



Deepfake Videos and Impacts on Education

Montgomery County Community College Tech Conference
September 24, 2021

Presenters



Conrad Tucker
Carnegie Mellon University
conradt@andrew.cmu.edu



Lance Bush
Challenger Center
lbush@challenger.org



Christopher Doss
RAND Corporation
cdoss@rand.org

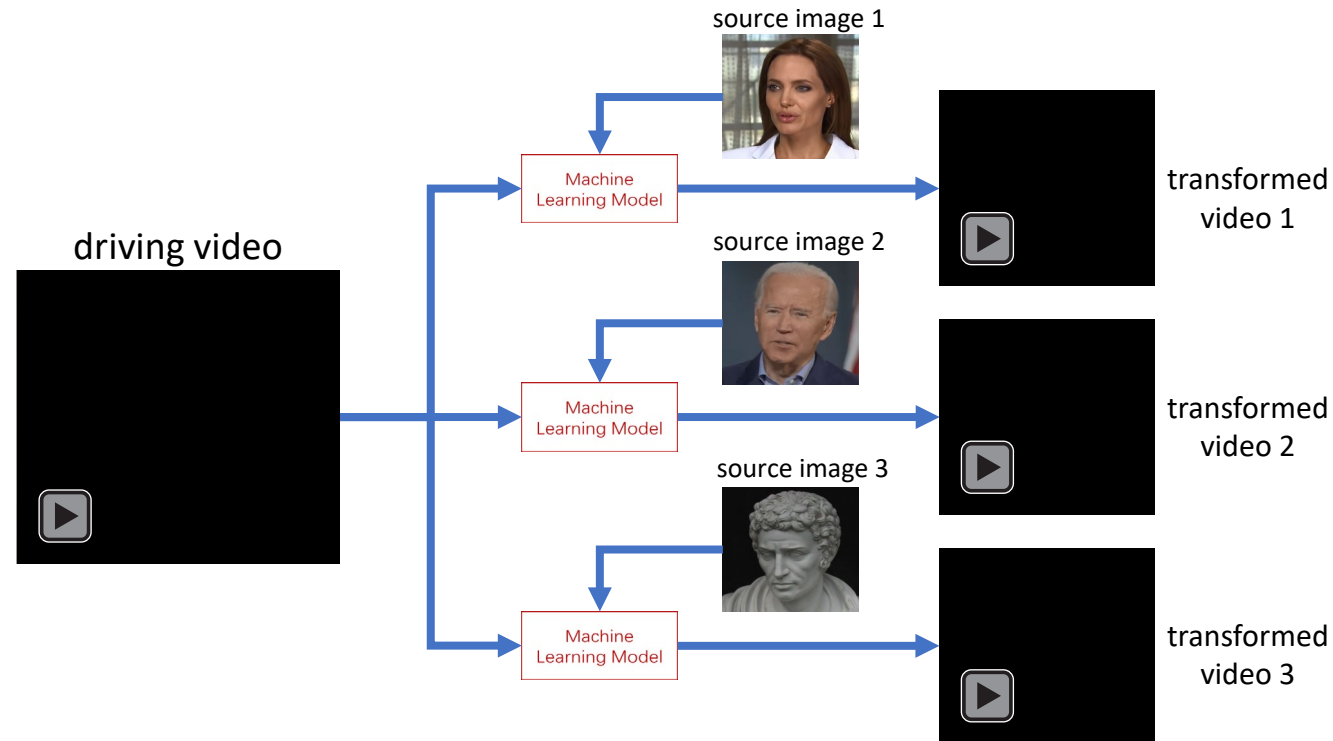
Dule Shu
Carnegie Mellon University
dules@andrew.cmu.edu

Denise Kopeccky
Challenger Center
dkopeccky@challenger.org

Valerie A. Fitton-Kane
Challenger Center
vfittonkane@challenger.org

Jared Mondschein
RAND Corporation
jmondsch@rand.org

AI has exacerbated the misinformation problem by making it **more accessible** and **scalable**



Model Reference: Siarohin, A., Lathuilière, S., Tulyakov, S., Ricci, E. and Sebe, N., 2020.
First order motion model for image animation. *arXiv preprint arXiv:2003.00196*.

Research Questions

- 1 How well can each population of educational stakeholders detect manipulated videos discussing climate change?
- 2 What characteristics of individuals moderate their ability to detect manipulated videos discussing climate change?
- 3 What characteristics of the videos do stakeholders analyze when deciding if the videos are authentic or manipulated?

Sample and Design

- Embedded randomized controlled trials into surveys
 - Each respondent is shown four videos
 - Each video is either randomly authentic or manipulated
- Fielded surveys to 5 populations of education stakeholders
 - ~800 middle school students (10-12 years old) in KY, ME, and MO
 - ~100 undergraduate and graduate students at Carnegie Mellon University
 - ~680 K-12 teachers from the American Teacher Panel
 - ~750 K-12 principals from the American School Leader Panel
 - ~750 US adults from American Life Panel



Research results are preliminary and not peer reviewed.

Survey Instrument



Tim Gallaudet, fmr. NOAA acting administrator



Richard Lindzen, Professor Emeritus of Meteorology at MIT



Greta Thunberg, climate change activist



Naomi Seibt, anti-climate change activist

Research results are preliminary and not peer reviewed.

Survey Instrument

1

What characteristics of digital content moderate an individual's ability to correctly distinguish between authentic and fake digital media?

Factors Investigated

- Visual and audio quality
- Foreground and background
- Veracity of media content

2

What characteristics of individuals moderate their vulnerability to deepfakes?

Factors Investigated

- *A priori* subject knowledge
- Political orientation
- Learning habits
- News consumption habits
- Social media usage
- Perception of the prevalence & impact of manipulated digital content

Research results are preliminary and not peer reviewed.



Try it Out

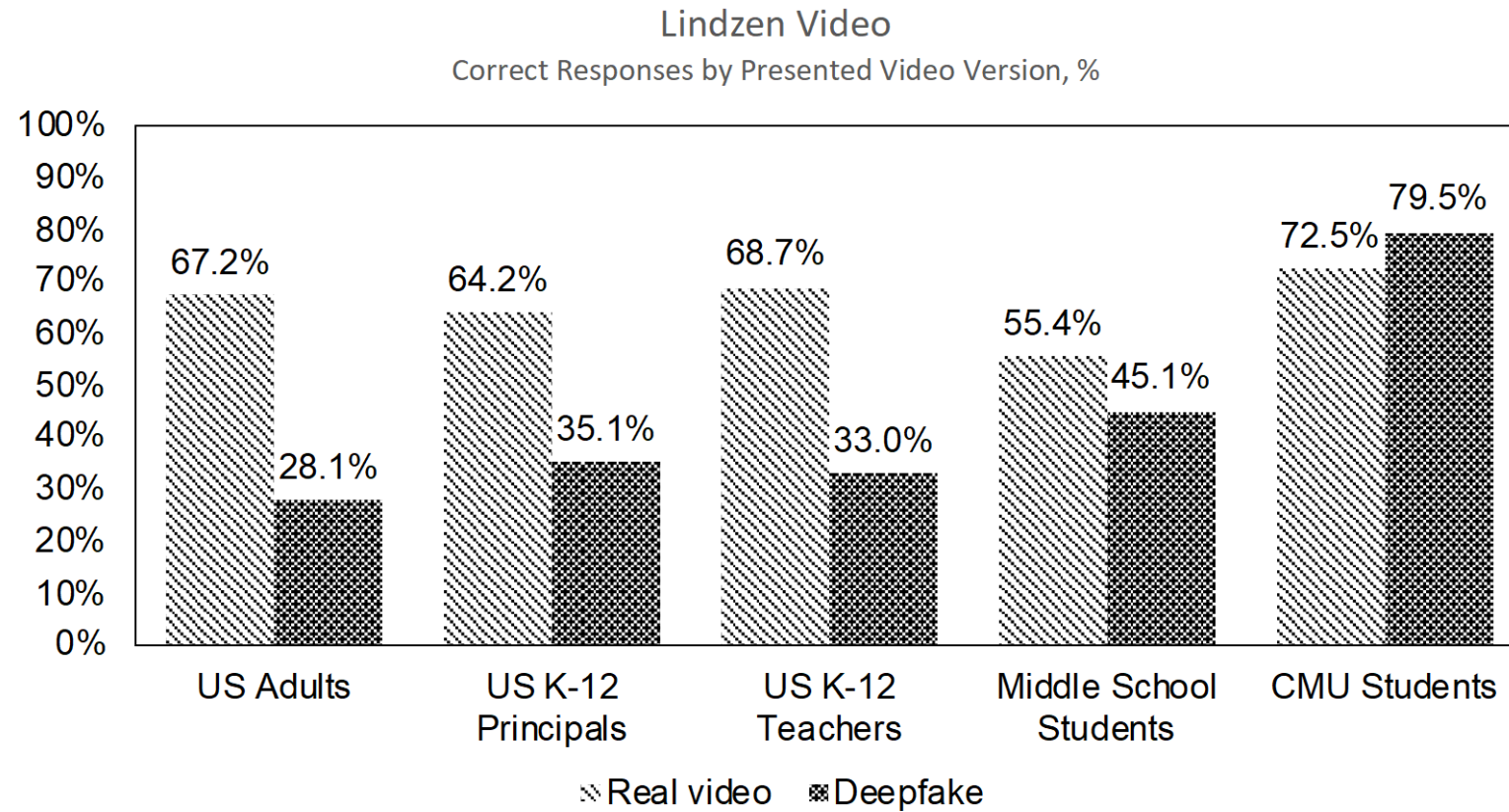
Richard Lindzen, Professor Emeritus of Meteorology at MIT

Do you think this video is authentic or manipulated?
What aspects of the video did you look at?



*Video replaced with still image as it may be used again in future research.
Research results are preliminary and not peer reviewed.*

Lindzen Results

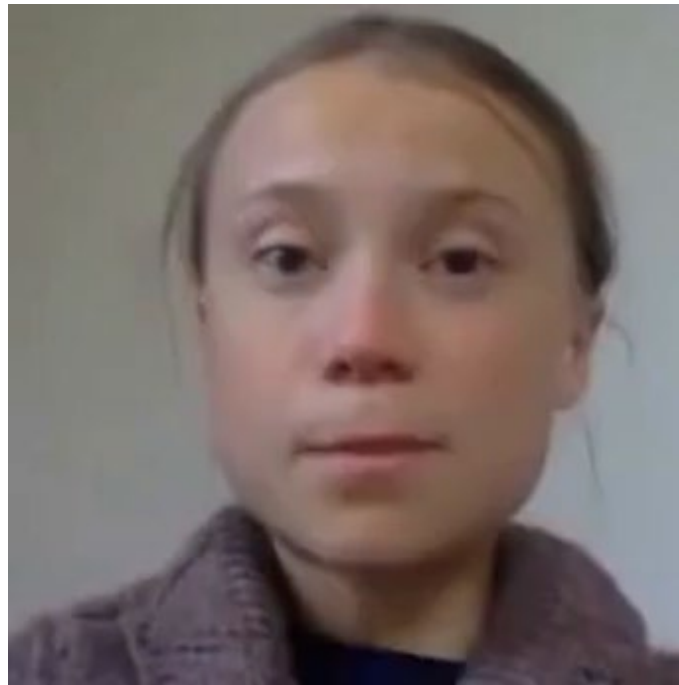


Research results are preliminary and not peer reviewed.

Try it Out

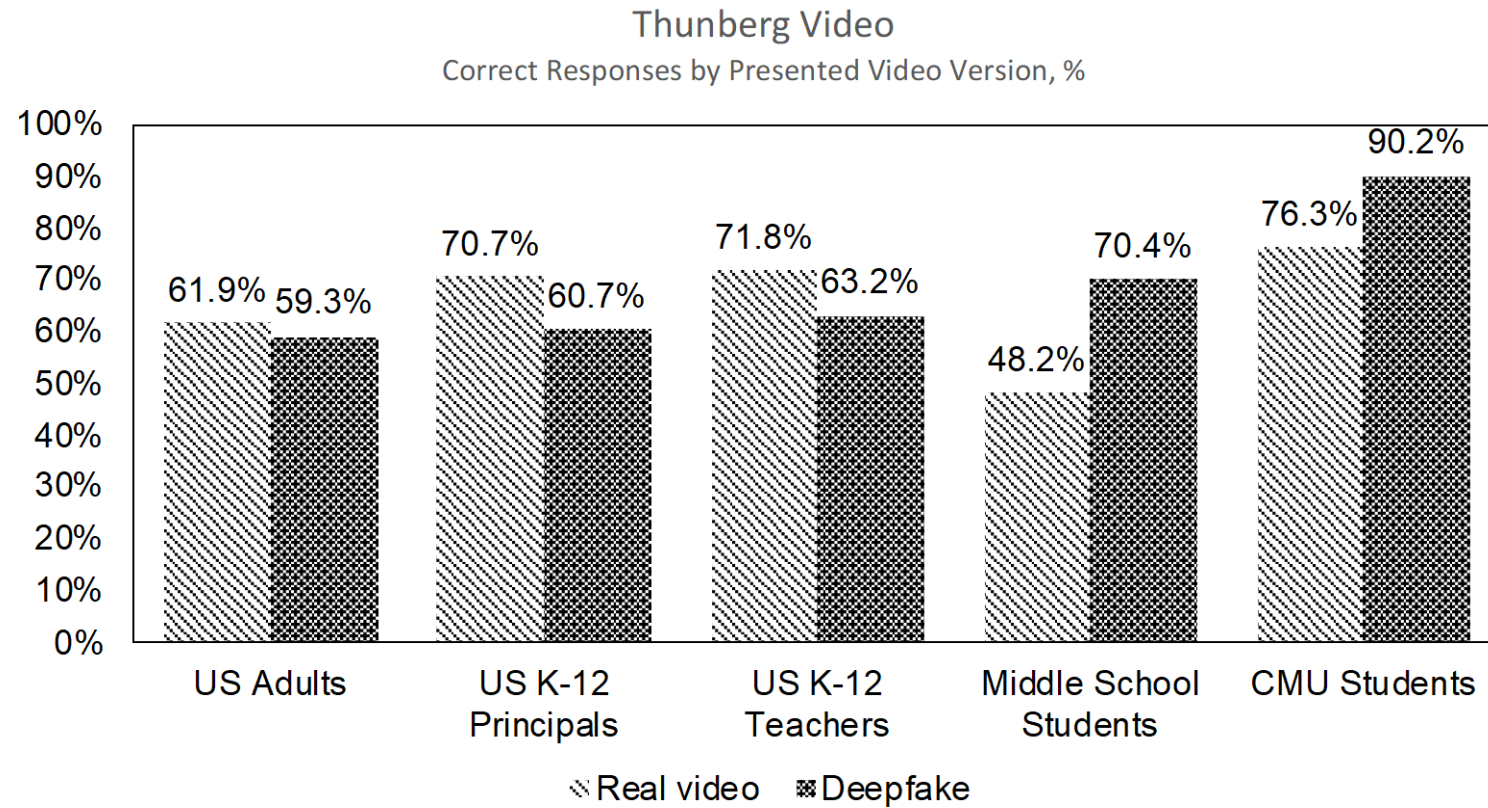
Greta Thunberg, Activist

Do you think this video is authentic or manipulated?
What aspects of the video did you look at?



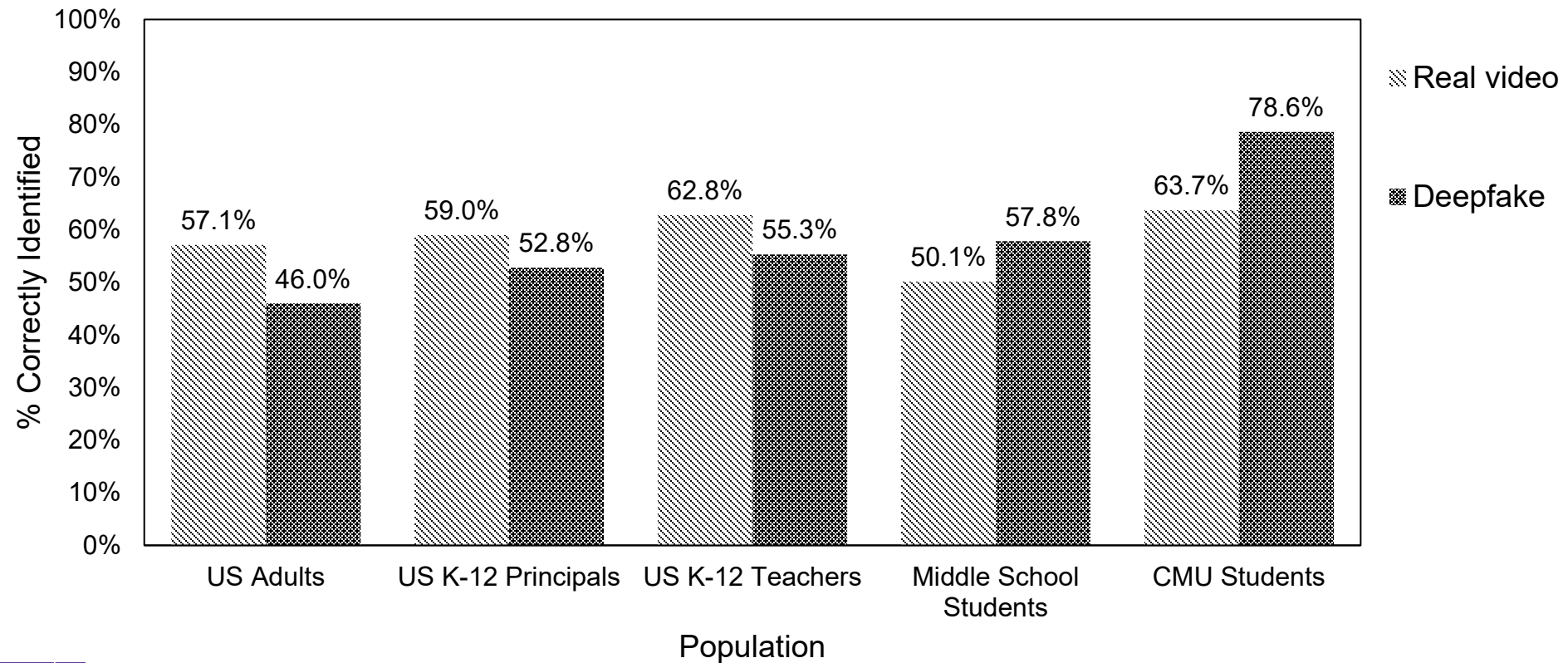
*Video replaced with still image as it may be used again in future research.
Research results are preliminary and not peer reviewed.*

Thunberg Results



Research results are preliminary and not peer reviewed.

Vulnerability of Youth and Adult Learners to Climate Change Deepfakes



Odds of correctly identifying deepfakes are essentially a coin toss

*Research results
are preliminary
and not peer
reviewed.*



Individual Characteristics That Moderate Vulnerability Interaction Terms on Probability of Correctly Identifying Fake Videos

	US Adults	Principals	Teachers	Middle Schoolers	Carnegie Mellon Students
Race (vs. all other)					
Black	-0.030 (0.106)	0.155*** (0.059)	0.015 (0.043)	0.044 (0.073)	0.586*** (0.211)
Hispanic	0.191** (0.080)	0.023 (0.078)	-0.011 (0.045)	0.125 (0.093)	-0.414* (0.237)
Political Orientation (vs. Liberal)					
Conservative	-0.156** (0.065)	-0.109** (0.050)	0.125*** (0.045)	N/A N/A	0.107 (0.205)
Moderate	-0.089 (0.073)	-0.126** (0.049)	0.151*** (0.052)	N/A N/A	0.010 (0.161)
Prefer Not to Say	-0.125 (0.109)	0.060 (0.065)	0.058 (0.064)	N/A N/A	-0.219 (0.158)
Trust in Information Sources					
Trust in Information Sources	-0.027 (0.024)	-0.046** (0.020)	-0.058*** (0.020)	-0.054*** (0.019)	-0.164*** (0.058)
Order of Videos Seen (vs. 1st)					
2nd	0.023 (0.078)	-0.117** (0.054)	-0.123** (0.059)		
3rd	0.042 (0.083)	-0.145*** (0.054)	-0.233*** (0.057)	Coming Soon	Coming Soon
4th	0.009 (0.085)	-0.140*** (0.053)	-0.195*** (0.054)		

Research results
are preliminary
and not peer
reviewed.

Individual Characteristics That Do Not Moderate Vulnerability

- Education levels
- Income
- Number of online platforms used to learn
- Knowledge of climate change
- Personal view of global warming
- Personal view of scientific consensus on global warming
- Personal view on risks posed by deepfakes
- Occupation (within ALP)
- Urbanicity
- Grades taught, subjects taught, title I eligibility of school (for educators)



Research results are preliminary and not peer reviewed.

Questions