Impact of AI-Generated Media on Election Security

Brandon Amacher
Director, Emerging Tech Policy Lab

Emerging Tech Policy Lab
Kaye Banner, Senior Analyst
Hope Fager, Strategic Research Team Lead
Leah Olsen, Senior Analyst
Maliq Rowe, Senior Analyst

Herbert Instititute for Public Policy
William Freedman, Research Assistant

BYU | HERBERT INSTITUTE FOR PUBLIC POLICY | CENTER FOR NATIONAL SECURITY STUDIES | EMTECH

Research Summary | June 2026

Impact of AI-Generated Media on Election Security Phase II

Executive Summary

AI-generated deepfakes represent a rapidly escalating threat to democratic processes and informed political participation. This study, the second in a series by a cross-disciplinary coalition of UVU faculty, staff, and students, extends prior research by examining whether synthetic video content can influence real political opinions in a realistic social media context, and whether ordinary citizens can reliably detect these deepfakes.

Key takeaways from this study include:

Deepfake videos influenced the opinions of potential voters just as effectively as real media.

No statistically significant difference in opinion change was observed between participants who watched real videos versus synthetic ones. Synthetic speakers received equal or higher credibility ratings across all three dimensions measured.

Participants across all tested demographics (political affiliation, age, etc.) could not reliably detect if they had been exposed to deepfake content.

Only 16.4% of participants correctly identified a deepfake after viewing it. A majority (58.7%) were convinced synthetic videos were real.

Familiarity with deepfakes does not protect you: Self-reported familiarity with deepfakes did not improve detection accuracy.

Subjects who claimed to be "Not at all familiar" with deepfakes were equally capable of correctly identifying whether they had seen deepfake content as those who claimed to be "Extremely familiar" with deepfakes (with 42% and 41% correct identification respectively).

Project Background & Structure

In fall 2024, the Center for National Security Studies' Emerging Tech Policy Lab (EMTECH) and the Gary R. Herbert Institute for Public Policy, both from Utah Valley University, conducted a study examining the impact of AI-generated deepfakes—defined as videos created or altered using artificial intelligence—compared to authentic video content. The study found that most participants could not reliably distinguish between real and AI-manipulated media. It also found that participants viewed deepfake content as equally impactful, trustworthy, and credible as authentic media. The findings of this initial study have significant implications for national security, economic development, and public policy.

One area of particular concern is public discourse and election security. In recent years, dozens of cases of AI-enabled election disinformation have emerged around the world, affecting contests ranging from federal elections to local and municipal ballot initiatives. However, the extent of their impact remains widely debated. Individuals targeted by deepfakes often argue that the content substantially influenced election outcomes, while electoral beneficiaries of deepfake disinformation frequently minimize its significance.

To further examine these concerns, EMTECH, in coordination with the UVU Herbert Institute for Public Policy, launched this follow-up study to analyze the effects deepfakes may have on elections. Specifically, this study examines the extent to which deepfakes can influence voter opinions compared to authentic media, whether certain communities are more vulnerable to deception by deepfakes, and which personal or societal factors may affect an individual's ability to be affected by manipulated content.

Study Design

In this study, we sought to replicate the way individuals are likely to encounter deepfakes in real-world settings, particularly while casually scrolling through social media without heightened suspicion or awareness. To mirror these conditions, participants were not informed at the outset that they might view AI-altered content. Instead, they were told that the study examined how individuals form and adjust opinions after viewing short-form media content. Only after subjects were shown a media sample, and measurements about the impact of the media on the viewer were taken, was the true nature of the study revealed.

All deepfake materials used in the study were made by students at the UVU Center for National Security Studies using publicly available tools. In order to mitigate the introduction of potential biases, we attempted to choose an issue that would not elicit strong emotional reactions. The text that appears in the survey (a fictional federal housing assistance initiative) was written to resemble an issue which might appear on a real ballot.

Research Questions

The primary objective of this study was to explore the effects a deepfake video could have on a viewer's opinion, and to discover how well a deepfake could influence someone to alter their opinion when compared to authentic media. The secondary objective of this study was to explore the effects previous opinions and other personal and societal factors could have on how a person perceives and detects deepfake content.

In total, we designed this study to address the following questions:

1. Is there a measurable difference in the way a deepfake video can influence someone's opinion versus a real video?
2. How well can deepfake videos be detected if the arguments agree or disagree with someone's existing opinion?
3. If a video is considered to be credible, is it less likely to be detected as a deepfake?
4. Is there any difference in an individual's ability to detect a deepfake if they trust or interact with news media consistently?
5. Does awareness/familiarity of deepfakes influence an individual's ability to detect them?

Methodology

This study used an online survey platform to conduct testing and collect data from a politically representative sample of the United States population. A total of 632 respondents participated in the study.

At the beginning of the survey, participants were asked 1) how they would vote on a down-ballot initiative, 2) to explain the reasoning behind their position, and 3) to indicate how firmly they held that view. Based on their responses, participants were categorized into one of three groups—those likely or very likely to vote "yes," those who were neutral or undecided, and those likely or very likely to vote "no" on the proposed initiative.

Within each group, participants were randomly assigned to view one of four videos:

- A real pro-initiative video
- An AI-generated pro-initiative video
- A real anti-initiative video
- An AI-generated anti-initiative video

After exposure to the video, participants were again shown the ballot text of the issue and asked the same questions about their opinion, reasoning, and flexibility of their vote.

Participants then evaluated the video using a seven-point Likert scale and accompanying free-response questions. The evaluation measures included

- The perceived knowledgeability of the speaker
- The perceived trustworthiness of the speaker
- The persuasiveness of the speaker
- The overall quality of the video

Figure 1

After completing these questions, respondents were informed that the video they viewed may have been either authentic or AI-generated. They were then asked to determine whether the video was real or synthetic, selecting from the following options: "definitely deepfake," "probably deepfake," "unsure," "probably real," or "definitely real." Participants also explained the reasoning behind their assessment and rated their confidence in that conclusion. Finally, respondents were asked about their familiarity with deepfakes, whether they believed that familiarity helped them identify manipulated content, and how concerned they were about the potential influence of deepfakes on the public.

Results

This study produced a wide range of findings on this topic. However, there are three key findings from this study that are of particular notability to policymakers and the public:

1 **Deepfake videos influenced the opinions of potential voters just as effectively as real media.**

This study found no statistical significance in the ability of a video to alter someone's previously held opinion if it was a deepfake or a legitimate video. The variance between the ability of real videos and their respective identical deepfakes to influence voters was statistically negligible. This finding is an alarming illustration of the potential efficacy of deepfake disinformation operations.

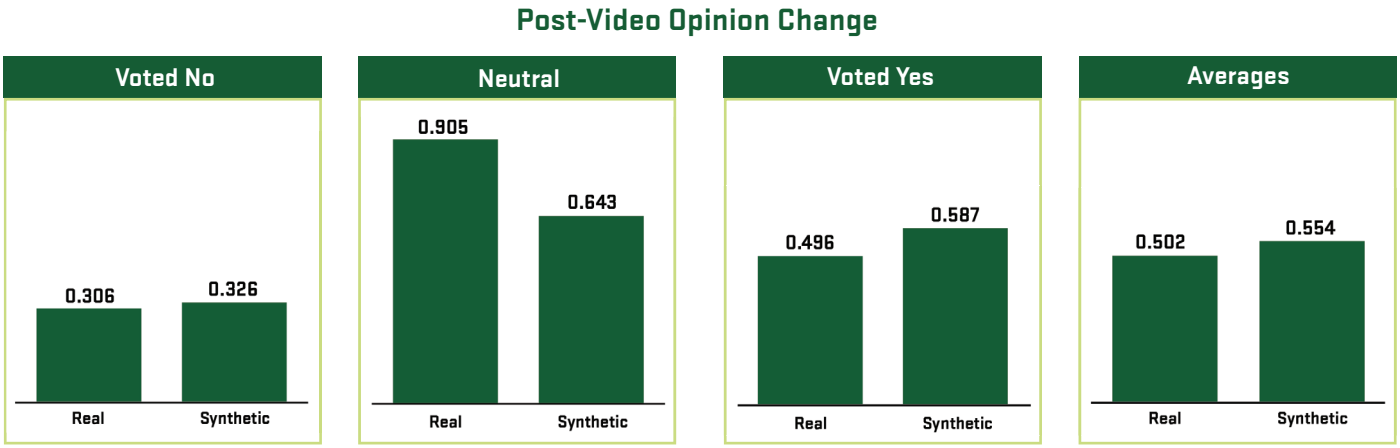

Figure 2

Participants were first asked how they would vote. Then, they watched a video and answered the same question again. To calculate the change in opinion, we subtracted the first response from the second response and used the absolute value of that difference.

2 **Participants across all tested demographics (political affiliation, age, etc.) could not reliably detect if they had been exposed to deepfake content.**

Even after being explicitly informed that they may have seen a deepfake, participants were largely unable to correctly identify AI-generated content. The detection results reveal a stark asymmetry.

Video Type	Correctly Identified	Misidentified	Unsure
Real Video	54.3%	22.2%	23.5%
Synthetic Video	16.4%	58.7%	24.9%

Figure 3—Correct Identification Rates by Media Type

This finding held consistently across demographic groups. Detection rates for synthetic videos were roughly equal across political affiliations—Democrats, Republicans, and Independents all identified deepfakes correctly at rates 15–19%. No age group, political affiliation, or other tested demographic showed a meaningful advantage. Deepfake detection failure is a universal vulnerability, not one concentrated in any particular group.

Among participants who said they were "very confident" or "extremely confident" in their guess while watching a synthetic video, incorrect identifications outnumbered correct ones by more than three to one (58% incorrect vs. 16% correct). Overconfidence compounds the risk; those most certain they could spot a deepfake were most likely to be wrong.

The default assumption is "real." Across all 632 participants and both video types, 56.5% of all guesses were that the video was real. Only 19.3% guessed deepfake.

Furthermore, across all of the perceived attributes we tested (trustworthiness, knowledgeability, persuasiveness, credibility, and quality), deepfake content received evaluations comparable to authentic content across all categories, with differences falling below the threshold for statistical significance.

3 **Familiarity with deepfakes does not protect you: Self-reported familiarity with deepfakes did not improve detection accuracy.**

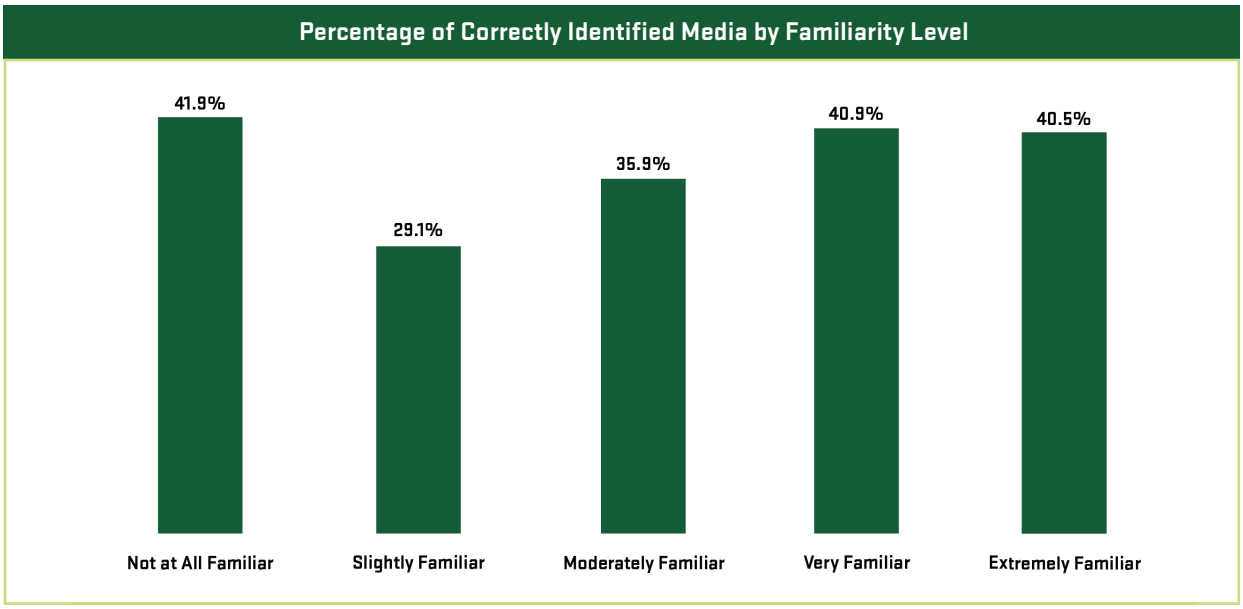

Figure 4—Media Identification by Familiarity Level

Participants self-reported their familiarity with deepfakes on a five-point Likert scale ranging from "Not at all familiar" to "Extremely familiar." Despite the wide range of familiarity levels represented, correct identification rates on the extremes were effectively the same.

Participants who described themselves as "Not at all familiar" with deepfakes correctly identified whether they had seen real or synthetic content 41.9% of the time. Those who described themselves as "Extremely familiar" performed at virtually the same rate: 40.5%. This difference is statistically negligible.

Furthermore, participants who reported being "extremely confident" that they had correctly identified deepfake content were, by a substantial margin, the most likely to have incorrectly classified the media. they viewed.

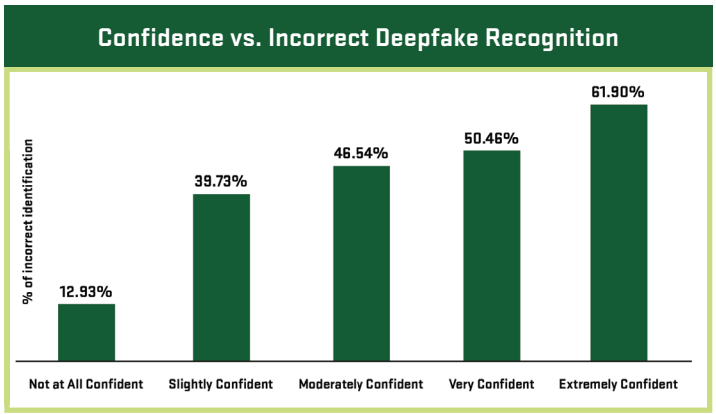

Figure 5—Confidence in Identification of Deepfake vs. Accuracy

" I am familiar with common signs of deepfakes, such as unnatural facial movements, inconsistencies in audio and distorted visual formatting. Knowing what to look for helped me more confidently assess the video and determine that it appeared authentic."

— Test participant who, in fact, viewed a deepfake video

This finding has significant implications for media literacy policy. Programs that aim to raise awareness of deepfakes as a concept appear insufficient on their own. Knowing that deepfakes exist does not give citizens the perceptual tools to identify them in practice.

Conclusion

The findings of this study present a clear and urgent challenge for policymakers, election security professionals, and the public: AI-generated deepfakes are already capable of influencing political opinion as effectively as authentic media, and the people most likely to encounter them are not equipped to identify them.

The implications of this are significant. Democratic systems depend on an informed electorate, and the integrity of that process rests in part on citizens' ability to evaluate the authenticity of the political content they consume. This study demonstrates that, under realistic conditions, that ability is essentially absent. Voters cannot tell the difference between a real political video and a synthetic one; and critically, they do not know that they cannot. Confident viewers were the most likely to be wrong, meaning that exposure to deepfake content may produce not only misinformed opinions, but a false sense of certainty in those opinions.

Equally concerning is the finding that familiarity with deepfakes provides no meaningful protection. This undermines a common assumption embedded in current media literacy policy—that educating the public about the existence and nature of deepfakes will reduce their impact. The data suggest otherwise. Awareness is not detection. A voter who knows deepfakes exist is no better equipped to spot one than a voter who has never heard of them.

No demographic group tested showed a meaningful advantage. The vulnerability cuts across age, political affiliation, and prior exposure to synthetic media. This rules out the possibility that deepfake influence is concentrated in identifiable, targetable populations, and instead frames it as a systemic risk to the democratic process as a whole.

The technology producing deepfakes continues to improve. The results documented here were produced using publicly available tools—the same tools accessible to anyone with a computer and an intent to deceive. As generation quality increases, detection will only become harder. The data leave little room for doubt: the potential threat to the information environment, voter confidence, and election integrity posed by deepfakes is immense and requires urgent mitigative action.

Gary R. Herbert Institute Staff and Advisors

LEADERSHIP TEAM
Gary R. Herbert, Founder, 17th Governor of Utah
Justin Jones, MS, Senior Director
Jonathan Barton, Project Specialist
David Connelly, PhD, Research Director, Associate Professor, History/Political Science
Karen Gill, Events Director
Jim Moss, EdD, JD, Professional in Residence, Constitutional Federalism Initiative
Erik Nystul, MPA, Program Director, Government Internships
Becca Aylworth Wright, Director of Academic Engagement and Publications

ASSOCIATED LEADERS
Dan Dimond, Sr. Director Institutional Advancement, UVU Foundation
Liv Moffat, Development Director, Herbert Foundation

FACULTY FELLOWS
Tara B. B. Bishop, PhD, Assist. Prof. Earth Science / Enviro Mgmt., Earth Sciences
Alan Parry, PhD, Assoc. Prof. Mathematics
George Rudolph, PhD, Chair/Professor, Computer Science
Majid Memari, PhD, Assistant Professor, Computer Science
David Benson, PhD, Assistant Professor of Strategy, Management
Hilary Hungerford, PhD, Associate Professor of Geography
Paul Monson, AIA, NCARB, Assistant Professor of Architecture

RESEARCH ASSISTANTS
Reese Boardman, Data Governance
Kayla Cullimore, Federalism Initiative
Kaya Frazier, GIS Mapping
Karissa Gillespie, Mail-in Ballot Study
Xander Greenwood, Utah Lake/Conservation
Josh Hadlock, Data Governance AI Developer
Jaxon Hill, Mail-in Ballot Study
Ethan Howlett, Data Governance AI Developer
Anthony Jankovic, Data Governance AI Developer
Suzette Johnson, Mail-in Ballot Study
Avery LaGiglia, Mail-in Ballot Study
Keomony Mary, Synthetic Data Project
Christ Hermes Louyoko Nianza, Federalism Intergovernmental/ GIS Mapping
Bradley Smith, Data Governance
Dori Whitehead, Mail-in Ballot Study

EVENT ASSISTANTS
Emilee Cook
Jade Haymore

COMMUNICATIONS/GRAPHIC DESIGN ASSISTANTS
Catalina Aroca Baquero, Social Media
Alexa Herrera, Communications
Kaitlyn Mayo, Graphic Design
Sophia Otis, Web Designer
Brooke Peterson, Graphic Design
Rebecca Whyte, Communications