysis of the survey data presented in Section III.

### A. Demographic Data

Most participants in the research were female between the ages of 15 and 24. The average familiarity score for general technology was 3.768, indicating that most respondents were moderately to highly familiar with technological concepts. In contrast, familiarity with AI had a lower average of 3.014, suggesting that participants were less knowledgeable about this specific area. Deepfake technology scored even lower, with an average familiarity of 2.476, reflecting widespread unfamiliarity among respondents.

## B. Quantitative Analysis

**Photo Samples:** The first question involved a photo of celebrity actor Tom Cruise and a deepfake photo to look like him. While I initially added this photo, thinking it would be the easiest to identify correctly, the results were very split. A significant number of participants were wrong about the photo. The only deepfake was Photo A, and only 43.7% (93 participants) answered it correctly.

The second photo sample test was similar to the first but featured two deepfakes of celebrity actor Tom Cruise. I wanted to see if the first photo sample test was chosen by guess or if the survey takers would correctly recognize Tom Cruise. Given all the choices, the responses were closely spread between the four options. However, only 15.5% (33 participants) were able to correctly identify that both Photos were deepfakes.

**Video Samples:** The first video clip test had two deepfakes and gave more choices; however, since the video was of the same person, it required more in-depth comparative observation and focus on details before the selection could be made. The results showed a more widespread identification of the video clip. The two deepfakes were Clip A and C. Only 9.4% (20 participants) found that Clip A was a deepfake, and 11.6% (25 participants) found that Clip C was a deepfake. An even lower number, only 10.8% (23 participants) correctly recognized that they were both deepfakes.

The second video clip test was similar to the first. However, in this test, the clips were of three different people, and the participant had to identify which were the deepfakes. There were two correct answers: both Clip A and B were deepfaked. This test required less comparative observation but more knowledge of deepfakes, visual effects,

speech patterns, and the technology behind how deepfakes are generated.

The results of this test were the most shocking as the most incorrect answer was the most chosen; 26.9% (57 participants) believed that clip C was the deepfake despite it being the only non-deepfaked video. I believe one possible reason for this is that participants were not knowledgeable of the sophisticated technology that is used in creating deepfakes. When analyzing the videos, they were unaware of how to fully account for visual effects, including variations in speech, accents, or other speech patterns. For instance, Clip C featured a young adult with an East Asian background, and her English was spoken with an accent. This could have influenced participants' perceptions, as her speech may have been less familiar or unclear compared to the spoken English in Clips A and B. Potentially, this may have led to her clip being misidentified as a deepfake.

**Audio Test Sample:** The audio test clip depicting the voice of former President Obama proved to be the easiest challenge for participants, with 33.8% (72 participants) accurately identifying clip A as a deepfake. I believed this was the easiest to identify as a deepfake due to President Obama's being a public figure and the actual content of the audio clips. Based on what the former president was depicted to be saying in Clip A, it was less likely for him to state that publicly than what was said in Clip B.

In addition, the participant could have been listening to his distinct voice qualities, including unique pitch, resonance, or vocal texture. As a popular president, most people had heard his voice multiple times and were familiar with the content of what he was saying. Thus, the deepfake voice of a popular person speaking on a more familiar topic might be more easily recognized than a less

familiar voice speaking on a less familiar topic.

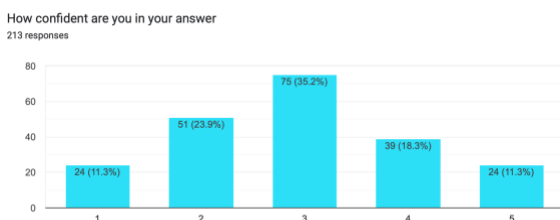
A summary of the performance data of all five tests revealed a significant challenge in the participants correctly identifying deepfakes. This was irrespective of whether it was visual or audio. Only 81 participants (38.2%) correctly identified at least one deepfake, while a mere 1.9% (4 individuals) correctly identified all five deepfakes. Conversely, 35.8% (76 participants) failed to identify any deepfake correctly.

### C. Qualitative Insights

The qualitative responses provided more profound insights into participants' trust levels, concerns about deepfakes, and their broader societal implications. On average, participants expressed moderate skepticism about the reliability of online media, with a mean trust score of 2.599 on a scale of 1 to 5. However, concern about the dangers posed by deepfakes was significantly higher, with an average score of 4.118. Many respondents viewed deepfakes as a critical threat due to their potential for misuse.

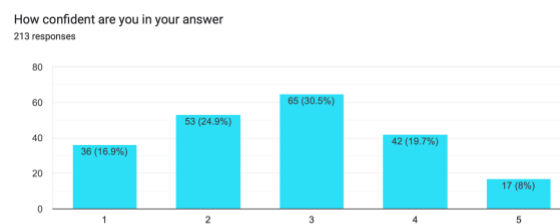
**Figure 18**

*Confidence in Photo Test #1 Answer*



**Figure 19**

*Confidence in Audio Test Answer*



Analysis of participants' confidence scores when identifying deepfake images and audio are reported in Figures 18 and 19. The results showed that most participants had low scores when reporting their confidence in identifying deepfakes. For the small percentage who rated their confidence as high (scores of 4 or 5), their performance in correctly identifying deepfakes, whether using images or audio, still scored fairly low. Hence, there was a discrepancy between participants' confidence and their ability to identify deepfakes.

Further analysis revealed that familiarity with deepfakes, while somewhat helpful, did not strongly correlate with performance. The correlation between familiarity with deepfakes and accurate identification was moderate (0.198), suggesting that while prior exposure to deepfake technology can slightly improve recognition, it is not a strong predictor of success. Conversely, familiarity with general technology (0.076) and AI concepts (0.146) had even weaker correlations with accuracy, further emphasizing that general technological knowledge does not necessarily translate into better recognition of deepfakes.

## V. Conclusion

The findings of this research revealed a concerning correlation between the participants' knowledge, abilities, and

perceptions when identifying deepfakes. The participants who had little technological background and were less familiar with deepfakes tended to express low confidence in their judgments of deepfakes. This showed their increased vulnerability to deepfake deceptions.

Additionally, the study has shown a significant knowledge gap between participants' understanding of general technology concepts and AI-specific developments. This highlights the need to educate the public in specific deepfake education to bridge the knowledge gap in identifying and understanding deepfakes. With this knowledge, individuals are better able to assess the authenticity of digital content. Whether through awareness campaigns or digital literacy programs, empowering individuals to navigate the challenges of sophisticated deepfake technologies will build public trust and confidence in technology and cybersecurity. The better prepared citizens are to recognize deepfakes and be confident about their abilities, the less severe the impacts will be from deepfake-based misinformation.

Along with raising awareness, cybersecurity experts must improve technology to foster and build the public trust and confidence in a more secure and trustworthy digital environment. Combining this with educational initiatives will result in improved detection strategies, better public engagement with decreased vulnerability, and lower implications and exploitation risks from deepfake media. Ultimately this will lead to a more informed, resilient society

with advanced deepfake technology and stronger defense strategies capable of navigating the challenges posed by deepfakes. This approach will not only protect individuals and organizations by safeguarding their security and privacy but also contribute to their trust and confidence in the digital environment, thereby having a great impact on modern society.

## References

1. Chadha, A., Kumar, V., Kashyap, S., Gupta, M. (2021). Deepfake: An Overview. In: Singh, P.K., Wierzchoń, S.T., Tanwar, S., Ganzha, M., Rodrigues, J.J.P.C. (eds) Proceedings of Second International Conference on Computing, Communications, and Cyber-Security. Lecture Notes in Networks and Systems, vol 203. Springer, Singapore. [https://doi.org/10.1007/978-981-16-0733-2\\_39](https://doi.org/10.1007/978-981-16-0733-2_39)
2. Doss, C., Mondschein, J., Shu, D., Wolfson, T., Kopecky, D., Fitton-Kane, V. A., Bush, L., & Tucker, C. (2023). Deepfakes and scientific knowledge dissemination. *Scientific Reports*, 13(1). <https://doi.org/10.1038/s41598-023-39944-3>
3. Firc, A., Malinka, K., & Hanáček, P. (2023). Deepfakes as a threat to a speaker and facial recognition: An overview of tools and attack vectors. *Heliyon*, 9(4), e15090. <https://doi.org/10.1016/j.heliyon.2023.e15090>
4. Gordon, R. (2024, March 15). 3 questions: What you need to know about audio deepfakes. MIT News.

- <https://news.mit.edu/2024/what-you-need-to-know-audio-deepfakes-0315>
5. Gupta, G., Raja, K., Gupta, M., Jan, T., Whiteside, S. T., & Prasad, M. (2023). A comprehensive review of Deepfake detection using advanced machine learning and Fusion Methods. *Electronics*, 13(1), 95. <https://doi.org/10.3390/electronics13010095>
  6. Jacobson, N. (2024, February 26). *Deepfakes and Their Impact on Society*. CPI OpenFox. <https://www.openfox.com/deepfakes-and-their-impact-on-society/>
  7. Kwok, A. O. J., & Koh, S. G. M. (2020). Deepfake: a social construction of technology perspective. *Current Issues in Tourism*, 24(13), 1798–1802. <https://doi.org/10.1080/13683500.2020.1738357>
  8. Mirsky, Y., & Lee, W. (2021). The Creation and Detection of Deepfakes: A Survey. *ACM Computing Surveys*, 54(1), 1–41. <https://doi.org/10.1145/3425780>
  9. Naitali, A., Ridouani, M., Salahdine, F., & Kaabouch, N. (2023). Deepfake attacks: Generation, detection, datasets, challenges, and Research Directions. *Computers*, 12(10), 216. <https://doi.org/10.3390/computers12100216>
  10. Schmitt, M., & Flechais, I. (2023). Digital deception: Generative artificial intelligence in Social Engineering and phishing. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4602790>
  11. Seow, J. W., Lim, M. K., Phan, R. C. W., & Liu, J. K. (2022). A comprehensive overview of Deepfake: Generation, detection, datasets, and opportunities. *Neurocomputing*, 513, 351–371. <https://doi.org/10.1016/j.neucom.2022.09.135>
  12. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deepfakes and beyond: A survey of face manipulation and fake detection. *Information Fusion*, 64, 131–148. <https://doi.org/10.1016/j.inffus.2020.06.014>
  13. U.S. Government Accountability Office. (2020, October 20). *Deconstructing deepfakes-how do they work and what are the risks?*. gao.gov. <https://www.gao.gov/blog/deconstructing-deepfakes-how-do-they-work-and-what-are-risks>
  14. U.S. Government Accountability Office. (2024, March 11). *Science & Tech spotlight: Combating Deepfakes*. gao.gov. <https://www.gao.gov/products/gao-24-107292>
  15. Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic media on political trust. *Political Communication*, 37(4), 603-623.
  16. Van der Sloot, B., & Wagenveld, Y. (2022). Deepfakes: Regulatory challenges for the synthetic society. *Computer Law & Security Review*, 46, 105716. <https://doi.org/10.1016/j.clsr.2022.105716>

17. Westerlund, M. 2019. The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review*, 9(11): 40-53. <http://doi.org/10.22215/timreview/1282>
18. Whittaker, Z. (2019). Deepfake scams: A growing threat to businesses. *Cybersecurity Today*, 5(3), 34-39.