

# Re-visiting the Political Deepfake Videos Misinformation the Public But No More than Other Fake Media (2021)\*

Yilin Zhu, Jialin Zhao, Renjing Liu

23/03/2021

## Contents

<b>1</b>	<b>Abstract</b>	<b>2</b>
<b>2</b>	<b>Introduction</b>	<b>2</b>
<b>3</b>	<b>Experimental Design</b>	<b>3</b>
<b>4</b>	<b>Data and Model</b>	<b>4</b>
<b>5</b>	<b>Results</b>	<b>6</b>
<b>6</b>	<b>Discussion</b>	<b>9</b>
6.1	First experiment . . . . .	9
6.2	Second experiment . . . . .	9
<b>7</b>	<b>Limitation and Future works</b>	<b>11</b>
<b>8</b>	<b>Appendix</b>	<b>12</b>
8.1	Appendix A . . . . .	12
8.2	Appendix B . . . . .	12
8.3	Appendix C . . . . .	12
8.4	Appendix D . . . . .	15
	<b>Reference</b>	<b>15</b>

---

\*Code and Data are available at: <https://github.com/PopeyesIsTheBest/PoliticalDeepFakeVideos.git>

# 1 Abstract

## 2 Introduction

With the continuous development of Artificial intelligence and deep learning in scientific research and applications, technologies have greatly affected many aspects of our lives. Recently, the deepfake, a term derived from the combination of “deep learning” and “fake”, which leverages open-source deep learning technology to produce a kind of inveroacious content, have been used to fabricate video and audio recordings of political elites around the world, showing and saying things they have never done. And some reports have reported these videos could even lead to a coup attempt, posing remarkable threats and potential harm on the political democracy (Bio 2019). Moreover, some also expressed concerns that deepfake technology may also undermine our ability to detect both honest and dishonest claims in the real word (Rini 2020).

Based on these impacts and concerns caused by deepfake, Soubhik Barari, Christopher Lucas and Kevin Munger (Barari, Lucas, and Munger 2021a) compared the power of persuasion between deepfake video and an identical political scandal presented in some other comparable media formats such as textual headlines or audio recordings, and they demonstrated that deepfakes are no more effective than the same misinformation produced by other media formats, being not uniquely powerful at deception or affective manipulation (Barari, Lucas, and Munger 2021a). Additionally, they (Barari, Lucas, and Munger 2021a) also investigated the heterogeneity by participants’ characteristics. It is noticeable that the authors (Barari, Lucas, and Munger 2021a) registered the older adults as the subgroups with hypothesized susceptibility to the deepfake detection, assuming they might be unable to evaluate the accuracy of digital information. And in their first experiment (exposure), which compares the attitudinal effects (the extent of fake in the clips and the change in effect to the target of the clips) of scandal clips by different subgroups of respondents, they found ‘although the elderly are more likely to be effectively triggered by the Warren fake media clippings relative to those below 65, there is no detectable difference in deception nor across media.’(Barari, Lucas, and Munger 2021a)

However, in the second experiment (Detection), which examines the ability of respondents to discriminate between real and fake clippings, they (Barari, Lucas, and Munger 2021a) did not perform comparisons across different age groups, but compared the outcome differences across several motivated reasoning factors (cognitive reflection, political knowledge, digital knowledge and partisan identity) instead (Barari, Lucas, and Munger 2021a).

On the basis of what Soubhik Barari, Christopher Lucas and Kevin Munger (Barari, Lucas, and Munger 2021a) have investigated for heterogeneity by the elderly and those below 65. We adjust the code, registering the young people (18-24 years old) as the subgroups with hypothesized susceptibility to deepfake detection, to perform comparisons in outcomes of dual experiments across young people and those older than 24. The reproducing materials provided by Soubhik Barari (Barari, Lucas, and Munger 2021b), allow us to produce this narrow replication of Soubhik Barari, Christopher Lucas and Kevin Munger (Barari, Lucas, and Munger 2021a).

There are some reasons motivating us to perform this relocation. According to (Barnwell 2020), young people ageing between 18 and 24 years old, are the largest age group who engage with countless images and videos that have been edited or digitally altered. For instance, a lot of young people are exposed to ironic TikTok videos. And hence, these people are more likely to touch with deepfake every day, compared with older adults. However, we might suspect whether they are aware of or have they thought deeply about the disinformation online? Moreover, we are also wondering about their ability to discriminate between real and fake information.

We acknowledge and are grateful for the replication materials provided by the authors of Soubhik Barari, Christopher Lucas and Kevin Munger (Barari, Lucas, and Munger 2021a). The remainder of this paper is formulated as follows. Section 2 will briefly talk about the experimental design. Section 3 will generally discuss the data and model relevant to our replication. Section 4 will summarize our associated findings, and finally Section 5 will provide more discussions for our findings.

### 3 Experimental Design

Barari, Lucas and Munger did 2 experiments in this study. The first experiment is an exposure experiment of a 2 x 6 pairing factorial design. Some of the randomly chosen respondents are informed about deepfakes while the rest are not. Then they are all exposed to a single fictitious news feed which is randomly chosen from any one of the following 6 situations: text, audio, skit, video, ad, no clip at all (the control group), and the videos used are deepfakes (Barari, Lucas, and Munger 2021a).

During the experiment, a natural “news-feed” environment is implemented: each respondent watches 6 media clips in order and the third news feed is a fictitious one, and the rest clips are real. An actress is paid to perform as Elizabeth Warren in the skit and uses the same audio recording as the audio condition. The deepfake video is constructed from the footage used in the skit condition. This setting is very important because it enhances the internal validity of this experiment. If video conditions present different results compared to the audio and the text conditions but not with the skit condition, then this means the audiovisual information is most persuasive, no matter if the video is fictionalized or not (Barari, Lucas, and Munger 2021a).

The following experiment is a detection experiment: same respondents are asked to scroll through 8 news videos which allows us to find the within-individual deception rate. Before the task, half of the respondents are debriefed about whether they are exposed to deepfake videos in the first experiment, while the other half are not. Later, half of the subjects are also provided an accuracy prime (Barari, Lucas, and Munger 2021a). For this experiment, since subjects may watch the same deepfake videos before, an upward bias may exist here although none of them mention this in open feedback. All respondents are randomly assigned into 3 situations with different levels of deepfakes used to construct the video: 75% (high-fake), 25% (low-fake) and 0% (no-fake) (Barari, Lucas, and Munger 2021a).

## 4 Data and Model

The experiments described above are two survey experiments on the Lucid survey research platform, which eventually collects a sample size of 5,750 valid respondents out of 17,501 total participants in September 2020 and October 2020 (Barari, Lucas, and Munger 2021a). Thus the dataset includes the responses from these subjects in both experiments, plus any characteristics that can potentially affect the deepfake deception and appeal. Namely these parameters include some demographic information such as age, gender, education, household income, race, and ethnicity. These characteristics are hypothesized to be highly relevant to the experiment results due to their correlation with digital literacy, internet usage, political knowledge and partisanship (Barari, Lucas, and Munger 2021a).

Since this is a Lucid survey experiment, Barari, Lucas and Munger also introduce a series of “technology checks” to ensure that respondents are actually able to watch and listen to videos. Besides the technology checks, pre-experimental attention checks are also applied here to do a brief review on whether the answers to some basic demographic characteristics (e.g., gender and age) from respondents are consistent with the characteristics provided by Lucid. If the two answers do not match up, these respondents are then labeled as “low-quality” respondents, which will be dropped later when applying the statistical models as a robustness measure (Barari, Lucas, and Munger 2021a).

Another safeguard for the validity of this experiment is the representativeness of the data. Barari, Lucas, & Munger adjust the distribution of the originally collected data by using raking to calculate post-stratification weights, which tries to duplicate the demographics traits in the most recent Current Population Survey (CPS). Similar to the “low-quality” labels, weighted regression also acts as a robustness measure in later analyses. Remember that some demographics are correlated to the experiment results, so this step is very important because it eliminates the bias in results and enhances the external validity of this experiment.

Three methods in total are used in each discussion of the experiment results and they are reproduced in a very similar way as Barari, Lucas and Munger did. While their research question focuses on comparing the deepfake videos with other media formats, our research narrows it down to examine the differences between age groups. So our models compare the differences between age groups instead of which media condition the subject exposed to as Barari, Lucas and Munger did.

First there is a figure showing the value of the estimate (namely, the mean value) and the 95% confidence interval of the estimate for different age groups. The 95% confidence interval is constructed by adding and subtracting  $1.96 \times \text{estimate}$ , which requires the assumption of normality. Since the data is sufficiently large, and each response is independent, so the normality assumption is relaxed here and thus this analysis is reliable. Next is a non-parametric test (t-test) to directly compare the difference in mean values among various age groups, which also leads to valid interpretation based on the same reasons as the confidence interval. Finally multiple linear regressions are performed as a robustness measure (Barari, Lucas, and Munger 2021a).

In the exposure experiment, the regression model examining the age effect on deception is estimated via (Barari, Lucas, and Munger 2021a):

$$\begin{aligned}
 \text{Believe}_{i} = & \\
 & \alpha + \beta_1 \text{Agegroup}_{i1} + \beta_2 \text{Agegroup}_{i2} + \beta_3 \text{Agegroup}_{i3} + \beta_4 \text{Agegroup}_{i4} + \beta_5 \text{Agegroup}_{i5} + \\
 & \beta_6 \text{Agegroup}_{i6} + \beta_7 \text{Agegroup}_{i7} + \beta_8 \text{Agegroup}_{i8} + \beta_9 \text{Agegroup}_{i9} + \beta_{10} \text{Agegroup}_{i10} + \beta_{11} \text{Agegroup}_{i11} + \beta_{12} \text{Agegroup}_{i12} + \beta_{13} \text{Agegroup}_{i13} + \beta_{14} \text{Agegroup}_{i14} + \beta_{15} \text{Agegroup}_{i15} + \beta_{16} \text{Agegroup}_{i16} + \beta_{17} \text{Agegroup}_{i17} + \beta_{18} \text{Agegroup}_{i18} + \beta_{19} \text{Agegroup}_{i19} + \beta_{20} \text{Agegroup}_{i20} + \beta_{21} \text{Agegroup}_{i21} + \beta_{22} \text{Agegroup}_{i22} + \beta_{23} \text{Agegroup}_{i23} + \beta_{24} \text{Agegroup}_{i24} + \beta_{25} \text{Agegroup}_{i25} + \beta_{26} \text{Agegroup}_{i26} + \beta_{27} \text{Agegroup}_{i27} + \beta_{28} \text{Agegroup}_{i28} + \beta_{29} \text{Agegroup}_{i29} + \beta_{30} \text{Agegroup}_{i30} + \beta_{31} \text{Agegroup}_{i31} + \beta_{32} \text{Agegroup}_{i32} + \beta_{33} \text{Agegroup}_{i33} + \beta_{34} \text{Agegroup}_{i34} + \beta_{35} \text{Agegroup}_{i35} + \beta_{36} \text{Agegroup}_{i36} + \beta_{37} \text{Agegroup}_{i37} + \beta_{38} \text{Agegroup}_{i38} + \beta_{39} \text{Agegroup}_{i39} + \beta_{40} \text{Agegroup}_{i40} + \beta_{41} \text{Agegroup}_{i41} + \beta_{42} \text{Agegroup}_{i42} + \beta_{43} \text{Agegroup}_{i43} + \beta_{44} \text{Agegroup}_{i44} + \beta_{45} \text{Agegroup}_{i45} + \beta_{46} \text{Agegroup}_{i46} + \beta_{47} \text{Agegroup}_{i47} + \beta_{48} \text{Agegroup}_{i48} + \beta_{49} \text{Agegroup}_{i49} + \beta_{50} \text{Agegroup}_{i50} + \beta_{51} \text{Agegroup}_{i51} + \beta_{52} \text{Agegroup}_{i52} + \beta_{53} \text{Agegroup}_{i53} + \beta_{54} \text{Agegroup}_{i54} + \beta_{55} \text{Agegroup}_{i55} + \beta_{56} \text{Agegroup}_{i56} + \beta_{57} \text{Agegroup}_{i57} + \beta_{58} \text{Agegroup}_{i58} + \beta_{59} \text{Agegroup}_{i59} + \beta_{60} \text{Agegroup}_{i60} + \beta_{61} \text{Agegroup}_{i61} + \beta_{62} \text{Agegroup}_{i62} + \beta_{63} \text{Agegroup}_{i63} + \beta_{64} \text{Agegroup}_{i64} + \beta_{65} \text{Agegroup}_{i65} + \beta_{66} \text{Agegroup}_{i66} + \beta_{67} \text{Agegroup}_{i67} + \beta_{68} \text{Agegroup}_{i68} + \beta_{69} \text{Agegroup}_{i69} + \beta_{70} \text{Agegroup}_{i70} + \beta_{71} \text{Agegroup}_{i71} + \beta_{72} \text{Agegroup}_{i72} + \beta_{73} \text{Agegroup}_{i73} + \beta_{74} \text{Agegroup}_{i74} + \beta_{75} \text{Agegroup}_{i75} + \beta_{76} \text{Agegroup}_{i76} + \beta_{77} \text{Agegroup}_{i77} + \beta_{78} \text{Agegroup}_{i78} + \beta_{79} \text{Agegroup}_{i79} + \beta_{80} \text{Agegroup}_{i80} + \beta_{81} \text{Agegroup}_{i81} + \beta_{82} \text{Agegroup}_{i82} + \beta_{83} \text{Agegroup}_{i83} + \beta_{84} \text{Agegroup}_{i84} + \beta_{85} \text{Agegroup}_{i85} + \beta_{86} \text{Agegroup}_{i86} + \beta_{87} \text{Agegroup}_{i87} + \beta_{88} \text{Agegroup}_{i88} + \beta_{89} \text{Agegroup}_{i89} + \beta_{90} \text{Agegroup}_{i90} + \beta_{91} \text{Agegroup}_{i91} + \beta_{92} \text{Agegroup}_{i92} + \beta_{93} \text{Agegroup}_{i93} + \beta_{94} \text{Agegroup}_{i94} + \beta_{95} \text{Agegroup}_{i95} + \beta_{96} \text{Agegroup}_{i96} + \beta_{97} \text{Agegroup}_{i97} + \beta_{98} \text{Agegroup}_{i98} + \beta_{99} \text{Agegroup}_{i99} + \beta_{100} \text{Agegroup}_{i100} + \beta_{101} \text{Agegroup}_{i101} + \beta_{102} \text{Agegroup}_{i102} + \beta_{103} \text{Agegroup}_{i103} + \beta_{104} \text{Agegroup}_{i104} + \beta_{105} \text{Agegroup}_{i105} + \beta_{106} \text{Agegroup}_{i106} + \beta_{107} \text{Agegroup}_{i107} + \beta_{108} \text{Agegroup}_{i108} + \beta_{109} \text{Agegroup}_{i109} + \beta_{110} \text{Agegroup}_{i110} + \beta_{111} \text{Agegroup}_{i111} + \beta_{112} \text{Agegroup}_{i112} + \beta_{113} \text{Agegroup}_{i113} + \beta_{114} \text{Agegroup}_{i114} + \beta_{115} \text{Agegroup}_{i115} + \beta_{116} \text{Agegroup}_{i116} + \beta_{117} \text{Agegroup}_{i117} + \beta_{118} \text{Agegroup}_{i118} + \beta_{119} \text{Agegroup}_{i119} + \beta_{120} \text{Agegroup}_{i120} + \beta_{121} \text{Agegroup}_{i121} + \beta_{122} \text{Agegroup}_{i122} + \beta_{123} \text{Agegroup}_{i123} + \beta_{124} \text{Agegroup}_{i124} + \beta_{125} \text{Agegroup}_{i125} + \beta_{126} \text{Agegroup}_{i126} + \beta_{127} \text{Agegroup}_{i127} + \beta_{128} \text{Agegroup}_{i128} + \beta_{129} \text{Agegroup}_{i129} + \beta_{130} \text{Agegroup}_{i130} + \beta_{131} \text{Agegroup}_{i131} + \beta_{132} \text{Agegroup}_{i132} + \beta_{133} \text{Agegroup}_{i133} + \beta_{134} \text{Agegroup}_{i134} + \beta_{135} \text{Agegroup}_{i135} + \beta_{136} \text{Agegroup}_{i136} + \beta_{137} \text{Agegroup}_{i137} + \beta_{138} \text{Agegroup}_{i138} + \beta_{139} \text{Agegroup}_{i139} + \beta_{140} \text{Agegroup}_{i140} + \beta_{141} \text{Agegroup}_{i141} + \beta_{142} \text{Agegroup}_{i142} + \beta_{143} \text{Agegroup}_{i143} + \beta_{144} \text{Agegroup}_{i144} + \beta_{145} \text{Agegroup}_{i145} + \beta_{146} \text{Agegroup}_{i146} + \beta_{147} \text{Agegroup}_{i147} + \beta_{148} \text{Agegroup}_{i148} + \beta_{149} \text{Agegroup}_{i149} + \beta_{150} \text{Agegroup}_{i150} + \beta_{151} \text{Agegroup}_{i151} + \beta_{152} \text{Agegroup}_{i152} + \beta_{153} \text{Agegroup}_{i153} + \beta_{154} \text{Agegroup}_{i154} + \beta_{155} \text{Agegroup}_{i155} + \beta_{156} \text{Agegroup}_{i156} + \beta_{157} \text{Agegroup}_{i157} + \beta_{158} \text{Agegroup}_{i158} + \beta_{159} \text{Agegroup}_{i159} + \beta_{160} \text{Agegroup}_{i160} + \beta_{161} \text{Agegroup}_{i161} + \beta_{162} \text{Agegroup}_{i162} + \beta_{163} \text{Agegroup}_{i163} + \beta_{164} \text{Agegroup}_{i164} + \beta_{165} \text{Agegroup}_{i165} + \beta_{166} \text{Agegroup}_{i166} + \beta_{167} \text{Agegroup}_{i167} + \beta_{168} \text{Agegroup}_{i168} + \beta_{169} \text{Agegroup}_{i169} + \beta_{170} \text{Agegroup}_{i170} + \beta_{171} \text{Agegroup}_{i171} + \beta_{172} \text{Agegroup}_{i172} + \beta_{173} \text{Agegroup}_{i173} + \beta_{174} \text{Agegroup}_{i174} + \beta_{175} \text{Agegroup}_{i175} + \beta_{176} \text{Agegroup}_{i176} + \beta_{177} \text{Agegroup}_{i177} + \beta_{178} \text{Agegroup}_{i178} + \beta_{179} \text{Agegroup}_{i179} + \beta_{180} \text{Agegroup}_{i180} + \beta_{181} \text{Agegroup}_{i181} + \beta_{182} \text{Agegroup}_{i182} + \beta_{183} \text{Agegroup}_{i183} + \beta_{184} \text{Agegroup}_{i184} + \beta_{185} \text{Agegroup}_{i185} + \beta_{186} \text{Agegroup}_{i186} + \beta_{187} \text{Agegroup}_{i187} + \beta_{188} \text{Agegroup}_{i188} + \beta_{189} \text{Agegroup}_{i189} + \beta_{190} \text{Agegroup}_{i190} + \beta_{191} \text{Agegroup}_{i191} + \beta_{192} \text{Agegroup}_{i192} + \beta_{193} \text{Agegroup}_{i193} + \beta_{194} \text{Agegroup}_{i194} + \beta_{195} \text{Agegroup}_{i195} + \beta_{196} \text{Agegroup}_{i196} + \beta_{197} \text{Agegroup}_{i197} + \beta_{198} \text{Agegroup}_{i198} + \beta_{199} \text{Agegroup}_{i199} + \beta_{200} \text{Agegroup}_{i200} + \beta_{201} \text{Agegroup}_{i201} + \beta_{202} \text{Agegroup}_{i202} + \beta_{203} \text{Agegroup}_{i203} + \beta_{204} \text{Agegroup}_{i204} + \beta_{205} \text{Agegroup}_{i205} + \beta_{206} \text{Agegroup}_{i206} + \beta_{207} \text{Agegroup}_{i207} + \beta_{208} \text{Agegroup}_{i208} + \beta_{209} \text{Agegroup}_{i209} + \beta_{210} \text{Agegroup}_{i210} + \beta_{211} \text{Agegroup}_{i211} + \beta_{212} \text{Agegroup}_{i212} + \beta_{213} \text{Agegroup}_{i213} + \beta_{214} \text{Agegroup}_{i214} + \beta_{215} \text{Agegroup}_{i215} + \beta_{216} \text{Agegroup}_{i216} + \beta_{217} \text{Agegroup}_{i217} + \beta_{218} \text{Agegroup}_{i218} + \beta_{219} \text{Agegroup}_{i219} + \beta_{220} \text{Agegroup}_{i220} + \beta_{221} \text{Agegroup}_{i221} + \beta_{222} \text{Agegroup}_{i222} + \beta_{223} \text{Agegroup}_{i223} + \beta_{224} \text{Agegroup}_{i224} + \beta_{225} \text{Agegroup}_{i225} + \beta_{226} \text{Agegroup}_{i226} + \beta_{227} \text{Agegroup}_{i227} + \beta_{228} \text{Agegroup}_{i228} + \beta_{229} \text{Agegroup}_{i229} + \beta_{230} \text{Agegroup}_{i230} + \beta_{231} \text{Agegroup}_{i231} + \beta_{232} \text{Agegroup}_{i232} + \beta_{233} \text{Agegroup}_{i233} + \beta_{234} \text{Agegroup}_{i234} + \beta_{235} \text{Agegroup}_{i235} + \beta_{236} \text{Agegroup}_{i236} + \beta_{237} \text{Agegroup}_{i237} + \beta_{238} \text{Agegroup}_{i238} + \beta_{239} \text{Agegroup}_{i239} + \beta_{240} \text{Agegroup}_{i240} + \beta_{241} \text{Agegroup}_{i241} + \beta_{242} \text{Agegroup}_{i242} + \beta_{243} \text{Agegroup}_{i243} + \beta_{244} \text{Agegroup}_{i244} + \beta_{245} \text{Agegroup}_{i245} + \beta_{246} \text{Agegroup}_{i246} + \beta_{247} \text{Agegroup}_{i247} + \beta_{248} \text{Agegroup}_{i248} + \beta_{249} \text{Agegroup}_{i249} + \beta_{250} \text{Agegroup}_{i250} + \beta_{251} \text{Agegroup}_{i251} + \beta_{252} \text{Agegroup}_{i252} + \beta_{253} \text{Agegroup}_{i253} + \beta_{254} \text{Agegroup}_{i254} + \beta_{255} \text{Agegroup}_{i255} + \beta_{256} \text{Agegroup}_{i256} + \beta_{257} \text{Agegroup}_{i257} + \beta_{258} \text{Agegroup}_{i258} + \beta_{259} \text{Agegroup}_{i259} + \beta_{260} \text{Agegroup}_{i260} + \beta_{261} \text{Agegroup}_{i261} + \beta_{262} \text{Agegroup}_{i262} + \beta_{263} \text{Agegroup}_{i263} + \beta_{264} \text{Agegroup}_{i264} + \beta_{265} \text{Agegroup}_{i265} + \beta_{266} \text{Agegroup}_{i266} + \beta_{267} \text{Agegroup}_{i267} + \beta_{268} \text{Agegroup}_{i268} + \beta_{269} \text{Agegroup}_{i269} + \beta_{270} \text{Agegroup}_{i270} + \beta_{271} \text{Agegroup}_{i271} + \beta_{272} \text{Agegroup}_{i272} + \beta_{273} \text{Agegroup}_{i273} + \beta_{274} \text{Agegroup}_{i274} + \beta_{275} \text{Agegroup}_{i275} + \beta_{276} \text{Agegroup}_{i276} + \beta_{277} \text{Agegroup}_{i277} + \beta_{278} \text{Agegroup}_{i278} + \beta_{279} \text{Agegroup}_{i279} + \beta_{280} \text{Agegroup}_{i280} + \beta_{281} \text{Agegroup}_{i281} + \beta_{282} \text{Agegroup}_{i282} + \beta_{283} \text{Agegroup}_{i283} + \beta_{284} \text{Agegroup}_{i284} + \beta_{285} \text{Agegroup}_{i285} + \beta_{286} \text{Agegroup}_{i286} + \beta_{287} \text{Agegroup}_{i287} + \beta_{288} \text{Agegroup}_{i288} + \beta_{289} \text{Agegroup}_{i289} + \beta_{290} \text{Agegroup}_{i290} + \beta_{291} \text{Agegroup}_{i291} + \beta_{292} \text{Agegroup}_{i292} + \beta_{293} \text{Agegroup}_{i293} + \beta_{294} \text{Agegroup}_{i294} + \beta_{295} \text{Agegroup}_{i295} + \beta_{296} \text{Agegroup}_{i296} + \beta_{297} \text{Agegroup}_{i297} + \beta_{298} \text{Agegroup}_{i298} + \beta_{299} \text{Agegroup}_{i299} + \beta_{300} \text{Agegroup}_{i300} + \beta_{301} \text{Agegroup}_{i301} + \beta_{302} \text{Agegroup}_{i302} + \beta_{303} \text{Agegroup}_{i303} + \beta_{304} \text{Agegroup}_{i304} + \beta_{305} \text{Agegroup}_{i305} + \beta_{306} \text{Agegroup}_{i306} + \beta_{307} \text{Agegroup}_{i307} + \beta_{308} \text{Agegroup}_{i308} + \beta_{309} \text{Agegroup}_{i309} + \beta_{310} \text{Agegroup}_{i310} + \beta_{311} \text{Agegroup}_{i311} + \beta_{312} \text{Agegroup}_{i312} + \beta_{313} \text{Agegroup}_{i313} + \beta_{314} \text{Agegroup}_{i314} + \beta_{315} \text{Agegroup}_{i315} + \beta_{316} \text{Agegroup}_{i316} + \beta_{317} \text{Agegroup}_{i317} + \beta_{318} \text{Agegroup}_{i318} + \beta_{319} \text{Agegroup}_{i319} + \beta_{320} \text{Agegroup}_{i320} + \beta_{321} \text{Agegroup}_{i321} + \beta_{322} \text{Agegroup}_{i322} + \beta_{323} \text{Agegroup}_{i323} + \beta_{324} \text{Agegroup}_{i324} + \beta_{325} \text{Agegroup}_{i325} + \beta_{326} \text{Agegroup}_{i326} + \beta_{327} \text{Agegroup}_{i327} + \beta_{328} \text{Agegroup}_{i328} + \beta_{329} \text{Agegroup}_{i329} + \beta_{330} \text{Agegroup}_{i330} + \beta_{331} \text{Agegroup}_{i331} + \beta_{332} \text{Agegroup}_{i332} + \beta_{333} \text{Agegroup}_{i333} + \beta_{334} \text{Agegroup}_{i334} + \beta_{335} \text{Agegroup}_{i335} + \beta_{336} \text{Agegroup}_{i336} + \beta_{337} \text{Agegroup}_{i337} + \beta_{338} \text{Agegroup}_{i338} + \beta_{339} \text{Agegroup}_{i339} + \beta_{340} \text{Agegroup}_{i340} + \beta_{341} \text{Agegroup}_{i341} + \beta_{342} \text{Agegroup}_{i342} + \beta_{343} \text{Agegroup}_{i343} + \beta_{344} \text{Agegroup}_{i344} + \beta_{345} \text{Agegroup}_{i345} + \beta_{346} \text{Agegroup}_{i346} + \beta_{347} \text{Agegroup}_{i347} + \beta_{348} \text{Agegroup}_{i348} + \beta_{349} \text{Agegroup}_{i349} + \beta_{350} \text{Agegroup}_{i350} + \beta_{351} \text{Agegroup}_{i351} + \beta_{352} \text{Agegroup}_{i352} + \beta_{353} \text{Agegroup}_{i353} + \beta_{354} \text{Agegroup}_{i354} + \beta_{355} \text{Agegroup}_{i355} + \beta_{356} \text{Agegroup}_{i356} + \beta_{357} \text{Agegroup}_{i357} + \beta_{358} \text{Agegroup}_{i358} + \beta_{359} \text{Agegroup}_{i359} + \beta_{360} \text{Agegroup}_{i360} + \beta_{361} \text{Agegroup}_{i361} + \beta_{362} \text{Agegroup}_{i362} + \beta_{363} \text{Agegroup}_{i363} + \beta_{364} \text{Agegroup}_{i364} + \beta_{365} \text{Agegroup}_{i365} + \beta_{366} \text{Agegroup}_{i366} + \beta_{367} \text{Agegroup}_{i367} + \beta_{368} \text{Agegroup}_{i368} + \beta_{369} \text{Agegroup}_{i369} + \beta_{370} \text{Agegroup}_{i370} + \beta_{371} \text{Agegroup}_{i371} + \beta_{372} \text{Agegroup}_{i372} + \beta_{373} \text{Agegroup}_{i373} + \beta_{374} \text{Agegroup}_{i374} + \beta_{375} \text{Agegroup}_{i375} + \beta_{376} \text{Agegroup}_{i376} + \beta_{377} \text{Agegroup}_{i377} + \beta_{378} \text{Agegroup}_{i378} + \beta_{379} \text{Agegroup}_{i379} + \beta_{380} \text{Agegroup}_{i380} + \beta_{381} \text{Agegroup}_{i381} + \beta_{382} \text{Agegroup}_{i382} + \beta_{383} \text{Agegroup}_{i383} + \beta_{384} \text{Agegroup}_{i384} + \beta_{385} \text{Agegroup}_{i385} + \beta_{386} \text{Agegroup}_{i386} + \beta_{387} \text{Agegroup}_{i387} + \beta_{388} \text{Agegroup}_{i388} + \beta_{389} \text{Agegroup}_{i389} + \beta_{390} \text{Agegroup}_{i390} + \beta_{391} \text{Agegroup}_{i391} + \beta_{392} \text{Agegroup}_{i392} + \beta_{393} \text{Agegroup}_{i393} + \beta_{394} \text{Agegroup}_{i394} + \beta_{395} \text{Agegroup}_{i395} + \beta_{396} \text{Agegroup}_{i396} + \beta_{397} \text{Agegroup}_{i397} + \beta_{398} \text{Agegroup}_{i398} + \beta_{399} \text{Agegroup}_{i399} + \beta_{400} \text{Agegroup}_{i400} + \beta_{401} \text{Agegroup}_{i401} + \beta_{402} \text{Agegroup}_{i402} + \beta_{403} \text{Agegroup}_{i403} + \beta_{404} \text{Agegroup}_{i404} + \beta_{405} \text{Agegroup}_{i405} + \beta_{406} \text{Agegroup}_{i406} + \beta_{407} \text{Agegroup}_{i407} + \beta_{408} \text{Agegroup}_{i408} + \beta_{409} \text{Agegroup}_{i409} + \beta_{410} \text{Agegroup}_{i410} + \beta_{411} \text{Agegroup}_{i411} + \beta_{412} \text{Agegroup}_{i412} + \beta_{413} \text{Agegroup}_{i413} + \beta_{414} \text{Agegroup}_{i414} + \beta_{415} \text{Agegroup}_{i415} + \beta_{416} \text{Agegroup}_{i416} + \beta_{417} \text{Agegroup}_{i417} + \beta_{418} \text{Agegroup}_{i418} + \beta_{419} \text{Agegroup}_{i419} + \beta_{420} \text{Agegroup}_{i420} + \beta_{421} \text{Agegroup}_{i421} + \beta_{422} \text{Agegroup}_{i422} + \beta_{423} \text{Agegroup}_{i423} + \beta_{424} \text{Agegroup}_{i424} + \beta_{425} \text{Agegroup}_{i425} + \beta_{426} \text{Agegroup}_{i426} + \beta_{427} \text{Agegroup}_{i427} + \beta_{428} \text{Agegroup}_{i428} + \beta_{429} \text{Agegroup}_{i429} + \beta_{430} \text{Agegroup}_{i430} + \beta_{431} \text{Agegroup}_{i431} + \beta_{432} \text{Agegroup}_{i432} + \beta_{433} \text{Agegroup}_{i433} + \beta_{434} \text{Agegroup}_{i434} + \beta_{435} \text{Agegroup}_{i435} + \beta_{436} \text{Agegroup}_{i436} + \beta_{437} \text{Agegroup}_{i437} + \beta_{438} \text{Agegroup}_{i438} + \beta_{439} \text{Agegroup}_{i439} + \beta_{440} \text{Agegroup}_{i440} + \beta_{441} \text{Agegroup}_{i441} + \beta_{442} \text{Agegroup}_{i442} + \beta_{443} \text{Agegroup}_{i443} + \beta_{444} \text{Agegroup}_{i444} + \beta_{445} \text{Agegroup}_{i445} + \beta_{446} \text{Agegroup}_{i446} + \beta_{447} \text{Agegroup}_{i447} + \beta_{448} \text{Agegroup}_{i448} + \beta_{449} \text{Agegroup}_{i449} + \beta_{450} \text{Agegroup}_{i450} + \beta_{451} \text{Agegroup}_{i451} + \beta_{452} \text{Agegroup}_{i452} + \beta_{453} \text{Agegroup}_{i453} + \beta_{454} \text{Agegroup}_{i454} + \beta_{455} \text{Agegroup}_{i455} + \beta_{456} \text{Agegroup}_{i456} + \beta_{457} \text{Agegroup}_{i457} + \beta_{458} \text{Agegroup}_{i458} + \beta_{459} \text{Agegroup}_{i459} + \beta_{460} \text{Agegroup}_{i460} + \beta_{461} \text{Agegroup}_{i461} + \beta_{462} \text{Agegroup}_{i462} + \beta_{463} \text{Agegroup}_{i463} + \beta_{464} \text{Agegroup}_{i464} + \beta_{465} \text{Agegroup}_{i465} + \beta_{466} \text{Agegroup}_{i466} + \beta_{467} \text{Agegroup}_{i467} + \beta_{468} \text{Agegroup}_{i468} + \beta_{469} \text{Agegroup}_{i469} + \beta_{470} \text{Agegroup}_{i470} + \beta_{471} \text{Agegroup}_{i471} + \beta_{472} \text{Agegroup}_{i472} + \beta_{473} \text{Agegroup}_{i473} + \beta_{474} \text{Agegroup}_{i474} + \beta_{475} \text{Agegroup}_{i475} + \beta_{476} \text{Agegroup}_{i476} + \beta_{477} \text{Agegroup}_{i477} + \beta_{478} \text{Agegroup}_{i478} + \beta_{479} \text{Agegroup}_{i479} + \beta_{480} \text{Agegroup}_{i480} + \beta_{481} \text{Agegroup}_{i481} + \beta_{482} \text{Agegroup}_{i482} + \beta_{483} \text{Agegroup}_{i483} + \beta_{484} \text{Agegroup}_{i484} + \beta_{485} \text{Agegroup}_{i485} + \beta_{486} \text{Agegroup}_{i486} + \beta_{487} \text{Agegroup}_{i487} + \beta_{488} \text{Agegroup}_{i488} + \beta_{489} \text{Agegroup}_{i489} + \beta_{490} \text{Agegroup}_{i490} + \beta_{491} \text{Agegroup}_{i491} + \beta_{492} \text{Agegroup}_{i492} + \beta_{493} \text{Agegroup}_{i493} + \beta_{494} \text{Agegroup}_{i494} + \beta_{495} \text{Agegroup}_{i495} + \beta_{496} \text{Agegroup}_{i496} + \beta_{497} \text{Agegroup}_{i497} + \beta_{498} \text{Agegroup}_{i498} + \beta_{499} \text{Agegroup}_{i499} + \beta_{500} \text{Agegroup}_{i500} + \beta_{501} \text{Agegroup}_{i501} + \beta_{502} \text{Agegroup}_{i502} + \beta_{503} \text{Agegroup}_{i503} + \beta_{504} \text{Agegroup}_{i504} + \beta_{505} \text{Agegroup}_{i505} + \beta_{506} \text{Agegroup}_{i506} + \beta_{507} \text{Agegroup}_{i507} + \beta_{508} \text{Agegroup}_{i508} + \beta_{509} \text{Agegroup}_{i509} + \beta_{510} \text{Agegroup}_{i510} + \beta_{511} \text{Agegroup}_{i511} + \beta_{512} \text{Agegroup}_{i512} + \beta_{513} \text{Agegroup}_{i513} + \beta_{514} \text{Agegroup}_{i514} + \beta_{515} \text{Agegroup}_{i515} + \beta_{516} \text{Agegroup}_{i516} + \beta_{517} \text{Agegroup}_{i517} + \beta_{518} \text{Agegroup}_{i518} + \beta_{519} \text{Agegroup}_{i519} + \beta_{520} \text{Agegroup}_{i520} + \beta_{521} \text{Agegroup}_{i521} + \beta_{522} \text{Agegroup}_{i522} + \beta_{523} \text{Agegroup}_{i523} + \beta_{524} \text{Agegroup}_{i524} + \beta_{525} \text{Agegroup}_{i525} + \beta_{526} \text{Agegroup}_{i526} + \beta_{527} \text{Agegroup}_{i527} + \beta_{528} \text{Agegroup}_{i528} + \beta_{529} \text{Agegroup}_{i529} + \beta_{530} \text{Agegroup}_{i530} + \beta_{531} \text{Agegroup}_{i531} + \beta_{532} \text{Agegroup}_{i532} + \beta_{533} \text{Agegroup}_{i533} + \beta_{534} \text{Agegroup}_{i534} + \beta_{535} \text{Agegroup}_{i535} + \beta_{536} \text{Agegroup}_{i536} + \beta_{537} \text{Agegroup}_{i537} + \beta_{538} \text{Agegroup}_{i538} + \beta_{539} \text{Agegroup}_{i539} + \beta_{540} \text{Agegroup}_{i540} + \beta_{541} \text{Agegroup}_{i541} + \beta_{542} \text{Agegroup}_{i542} + \beta_{543} \text{Agegroup}_{i543} + \beta_{544} \text{Agegroup}_{i544} + \beta_{545} \text{Agegroup}_{i545} + \beta_{546} \text{Agegroup}_{i546} + \beta_{547} \text{Agegroup}_{i547} + \beta_{548} \text{Agegroup}_{i548} + \beta_{549} \text{Agegroup}_{i549} + \beta_{550} \text{Agegroup}_{i550} + \beta_{551} \text{Agegroup}_{i551} + \beta_{552} \text{Agegroup}_{i552} + \beta_{553} \text{Agegroup}_{i553} + \beta_{554} \text{Agegroup}_{i554} + \beta_{555} \text{Agegroup}_{i555} + \beta_{556} \text{Agegroup}_{i556} + \beta_{557} \text{Agegroup}_{i557} + \beta_{558} \text{Agegroup}_{i558} + \beta_{559} \text{Agegroup}_{i559} + \beta_{560} \text{Agegroup}_{i560} + \beta_{561} \text{Agegroup}_{i561} + \beta_{562} \text{Agegroup}_{i562} + \beta_{563} \text{Agegroup}_{i563} + \beta_{564} \text{Agegroup}_{i564} + \beta_{565} \text{Agegroup}_{i565} + \beta_{566} \text{Agegroup}_{i566} + \beta_{567} \text{Agegroup}_{i567} + \beta_{568} \text{Agegroup}_{i568} + \beta_{569} \text{Agegroup}_{i569} + \beta_{570} \text{Agegroup}_{i570} + \beta_{571} \text{Agegroup}_{i571} + \beta_{572} \text{Agegroup}_{i572} + \beta_{573} \text{Agegroup}_{i573} + \beta_{574} \text{Agegroup}_{i574} + \beta_{575} \text{Agegroup}_{i575} + \beta_{576} \text{Agegroup}_{i576} + \beta_{577} \text{Agegroup}_{i577} + \beta_{578} \text{Agegroup}_{i578} + \beta_{579} \text{Agegroup}_{i579} + \beta_{580} \text{Agegroup}_{i580} + \beta_{581} \text{Agegroup}_{i581} + \beta_{582} \text{Agegroup}_{i582} + \beta_{583} \text{Agegroup}_{i583} + \beta_{584} \text{Agegroup}_{i584} + \beta_{585} \text{Agegroup}_{i585} + \beta_{586} \text{Agegroup}_{i586} + \beta_{587} \text{Agegroup}_{i587} + \beta_{588} \text{Agegroup}_{i588} + \beta_{589} \text{Agegroup}_{i589} + \beta_{590} \text{Agegroup}_{i590} + \beta_{591} \text{Agegroup}_{i591} + \beta_{592} \text{Agegroup}_{i592} + \beta_{593} \text{Agegroup}_{i593} + \beta_{594} \text{Agegroup}_{i594} + \beta_{595} \text{Agegroup}_{i595} + \beta_{596} \text{Agegroup}_{i596} + \beta_{597} \text{Agegroup}_{i597} + \beta_{598} \text{Agegroup}_{i598} + \beta_{599} \text{Agegroup}_{i599} + \beta_{600} \text{Agegroup}_{i600} + \beta_{601} \text{Agegroup}_{i601} + \beta_{602} \text{Agegroup}_{i602} + \beta_{603} \text{Agegroup}_{i603} + \beta_{604} \text{Agegroup}_{i604} + \beta_{605} \text{Agegroup}_{i605} + \beta_{606} \text{Agegroup}_{i606} + \beta_{607} \text{Agegroup}_{i607} + \beta_{608} \text{Agegroup}_{i608} + \beta_{609} \text{Age$$

18-24 acts as a reference category in the regression results.  $\mathbf{X}_i$  is a vector of covariates including the device platform, media condition (treat), gender, education, cognitive resources (CRT), measures of digital literacy, political knowledge and internet usage, if the respondent is a sexist.  $\varepsilon_i$  is the error term and  $\beta$  is a coefficient vector for the covariates. The model examining the age effect on affect is equivalent to that for measuring deception, except  $\text{Favor}_i$ , the favorability as the outcome.

In the detection experiment, the key model (Barari, Lucas, and Munger 2021a) we use to test the effect of age on detection accuracy via the specification is also similar with what we have defined, except  $\text{DetectAcc}_i$ , the ability to detect between real and fake clippings. Barari, Lucas, & Munger uses R, and our reproduction is also done in R (R Core Team 2020).

## 5 Results

1) shows the mean level of deception in each age group where the outcome is the extent of believing the clipping is fake or doctored, on a scale of 1-5. The p-values are respectively 0.392, 0.7338, 0.8289, and 0.8863 when we compare the 18-24 age group with age 25-34, age 35-44, age 45-64, and age 65+. Finally the regression results of extent of belief on age groups and covariates are presented in [ref: Table 1].

We then want to examine the favorability heterogeneity for different subgroups. From the topline null results shown in the ??), we can see that young respondents are easier to be affectively triggered by fake clippings than participants aging between 25 and 44, especially presenting a significant difference with the 35-44 age subgroup ( $\delta = -8.66$ ,  $t = -12.16$ ,  $p$  value  $< 0.05$ ). The results also indicate the same finding with what authors have investigated, showing that the elderly are more likely to be affectively triggered by the inveracious media clippings than the young people ( $\delta = 3.42$ ,  $t = 2.0196$ ,  $p$  value  $< 0.05$ ). The results regression model [Table 2] also shows the same insights.

##Detection experiment results

2) shows the overall accuracy of detecting fake videos as well as accuracy in the five age groups (18-24, 25-34, 35-44, 45-64, 65+), along with associated false positive and false negative rates in the second detection experiment. Table[4] consists of 6 regression models and shows the predictors of the second experiment detection accuracy including six age groups that involved in the detection experiment. Since our research question is to examine the differences between age groups and detection accuracy of deepfake videos, only the 45 to 65 age group is statistically significant compared to its peers.

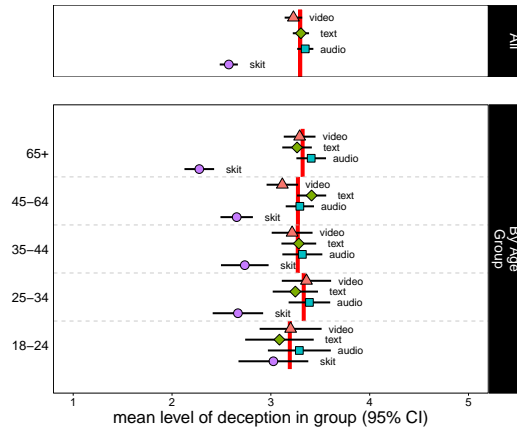


Figure 1: Marginal Means in Exposure Experiment Deception Outcomes

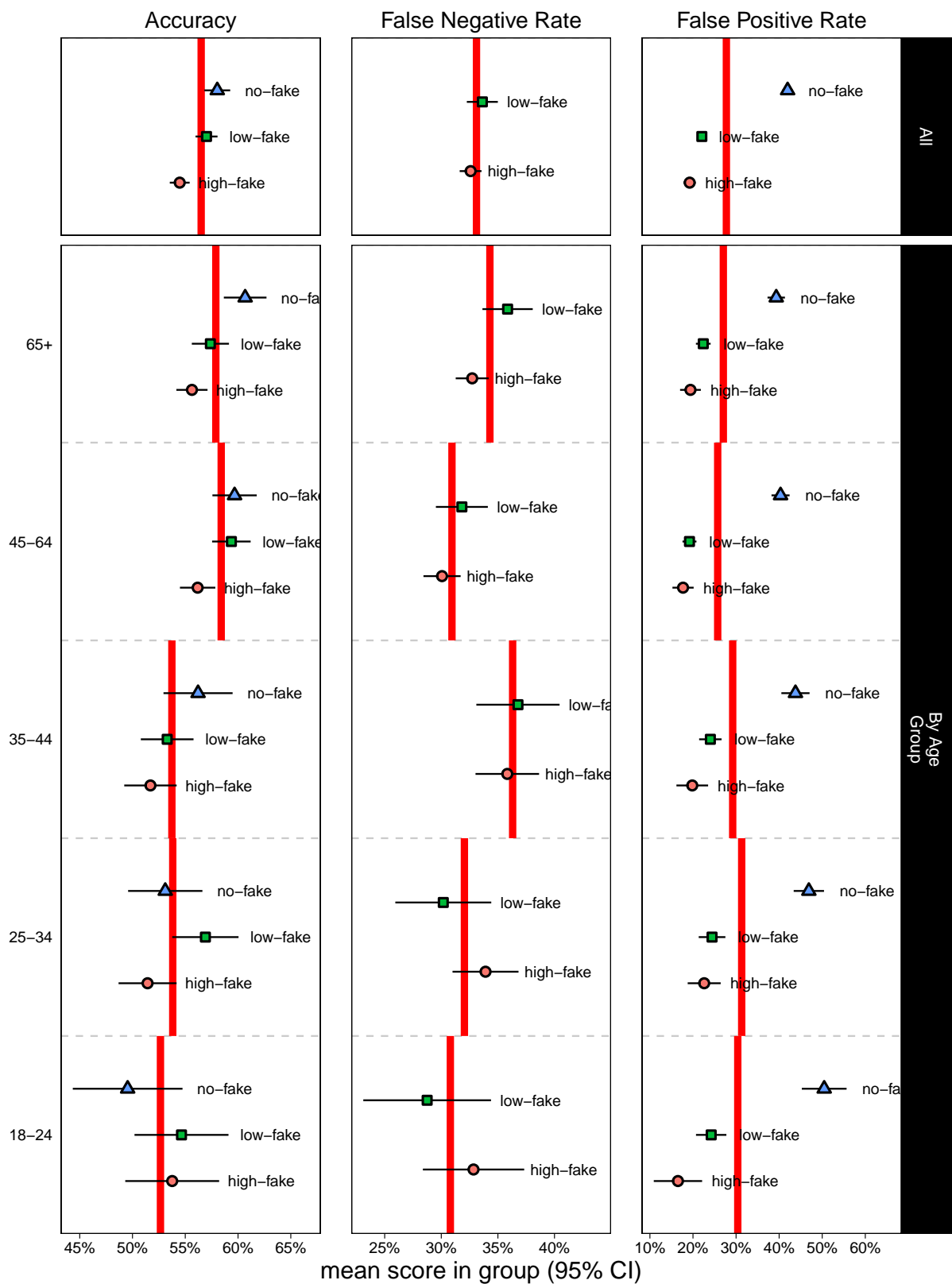


Figure 2: Marginal Means in Detection Experiment Outcomes



## 6 Discussion

### 6.1 First experiment

One interesting point from 1) is that all age groups are statistically better at detecting the fake skit compared to video, text or audio while no significant difference is found between the later three media formats. This may imply that skit is as persuasive as the video for the youngest generation in our experiment although the skit is obviously fictionalized. However, no significant differences in deception levels between the 18-24 age group with any other groups regarding the two-sample t-tests results. The regression results shown in [ref: Table 1] also agree that there are no statistically significant differences.

### 6.2 Second experiment

It is clear from the graph that the best detection accuracy is observed in the 45 to 64 age group with the highest accuracy rate of detecting low-fake and high-fake videos that are around 58% and 56% respectively. There is a statistically significant difference between the 45-64 age group and second-stage detection accuracy since the p-value for regression models are less than 0.1 and 0.05[Table 4]. Compared to the false positive rate by different age groups in 2), the 45 to 64 age group has the lowest false positive rate which indicating participants in this age group are less likely to identify real videos as fake clips than their peers. More specifically, respondents in the 35-44, 45-64 and 65+ age groups have shared the similarity of higher accuracy in detecting no-fake videos than low-fake and high-fake clips. Although Barari, Lucas and Munger have mentioned an interesting finding in their deepfake research paper that “users over 65 shared nearly 7 times as many as articles from fake news domains as the youngest age groups” on Facebook (Guess et al. 2020). Based on the result shown in 2) elder generation especially those who are over 65 years old are more likely to achieve the best detection accuracy of real videos as well as no-fake videos than the other four age groups.

By contrast, there is no statistically significant difference between the younger generations and second-stage detection accuracy. However, younger participants who are between 18 to 34 years old are much better at classifying low-fake videos correctly which leads to a greater value of accuracy in low-fake videos than no-fake and high-fake clips and correspondingly the lowest false-negative rate of detecting low-fake videos is observed. Participants who are under 24 years old achieved excellent detection accuracy in the low or high level of fake clips with approximately 55% and 54% respectively. However, younger generations seem less able to recognize real videos from a list of combinations of real and fake news videos, resulting in a high false-positive rate in 18-24 and 25-34 age groups where a larger difference between no-fake and high-fake is found especially in the 18-24 age group.

Millennials and Gen Z’ers who are often known as the youngest generation born after 1996 and is the generation who has first access to the internet technology and was largely exposed to the web revolution that occurred in the 1990s. For adolescents in the age group between 15 to 20 and adults around 21 to 30 years old, the major media consumptions for them are mainly taking place on online media and social media compared to older generations (Manalu, Pradekso, and Setyabudi 2018). Based on a study of 6000 college students in the US from 11 universities, around 89% of the respondents reported that they relied on social media to receive news (Head et al. 2018). Meanwhile, insufficient information and news feed are already overwhelmed. Social media such as Twitter, Facebook, Instagram and Tiktok are public platforms which mean everyone could post or share information without validating the sources and contents regarding ethical, truthful and high quality. Fake news on social media is something that can not be avoided entirely and is a challenge that these top tech companies make an effort to minimize the risk from this type of misinformation and disinformation and limit the spreading of harmful and misleading information since ignoring the possible effects could lead to an unintended impact on credible and validity of information shared on social media networks. Around 88% of 18 to 29 years old respondents reported using social media frequently in a 2018 survey (Smith, Anderson, and others 2018). However, people often consider Gen Z’ers are familiar with web technology and social media platforms, approximately 97% of college students reported that they have shared misinformation on social media in a 2015 study. Moreover, based on a study in 2019, colleague students are better at classifying fake news rather than real news with 60.58% accuracy of real news identification and

64.29% accuracy of fake news identification (Leeder 2019). A similar outcome is found in our research, younger generations from 18 to 24 years old are less likely to identify real video clips from a list of videos consist of both real and fake clips with less than 50% of detection accuracy. In contrast, participants from 65+ age groups performed the best identification of real videos with more than 60% of detection accuracy.

## 7 Limitation and Future works

However, one potential weakness of the survey experiments comes with the collection of data just as other web surveys. As we discussed in the data section, among all 17,501 subjects who launched the survey, only 5,750 (32.86%) subjects passed a series of checks or completed the full experiment. If there exist some similarities within those people who drop the survey experiment and these shared characteristics also relate to the experiment result, at the same time they are not captured by the demographics in the dataset (not act as controls in the experiment), then there might be a selection bias in the results. For example, if only subjects that are patient enough finish the survey experiment, while impatient subjects leave the research platform before completing it, then the data only includes patient respondents. If more patient subjects are more likely to detect the use of deepfake, plus we can hardly foresee “the level of patience” from demographics (they may be more careful and cautious), then it is unobserved here and the results will underestimate the threat of deepfakes since the randomization is not applied at the very first stage of choosing the respondents. Although web surveys are a reachable and cheap approach to gather relatively large data, self-selection (i.e., individuals select themselves to complete the survey) and under-coverage (i.e., some groups in the population are under-represented due to less access to the internet) (Bethlehem 2010). Bethlehem’s study also points that both of them threaten the weighting adjustment techniques and the representativeness of the sufficiently large dataset, which can eventually lead to biased estimates.

## 8 Appendix

### 8.1 Appendix A

Table 1: Models of Belief in Exposure-Stage (News Feed) Scandal Clipping

	Extent of belief that clipping was not fake or doctored [1-5]						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age 25-34	0.07 (0.17)	0.19 (0.15)	0.04 (0.18)	0.001 (0.17)	0.09 (0.15)	0.02 (0.18)	0.10 (0.16)
Age 35-44	-0.11 (0.16)	-0.02 (0.14)	-0.11 (0.17)	-0.13 (0.17)	-0.15 (0.15)	-0.11 (0.17)	-0.06 (0.15)
Age 45-64	-0.03 (0.15)	0.05 (0.13)	-0.02 (0.16)	-0.03 (0.16)	-0.13 (0.14)	-0.04 (0.17)	-0.02 (0.14)
Age 65+	0.01 (0.15)	0.10 (0.13)	0.02 (0.16)	-0.002 (0.16)	-0.05 (0.14)	0.02 (0.17)	0.07 (0.15)
On Mobile				0.04 (0.08)	0.07 (0.08)	0.15* (0.09)	0.25*** (0.09)
Audio				0.10 (0.09)	0.12 (0.09)	0.12 (0.09)	0.17* (0.09)
Text				-0.01 (0.09)	-0.14 (0.09)	0.02 (0.09)	-0.06 (0.09)
Skit				-0.68*** (0.10)	-0.65*** (0.10)	-0.66*** (0.11)	-0.59*** (0.10)
High School				-0.03 (0.34)	-0.25 (0.20)	-0.24 (0.34)	-0.28 (0.20)
College				-0.02 (0.33)	-0.24 (0.21)	-0.20 (0.34)	-0.21 (0.20)
Postgrad				-0.18 (0.34)	-0.34 (0.22)	-0.36 (0.34)	-0.36 (0.22)
Independent PID					0.28*** (0.10)	0.11 (0.11)	0.24** (0.11)
Republican PID					0.62*** (0.07)	0.56*** (0.08)	0.62*** (0.08)
CRT				-0.15 (0.14)	-0.15 (0.14)	-0.04 (0.15)	0.08 (0.15)
Male				-0.001 (0.07)	0.02 (0.07)	0.03 (0.07)	0.05 (0.07)
Political Knowledge				0.07 (0.15)	0.32** (0.15)	0.11 (0.16)	0.33** (0.16)
Internet Usage				0.05 (0.05)	0.08 (0.05)	0.06 (0.05)	0.07 (0.05)
Ambivalent Sexism				0.16*** (0.04)	0.03 (0.04)	0.05 (0.04)	0.06 (0.04)
Constant	3.33*** (0.14)	3.27*** (0.12)	3.34*** (0.15)	2.71*** (0.51)	2.68*** (0.43)	2.74*** (0.53)	2.38*** (0.45)
Weighted? Low-Quality Dropped?							
N	1,619	1,619	1,445	1,619	1,619	1,445	1,445
R <sup>2</sup>	0.002	0.002	0.001	0.05	0.09	0.09	0.09
Adjusted R <sup>2</sup>	-0.001	-0.0003	-0.001	0.04	0.08	0.08	0.08

\*p < .1; \*\*p < .05; \*\*\*p < .01

Notes: Reference category for medium is Video. CRT is scaled 0-1, political knowledge and ambivalent sexism are 0-1, internet usage is 1-7. Sample did not receive information in the first stage.

### 8.2 Appendix B

### 8.3 Appendix C

Table 2: Models of Scandal Target Affect

	Elizabeth Warren Feeling Thermometer						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Age 25-34	2.94 (2.27)	3.18 (1.95)	3.91 (2.46)	4.95*** (1.82)	4.58*** (1.61)	6.23*** (1.95)	6.03*** (1.72)
Age 35-44	8.66*** (2.17)	10.16*** (1.90)	10.03*** (2.34)	11.81*** (1.80)	12.51*** (1.64)	12.33*** (1.92)	12.80*** (1.75)
Age 45-64	-2.61 (2.03)	-3.84** (1.73)	-1.25 (2.18)	2.73 (1.73)	2.47 (1.52)	3.64** (1.84)	2.56 (1.61)
Age 65+	-3.42* (2.03)	-5.15*** (1.80)	-1.37 (2.17)	0.05 (1.77)	-0.32 (1.62)	1.30 (1.89)	0.07 (1.71)
Info Provided				0.74 (0.72)	0.68 (0.74)	0.69 (0.76)	0.77 (0.78)
On Mobile				-1.91* (0.98)	-1.02 (0.99)	-3.04*** (1.03)	-2.97*** (1.05)
Video				-3.42*** (1.28)	-3.98*** (1.30)	-2.65* (1.36)	-3.45** (1.38)
Audio				-3.32*** (1.25)	-4.29*** (1.28)	-2.77** (1.32)	-3.78*** (1.36)
Text				-2.60** (1.25)	-1.37 (1.28)	-2.34* (1.32)	-1.76 (1.34)
Skit				-3.33*** (1.25)	-3.53*** (1.27)	-3.00** (1.32)	-4.07*** (1.34)
Attack Ad				-4.01*** (1.26)	-3.76*** (1.26)	-3.45*** (1.33)	-3.77*** (1.33)
High School				-1.55 (3.58)	-1.37 (2.29)	-2.66 (3.70)	-2.03 (2.32)
College				0.11 (3.58)	1.27 (2.33)	-2.00 (3.70)	-0.53 (2.38)
Postgrad				8.70** (3.64)	11.81*** (2.48)	6.28* (3.78)	9.89*** (2.56)
Independent PID				-26.82*** (1.20)	-26.70*** (1.21)	-26.48*** (1.26)	-27.03*** (1.27)
Republican PID				-40.07*** (0.83)	-38.14*** (0.84)	-40.54*** (0.88)	-39.10*** (0.88)
CRT				-1.37 (1.61)	-1.54 (1.65)	-0.32 (1.71)	-0.29 (1.74)
Male				0.09 (0.82)	-1.08 (0.81)	-0.04 (0.87)	-0.74 (0.85)
Political Knowledge				1.18 (1.72)	0.31 (1.72)	1.20 (1.81)	0.44 (1.80)
Internet Usage				0.73 (0.57)	0.87 (0.58)	0.81 (0.61)	1.09* (0.63)
Ambivalent Sexism				-3.95*** (0.47)	-3.38*** (0.47)	-4.66*** (0.50)	-3.90*** (0.50)
Constant	43.41*** (1.86)	42.94*** (1.55)	41.40*** (2.01)	67.13*** (5.62)	64.50*** (4.92)	68.68*** (5.97)	65.16*** (5.25)
Weighted? Low-Quality Dropped?							
N	5,524	5,524	4,895	5,523	5,523	4,894	4,894
R <sup>2</sup>	0.02	0.03	0.02	0.39	0.36	0.40	0.38
Adjusted R <sup>2</sup>	0.02	0.03	0.01	0.39	0.36	0.40	0.38

\*p &lt; .1; \*\*p &lt; .05; \*\*\*p &lt; .01

Notes: Reference category for medium is Control.

Table 3: Predictors of Second-Stage Detection Accuracy

	Detection Accuracy (% Correctly Classified)					
	(1)	(2)	(3)	(4)	(5)	(6)
Digital Literacy		0.25*** (0.02)	0.22*** (0.02)	0.23*** (0.02)	0.20*** (0.02)	0.21*** (0.02)
Accuracy Prompt	−0.002 (0.01)		−0.01 (0.01)	−0.004 (0.01)	−0.005 (0.01)	−0.002 (0.01)
Stage 1 Debrief			0.01* (0.01)	0.01* (0.01)	0.01* (0.01)	0.01* (0.01)
Stage 1 Info Provided			−0.01 (0.01)	−0.001 (0.01)	−0.01 (0.01)	−0.004 (0.01)
Political Knowledge			0.18*** (0.01)	0.19*** (0.01)	0.18*** (0.02)	0.19*** (0.02)
Internet Usage			−0.001 (0.005)	−0.01 (0.005)	0.002 (0.01)	−0.001 (0.01)
Low-fake Env.			0.03*** (0.01)	0.04*** (0.01)	0.03*** (0.01)	0.05*** (0.01)
No-fake Env.			0.04*** (0.01)	0.04*** (0.01)	0.04*** (0.01)	0.05*** (0.01)
Age groups 25-34			−0.01 (0.02)	0.003 (0.01)	−0.004 (0.02)	0.01 (0.01)
Age groups 35-44			−0.02 (0.02)	−0.01 (0.01)	−0.02 (0.02)	−0.01 (0.01)
Age groups 45-64			0.01 (0.01)	0.02* (0.01)	0.02 (0.02)	0.03** (0.01)
Age groups 65+			0.004 (0.01)	0.01 (0.01)	0.01 (0.02)	0.01 (0.01)
High School			0.01 (0.03)	0.01 (0.02)	0.03 (0.03)	0.02 (0.02)
College			0.02 (0.03)	0.02 (0.02)	0.04 (0.03)	0.03 (0.02)
Postgrad			−0.005 (0.03)	−0.01 (0.02)	0.01 (0.03)	0.001 (0.02)
Republican			0.06*** (0.02)	0.08*** (0.02)	0.07*** (0.02)	0.08*** (0.02)
CRT			−0.06** (0.03)	−0.06** (0.03)	−0.06** (0.03)	−0.06** (0.03)
Republican x CRT			0.09*** (0.01)	0.07*** (0.01)	0.09*** (0.01)	0.08*** (0.01)
Ambivalent Sexism			0.001 (0.004)	−0.002 (0.004)	0.002 (0.004)	0.0004 (0.004)
Constant	0.57*** (0.005)	0.36*** (0.02)	0.16*** (0.05)	0.18*** (0.04)	0.13** (0.05)	0.14*** (0.05)
Weighted?						
Low-Quality Dropped?						
N	5,497	5,497	5,496	5,496	4,870	4,870
R <sup>2</sup>	0.0000	0.02	0.09	0.10	0.09	0.10
Adjusted R <sup>2</sup>	−0.0002	0.02	0.09	0.09	0.09	0.10

\*p &lt; .1; \*\*p &lt; .05; \*\*\*p &lt; .01

Notes: Reference category for environment is High-fake. Reference category for age group is 18-24. PID pooled for brevity.

## 8.4 Appendix D

Table 4: Predictors of Second-Stage False Positive Rate (FPR)

	Detection FPR (% Real Videos Classified as Deepfakes)					
	(1)	(2)	(3)	(4)	(5)	(6)
Digital Literacy		−0.11*** (0.03)	−0.09*** (0.03)	−0.12*** (0.03)	−0.06** (0.03)	−0.09*** (0.03)
Accuracy Prompt	−0.01 (0.01)		0.004 (0.01)	−0.01 (0.01)	0.003 (0.01)	−0.01 (0.01)
Stage 1 Debrief			−0.03*** (0.01)	−0.03*** (0.01)	−0.03*** (0.01)	−0.03*** (0.01)
Stage 1 Info Provided			0.01 (0.01)	−0.0003 (0.01)	0.01 (0.01)	0.004 (0.01)
Political Knowledge			−0.12*** (0.02)	−0.11*** (0.02)	−0.12*** (0.02)	−0.12*** (0.02)
Internet Usage			−0.004 (0.01)	−0.002 (0.01)	−0.004 (0.01)	−0.002 (0.01)
Low-fake Env.			0.03*** (0.01)	0.02*** (0.01)	0.02*** (0.01)	0.01 (0.01)
No-fake Env.			0.23*** (0.01)	0.23*** (0.01)	0.22*** (0.01)	0.21*** (0.01)
Age groups 25-34			0.02 (0.02)	−0.001 (0.02)	0.01 (0.02)	−0.01 (0.02)
Age groups 35-44			0.01 (0.02)	−0.03* (0.02)	0.005 (0.02)	−0.04** (0.02)
Age groups 45-64			−0.02 (0.02)	−0.04*** (0.01)	−0.02 (0.02)	−0.05*** (0.02)
Age groups 65+			−0.0001 (0.02)	−0.03* (0.02)	0.002 (0.02)	−0.04** (0.02)
High School			0.002 (0.03)	−0.001 (0.02)	−0.02 (0.04)	−0.02 (0.02)
College			0.01 (0.03)	0.01 (0.02)	−0.02 (0.04)	−0.01 (0.02)
Postgrad			0.03 (0.04)	0.02 (0.02)	−0.001 (0.04)	0.01 (0.02)
Republican			−0.07*** (0.02)	−0.09*** (0.02)	−0.07*** (0.02)	−0.08*** (0.02)
CRT			0.03 (0.03)	0.04 (0.03)	0.04 (0.03)	0.03 (0.03)
Republican x CRT			−0.07*** (0.01)	−0.06*** (0.01)	−0.08*** (0.01)	−0.07*** (0.01)
Ambivalent Sexism			−0.0002 (0.005)	0.002 (0.005)	−0.002 (0.005)	0.001 (0.005)
Constant	0.28*** (0.01)	0.37*** (0.02)	0.43*** (0.06)	0.48*** (0.05)	0.44*** (0.06)	0.49*** (0.05)
Weighted?						
Low-Quality Dropped?						
N	5,495	5,495	5,494	5,494	4,869	4,869
R <sup>2</sup>	0.0002	0.003	0.16	0.17	0.16	0.16
Adjusted R <sup>2</sup>	−0.0000	0.003	0.16	0.16	0.15	0.16

\*p < .1; \*\*p < .05; \*\*\*p < .01

Notes: Reference category for environment is High-fake. Reference category for age group is 18-24. PID pooled for brevity.

## Reference

Barari, Soubhik, Christopher Lucas, and Kevin Munger. 2021a. “Political Deepfake Videos Misinform the Public, but No More Than Other Fake Media.”

Table 5: Predictors of Second-Stage False Negative Rate (FNR)

	Detection FNR (% Deepfakes Classified as Real Videos)					
	(1)	(2)	(3)	(4)	(5)	(6)
Digital Literacy		-0.03 (0.03)	-0.03 (0.03)	-0.01 (0.03)	-0.05 (0.03)	-0.02 (0.04)
Accuracy Prompt	-0.01 (0.01)		-0.004 (0.01)	-0.01 (0.01)	0.003 (0.01)	-0.002 (0.01)
Stage 1 Debrief			0.01 (0.01)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
Stage 1 Info Provided			-0.002 (0.01)	0.003 (0.01)	-0.01 (0.01)	-0.002 (0.01)
Political Knowledge			-0.002 (0.02)	-0.03 (0.02)	-0.01 (0.02)	-0.03 (0.02)
Internet Usage			0.02*** (0.01)	0.01 (0.01)	0.02*** (0.01)	0.02** (0.01)
Low-fake Env.			0.01 (0.01)	0.0002 (0.01)	0.01 (0.01)	-0.005 (0.01)
No-fake Env.			-0.001 (0.02)	0.02 (0.02)	-0.01 (0.02)	0.01 (0.02)
Age groups 25-34			0.03 (0.02)	0.05** (0.02)	0.02 (0.02)	0.04* (0.02)
Age groups 35-44			-0.01 (0.02)	0.003 (0.02)	-0.01 (0.02)	0.003 (0.02)
Age groups 45-64			0.03 (0.02)	0.04** (0.02)	0.02 (0.02)	0.04** (0.02)
Age groups 65+			0.01 (0.04)	0.02 (0.03)	0.01 (0.04)	0.01 (0.03)
High School			0.03 (0.04)	0.04 (0.03)	0.03 (0.04)	0.03 (0.03)
College			0.08* (0.04)	0.11*** (0.03)	0.07 (0.05)	0.10*** (0.03)
Postgrad			-0.05** (0.02)	-0.04 (0.03)	-0.06** (0.03)	-0.04 (0.03)
Republican			0.09** (0.04)	0.08* (0.04)	0.11*** (0.04)	0.10** (0.04)
CRT			0.02*** (0.01)	0.02*** (0.01)	0.01* (0.01)	0.01 (0.01)
Republican x CRT			-0.04*** (0.01)	-0.03** (0.02)	-0.05*** (0.02)	-0.04** (0.02)
Ambivalent Sexism	0.34*** (0.01)	0.36*** (0.03)	0.15** (0.07)	0.19*** (0.06)	0.18** (0.07)	0.22*** (0.07)
Weighted?						
Low-Quality Dropped?						
N	3,690	3,690	3,690	3,690	3,266	3,266
R <sup>2</sup>	0.0002	0.0003	0.02	0.03	0.02	0.02
Adjusted R <sup>2</sup>	-0.0000	0.0000	0.02	0.03	0.01	0.02

\*p &lt; .1; \*\*p &lt; .05; \*\*\*p &lt; .01

Notes: Reference category for environment is High-fake. Reference category for age group is 18-24. PID pooled for brevity.



- . 2021b. “Political Deepfake Videos Misinform the Public, but No More Than Other Fake Media.” <https://github.com/soubhikbarari/Political-Deepfakes-Fx>.
- Barnwell, Paul. 2020. “Are Deepfake Videos a Threat to Democracy?” <https://www.commonsense.org/education/articles/are-deepfake-videos-a-threat-to-democracy>.
- Bethlehem, Jelke. 2010. “Selection Bias in Web Surveys.” *International Statistical Review* 78 (2): 161–88.
- Bio. 2019. “The Bizarre and Terrifying Case of the ‘Deepfake’ Video That Helped Bring an African Nation to the Brink.” <https://www.motherjones.com/politics/2019/03/deepfake-gabon-ali-bongo/>.
- Guess, Andrew M, Michael Lerner, Benjamin Lyons, Jacob M Montgomery, Brendan Nyhan, Jason Reifler, and Neelanjana Sircar. 2020. “A Digital Media Literacy Intervention Increases Discernment Between Mainstream and False News in the United States and India.” *Proceedings of the National Academy of Sciences* 117 (27): 15536–45.
- Head, Alison, John Wihbey, P Takis Metaxas, Margy MacMillan, and Dan Cohen. 2018. “How Students Engage with News: Five Takeaways for Educators, Journalists, and Librarians.” *Project Information Literacy Research Institute*.
- Leeder, Chris. 2019. “How College Students Evaluate and Share ‘Fake News’ Stories.” *Library & Information Science Research* 41 (3): 100967.
- Manalu, Rouli, Tandiyo Pradekso, and Djoko Setyabudi. 2018. “Understanding the Tendency of Media Users to Consume Fake News.”
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rini, Regina. 2020. “Deepfakes and the Epistemic Backstop.” *Philosopher’s Imprint* 20.
- Smith, Aaron, Monica Anderson, and others. 2018. “Social Media Use in 2018.” *Pew Research Center* 1: 1–4.