S. HRG. 118–130

AI AND THE FUTURE OF OUR ELECTIONS

HEARING

BEFORE THE

COMMITTEE ON RULES AND ADMINISTRATION UNITED STATES SENATE

ONE HUNDRED EIGHTEENTH CONGRESS

FIRST SESSION

WEDNESDAY, SEPTEMBER 27, 2023

Printed for the use of the Committee on Rules and Administration

Available on http://www.govinfo.gov

U.S. GOVERNMENT PUBLISHING OFFICE
53–678 WASHINGTON : 2023

COMMITTEE ON RULES AND ADMINISTRATION

FIRST SESSION

AMY KLOBUCHAR, Minnesota, *Chairwoman*

DIANNE FEINSTEIN, California	DEB FISCHER, Nebraska
CHARLES E. SCHUMER, New York	MITCH MCCONNELL, Kentucky
MARK R. WARNER, Virginia	TED CRUZ, Texas
JEFF MERKLEY, Oregon	SHELLEY MOORE CAPITO, West Virginia
ALEX PADILLA, California	ROGER WICKER, Mississippi
JON OSSOFF, Georgia	CINDY HYDE-SMITH, Mississippi
MICHAEL F. BENNET, Colorado	BILL HAGERTY, Tennessee
PETER WELCH, Vermont	KATIE BOYD BRITT, Alabama

ELIZABETH FARRAR, *Staff Director*
JACKIE BARBER, *Republican Staff Director*

C O N T E N T S

AI AND THE FUTURE OF OUR ELECTIONS

WEDNESDAY, SEPTEMBER 27, 2023

UNITED STATES SENATE
COMMITTEE ON RULES AND ADMINISTRATION
Washington, DC.

The Committee met, pursuant to notice, at 3:31 p.m., in Room 301, Russell Senate Office Building, Hon. Amy Klobuchar, Chairwoman of the Committee, presiding.

Present: Senators Klobuchar, Fischer, Schumer, Warner, Merkley, Padilla, Ossoff, Bennet, Welch, Hagerty, and Britt.

OPENING STATEMENT OF HONORABLE AMY KLOBUCHAR, CHAIRWOMAN, A UNITED STATES SENATOR FROM THE STATE OF MINNESOTA

Chairwoman KLOBUCHAR. Okay. Good afternoon, everyone. I am honored to call this hearing to order. I am pleased to be here with my colleague, Senator Fischer, wearing her pin with the ruby red slippers, which symbolizes there is no place like home.

Senator FISCHER. On top of my heels.

Chairwoman KLOBUCHAR. Yes, this week in Washington, it is kind of on our minds. Thank you as well, Senator Merkley, for being here. I know we have other Members attending as well. I want to thank Ranking Member Fischer and her staff for working with us on this hearing on Artificial Intelligence and the Future of our Elections.

I want to introduce—I will introduce our witnesses shortly, but we are joined by Minnesota's Secretary of State, Steve Simon, with vast experience running elections and is well respected in our state and nationally.

Trevor Potter, the President of the Campaign Legal Center, and former FEC Commissioner and Chair. Thank you for being here. Maya Wiley, President and CEO of the Leadership Conference on Civil and Human Rights. We are also going to hear, I know that Ranking Member Fischer will be introducing our two remaining witnesses. We thank you for being here, Neil Chilson, Senior Research Fellow at the Center for Growth and Opportunity, and Ari Cohn, Free Speech Counsel at TechFreedom.

Like any emerging technology, AI comes with significant risks, and our laws need to keep up. Some of the risks are already clear, starting with security, which includes protecting our critical infrastructure, guarding against cyber-attacks, and staying ahead of foreign adversaries. We must also protect our innovation economy, including the people who produce content, and countering the alarming rise in criminals using AI to scam people.

Confronting these issues is a major bipartisan focus here in the Senate, where two weeks ago we convened the first in a series of forums organized by Leader Schumer, and Senators Rounds and Young, and Senator Heinrich to discuss this technology with experts of all backgrounds, industry, union, nonprofit, across the spectrum in their views.

Today, we are here to focus, hone in on a particular risk of AI. That is the risk that it poses for our elections and how we address them. Given the stakes for our democracy, we cannot afford to wait. The hope is we can move on some of this by year end with some of the legislation which already has bipartisan support, to be able to get it done with some larger legislation.

As I noted, we are already seeing this technology being used to generate viral, misleading content, to spread disinformation, and deceive voters. There was an AI-generated video, for instance, posted on Twitter of one of my colleagues, Senator Warren, in which a fake Senator Warren said that people from the opposing party should not be able to vote. She never said that, but it looked like her.

The video was seen by nearly 200,000 users in a week, and AI-generated content has already begun to appear in political ads. There was one AI-generated image of former President Trump hugging Dr. Fauci that was actually a fake.

The problem for voters is that people are not going to be able to distinguish if it is the opposing candidate or their own candidate, if it is them talking or not. That is untenable in a democracy. Plus, new services like Banter AI have hit the market, which can create voice recordings that sound like, say, President Biden or other elected officials from either party.

This means that anyone with a computer can put words in the mouth of a leader. That would pose a problem during an emergency situation like a natural disaster, and it is not hard to imagine it being used to confuse people. We also must remember that the risks posed by AI are not just about candidates. It is also about people being able to vote. In the Judiciary hearing, I actually just simply asked ChatGPT to write me a tweet about a polling location in Bloomington, Minnesota. I noted that sometimes there were lines at that location, what should voters do? It just quickly spit out, go to 1234 Elm Street. There is no such location in Bloomington, Minnesota.

You have the problem of that too, more likely to occur as we get closer to an election. With AI, the rampant disinformation we have seen in recent years will quickly grow in quantity and quality.

We need guardrails to protect our elections. What do we do? I hope that will be some of the subject, in addition to admiring the problem that we can discuss today. Senator Hawley and I worked over the last two months on a bill together that we are leading together—hold your beer, that is correct. On a bill that we are leading together to get at deepfake videos, like the ones I just talked about used against former President Trump, and against Elizabeth Warren. Those are ads that are not really the people. Senator Collins and Senator Coons, Senator Bennet, Senator Ricketts have joined us already on that bill.

We just introduced it. It creates a framework that is Constitutionally all right, based on past and recent precedent, with exceptions for things like parody and satire, that allows those to be banned. Another key part of transparency when it comes to this technology is disclaimers for other types of ads.

That is another bill, Congresswoman Yvette Clarke is leading it in the House, which would require a disclaimer on ads that include AI-generated images so at least voters know that AI is being used in the campaign ads.

Finally, I see Commissioner Dickerson out there. Finally—are you happy about that, Mr. Cohn? There you go. Finally, it is important that the Federal Election Commission be doing their part in taking on these threats.

While the FEC is now accepting public comments on whether it can regulate the deceptive AI-generated campaign ads after deadlocking on the issue earlier this summer, we must remain focused on taking action in time for the next election. Whether you agree or not, agree that the FEC currently has the power to do that, there is nothing wrong with spelling it out if that is the barrier.

We are working with Republicans on that issue as well. I kind of look at it three-pronged. The most egregious that must be banned under the—with the Constitutional limitations, the disclaimers, and then giving the FEC the power that they need, as well as a host of state laws, one of which I am sure we will hear about from Steve Simon.

With bipartisan cooperation put in place, and we will get the guardrails that we need. We can harness the potential of AI, the great opportunities, while controlling the threats we now see emerging and safeguard our democracy from those who would use this technology to spread disinformation and upend our elections, whether it is abroad, whether it is domestic.

I believe strongly in the power of elections. I also believe in innovation, and we have got to be able to draw that line to allow voters to vote and make good decisions, while at least putting the guardrails in place. With that, I turn it over to my friend, Senator Fischer. Thank you.

OPENING STATEMENT OF HONORABLE DEB FISCHER, A UNITED STATES SENATOR FROM THE STATE OF NEBRASKA

Senator FISCHER. Thank you, Chairwoman Klobuchar. Thank you to our witnesses today for being here. I do look forward to hearing your testimony.

Congress often examines issues that affect Americans on a daily basis. Artificial intelligence has become one of those issues. AI is not new, but significant increases in computing power have revolutionized its capabilities. It has quickly moved from the stuff of science fiction to being a part of our daily lives.

There is no question that AI is transformative and is poised to evolve rapidly. This makes understanding AI all the more important. In considering whether legislation is necessary, Congress should weigh the benefits and the risks of AI.

We should look at how innovative uses of AI could improve the lives of our constituents, and also the dangers that AI could pose.

We should consider the possible economic advantages and pitfalls. We should thoughtfully examine existing laws and regulations, and how they might apply to AI.

Lately, AI has been a hot topic here in Washington. I know many of my colleagues and Committees in both chambers are exploring this issue. The Rules Committee's jurisdiction includes federal laws governing elections and campaign finance, and we are here today to talk about how AI impacts campaign, politics, and elections.

The issues surrounding the use of AI in campaigns and elections are complicated. On one hand, there are concerns about the use of AI to create deceptive or fraudulent campaign ads. On the other hand, AI can allow campaigns to more efficiently and effectively reach voters. AI driven technology can also be used to check images, video, and audio for authenticity.

As we learn more about this technology, we must also keep in mind the important protections our Constitution provides for free speech in this country. Those protections are vital to preserving our democracy.

For a long time, we did not have many reasons to consider the sources of speech, or if it mattered whether AI was helping to craft it. Our First Amendment prohibits the Government from policing protected speech, so we must carefully scrutinize any policy proposals that would restrict that speech.

As Congress examines this issue, we need to strike a careful balance between protecting the public, protecting innovation, and protecting speech. Well-intentioned regulations rushed into law can stifle both innovation and our Constitutional responsibilities.

Again, I am grateful that we have the opportunity to discuss these issues today and to hear from our expert witnesses. Thank you.

Chairwoman KLOBUCHAR. Thank you very much, Senator Fischer. I am going to introduce our witnesses. Our first witness is Minnesota Secretary of State Steve Simon. Secretary Simon has served as Minnesota's Chief Elections Administrator since 2015.

He previously served in the Minnesota House of Representatives and was an Assistant Attorney General. He earned his law degree from the University of Minnesota and his bachelor's degree from Tufts.

Our second witness is Trevor Potter, President of the Campaign Legal Center, which he founded in 2002, and former Republican Chairman of the Federal Election Commission, after his last appointment by President H.W. Bush.

He appeared before this Committee last in March of 2021 and did not screw up, so we invited him back again. Mr. Potter also served as General Counsel to my friend and former colleague, Senator John McCain's 2000 and 2008 Presidential campaign, and has taught campaign finance at the University of Virginia and at Oxford. He earned his law degree from The University of Virginia, and bachelor's degree from Harvard.

Our third witness is Maya Wiley, President and CEO of The Leadership Conference on Civil and Human Rights. Ms. Wiley is also a Professor of Public and Urban Policy at The New School. Previously, she served as Counsel to the Mayor of New York City and was the Founder and President of the Center for Social Inclu-

sion. She earned her law degree from Columbia Law School and her bachelor's degree from Dartmouth. With that, I will have Senator Fischer introduce our remaining two witnesses.

Senator FISCHER. Thank you, Senator Klobuchar. Again, I thank our witnesses for all being here today. We have with us also Neil Chilson, who serves as a Senior Research Fellow at the Center for Growth and Opportunity, a nonpartisan think tank at Utah State University that focuses on technology and innovation.

Mr. Chilson has previously served as Acting Chief Technologist at the Federal Trade Commission.

We also have Ari Cohn, who serves as free speech Counsel at TechFreedom, a nonpartisan nonprofit devoted to technology, law, and policy, and the preservation of civil liberties. Mr. Cohn is a nationally recognized expert in First Amendment law and defamation law, and co-authored amicus briefs to state and federal courts across the country on vital First Amendment issues. Welcome to all of you.

Chairwoman KLOBUCHAR. Very good. If the witnesses could please stand.

Chairwoman KLOBUCHAR. Do you swear the testimony you are going to give before the Committee shall be the truth, the whole truth, and nothing but the truth, so help you God?

Mr. SIMON. I do.

Mr. POTTER. I do.

Ms. WILEY. I do.

Mr. CHILSON. I do.

Mr. COHN. I do.

Chairwoman KLOBUCHAR. Thank you. Please be seated. We are going to proceed.

OPENING STATEMENT OF HONORABLE STEVE SIMON, SECRETARY OF STATE, STATE OF MINNESOTA, ST. PAUL, MINNESOTA

Mr. SIMON. Thank you, Madam Chair, Ranking Member Fischer, and Members of the Committee. Thank you for this opportunity. I am Steve Simon. I have the privilege of serving as Minnesota's Secretary of State. I am grateful for your willingness to engage on this important topic, and I really am honored to be here.

Artificial intelligence is not a threat to American democracy in and of itself, but it is an emerging and powerful amplifier of existing threats. All of us who touch the election process must be watchful and proactive, especially as the 2024 Presidential contest approaches.

A year ago, we were not talking so much about generative AI. The release of the newly accessible tools such as ChatGPT challenged all that. In the hands of those who want to mislead, AI is a new and improved tool. Instead of stilted communications with poor grammar, generative AI can provide apparent precision and clarity. The potential threat to the administration of elections is real.

We are talking about an old problem, namely election misinformation and disinformation, that can now more easily be amplified. One possible danger could come from an innocent cir-

cumstance. AI software simply might fail to grasp the nuances of our state by state election system.

A prominent computer scientist in Minnesota named Max Hailperin made this point in an article several months ago. He asked ChatGPT questions about Minnesota election law, much as Senator Klobuchar said that she did, and the program gave the wrong answers to several questions. Now, was that intentional misdirection? Probably not. Still, it is a danger to voters who may get bad information about critical election rules.

In the wrong hands, AI could be used to misdirect intentionally and in ways that are far more advanced than ever. I remember seeing a paper leaflet from an election about 20 or more years ago, distributed in a particular neighborhood, that told residents that in the coming election voting would occur on Tuesday for those whose last names begin with the letters A through L, while everyone else would vote on Wednesday.

Now, that was a paper leaflet from a couple or more decades ago. Now imagine a convincing seeming email or deepfake conveying that kind of disinformation in 2024. The perpetrators could be domestic or foreign. In fact, the Department of Homeland Security has warned recently that our foreign adversaries may use AI to sharpen their attacks on our democracy.

One last point on potential consequences. The Brennan Center recently identified a so-called liar's dividend from the very use of AI. Simply put, the mere existence of AI can lead to undeserved suspicion of messages that are actually true. A video, for example, that contradicts a person's preconceived ideas may now be simply dismissed as a deepfake. The bottom line is that misdirection in elections can cause disruption.

If AI misdirects, it could become an instrument of that disruption. What can be done about it? Well, in our office, we are trying to be proactive. First, we are leading with the truth. That means pushing out reliable and accurate information while also standing up to mis and disinformation quickly.

Second, we have been working with local and federal partners to monitor and respond to inaccuracies that could morph into conspiracy theories on election related topics.

Third, we have emphasized media literacy. The National Association of Secretaries of State has helped with its Trusted Sources Initiative, urging Americans to seek out sources of election information from Secretaries of State and local election administrators.

Fourth, our cyber defenses are strong. We have invested time and resources in guarding against intrusions that could introduce misleading information to voters.

As for possible legislation, I do believe that a federal approach would be helpful. The impact of AI will be felt at a national level. I applaud bipartisan efforts such as the Protect Elections from Deceptive AI Act and the Real Political Ads Act.

Recently, the Minnesota Legislature enacted similar legislation with broad bipartisan support. There is a critical role for the private sector, too. Companies have a responsibility to the public to make sure their AI products are secure and trustworthy. I support the efforts already underway to encourage adherence to basic standards. But let me end on a note of some cautious optimism.

AI is definitely a challenge. It is a big challenge. But in some ways, we have confronted similar challenges before with each technological leap. We have generally been able to manage the potential disruptions to the way we receive and respond to information.

The move to computerization, the arrival of the internet, the emergence of social media all threatened to destabilize information pathways. But in short order, the American people got smart about those things.

They adapted, and Congress helped. AI may be qualitatively different from those other advances, but if we get better at identifying false information and if we continue to rely on trusted sources for election information, and if Congress can help, we can overcome many of the threats that AI poses, while harnessing its benefits to efficiency and productivity.

Thank you for inviting me to testify today. I look forward to our continued partnership.

[The prepared statement of Mr. Simon was submitted for the record.]

Chairwoman KLOBUCHAR. Thank you very much. Appreciate it. Mr. Potter.

OPENING STATEMENT OF HONORABLE TREVOR POTTER, FORMER COMMISSIONER AND CHAIRMAN OF THE FEDERAL ELECTION COMMISSION, FOUNDER AND PRESIDENT, CAMPAIGN LEGAL CENTER, WASHINGTON, DC

Mr. POTTER. Good afternoon and thank you for the honor of appearing before you today to testify about artificial intelligence and elections. My testimony will focus on how political communications generated through AI relate to the conduct of campaigns and why federal regulation is urgently needed to address the impact of some aspects of this technology on our democracy.

To summarize the overarching concern, AI tools can increasingly be used to design and spread fraudulent or deceptive political communications that infringe on voters' fundamental right to make informed decisions at the ballot box.

Every election cycle, billions of dollars are spent to create and distribute political communications. Before voters cast their ballots they must parse through these many messages and decide what to believe. Our campaign laws are intended to protect and assist voters by requiring transparency about who is paying to influence their election choices and who is speaking to them.

However, AI could make voters' task much more difficult because of its unprecedented ability to easily create realistic false content. Unchecked, the deceptive use of AI could make it virtually impossible to determine who is truly speaking in a political communication, whether the message being communicated is authentic, or even whether something being depicted actually happened.

This could leave voters unable to meaningfully evaluate candidates, and candidates unable to convey their desired message to voters, undermining our democracy. It opens the door to malign—even foreign—actors to manipulate our elections with false information. Foreign adversaries may not favor specific candidates, they may just seek to create chaos and sow distrust in our elections, thereby harming both parties and the whole country.

I believe there are three concurrent paths to proactively addressing these risks, three paths flagged by the Chair in her opening remarks.

First, Congress could strengthen the FEC's power to protect elections against fraud. Under current, existing law, the FEC can stop federal candidates and their campaigns from fraudulently misrepresenting themselves as speaking for another candidate or party on a matter which is damaging to that candidate or party.

I believe the FEC should explicitly clarify, through the rulemaking process, that the use of AI is included in this prohibition. Then Congress should expand this provision to prohibit any person, not just a candidate, from fraudulently misrepresenting themselves as speaking for a candidate.

Second, Congress should pass a new law specifically prohibiting the use of AI to engage in electoral fraud or manipulation. This would help protect voters from the most pernicious uses of AI. While any regulation of campaign speech raises First Amendment concerns that must be addressed, let me also say this, the Government has a clear, compelling interest in protecting the integrity of the electoral process.

In addition, voters have a well-recognized First Amendment right to meaningfully participate in elections, including being able to assess the political messages they see and know who the actual speaker is. There is no countervailing First Amendment right to intentionally defraud voters in elections, so a narrow law prohibiting the use of AI to deceptively undermine our elections through fake speech would rest on firm Constitutional footing.

Third, and finally, Congress should also expand existing disclosure requirements to ensure voters know when electoral content has been materially altered or falsified by AI. This would at least ensure voters can treat such content with appropriate skepticism.

These proposals are not mutually exclusive or exhaustive. Congress could decide to use a combination of tools, while a single solution is unlikely to remain relevant for long. Congress should carefully consider how each policy could be most effectively enforced, with options including overhauling the often gridlocked and slow FEC enforcement process, new criminal penalties enforceable by the Justice Department, and a private right of action, allowing candidates targeted by deceptive AI to seek rapid relief in federal court.

Thank you for the opportunity to testify today. I look forward to your questions.

[The prepared statement of Mr. Potter was submitted for the record.]

Chairwoman KLOBUCHAR. Thank you very much, Mr. Potter. The Rules Committee, as Senator Fischer knows, is the only Committee on which both Senator Schumer and Senator McConnell serve.

This makes our jobs very important. We are pleased that Senator Schumer is here, and we are going to give him the opportunity to say a few words. Thank you.

Senator SCHUMER. Well, thank you, Senator Klobuchar. Whatever Committee you Chair will always be important. Same with Senator Fischer. I would like to congratulate you, Mr. Potter. You made it as a witness without being from Minnesota.

[Laughter.]

Senator SCHUMER. Anyway, thank you. I want to thank my colleagues for being here. As you all know, AI, artificial intelligence is already reshaping life on earth in dramatic ways. It is transforming how we fight diseases, tackle hunger, manage our lives, enrich our minds, ensure peace, and very much more.

But we cannot ignore AI's dangers, workforce disruptions, misinformation, bias, new weapons. Today, I am pleased to talk to you about a more immediate problem, how AI could be used to jaundice, even totally discredit, our elections as early as next year. Make no mistake, the risks of AI on our elections is not just an issue for Democrats, nor just Republicans. Every one of us will be impacted. No voter will be spared.

No election will be unaffected. It will spread to all corners of democracy, and thus it demands a response from all of us. That is why I firmly believe that any effort by Congress to address AI must be bipartisan, and I can think of few issues that should both—unite both parties faster than safeguarding our democracy.

We do not need to look very hard to see how AI can warp our democratic systems this year. We have already seen instances of AI-generated deepfakes and misinformation reach the voters. Political ads have been released this year, right now, using AI-generated images and text to voice converters to depict certain candidates in a negative light.

Uncensored chat bots can already be deployed at a massive scale to target millions of individual voters for political persuasion. Once damaging information is sent to 100 million homes, it is hard, oftentimes impossible, to put that genie back in the bottle.

Everyone has experienced these rampant rumors that once they get out there, no matter how many times you refute them, still stick around. If we do not act, we could soon live in a world where political campaigns regularly deploy totally fabricated but also totally believable images and footage of Democratic or Republican candidates, distorting their statements and greatly harming their election chances.

What then is to stop foreign adversaries from taking advantage of this technology to interfere with our elections? This is the problem we now face. If left unchecked, AI's use in our elections could erode our democracy from within and from abroad, and the damage, unfortunately, could be irreversible.

As Americans prepare to go to the polls in 2024, we have to move quickly to establish safeguards to protect voters from AI related misinformation. It will not be easy. For Congress to legislate on AI is for us to engage in perhaps the most complex subject this body has ever faced. I am proud of the Rules Committee and the leadership on this issue.

Thank you, Chairwoman Klobuchar, for your continuing work on important legislative efforts to protect our elections from the potential harms on AI. Thank you again for organizing this hearing. Holding this hearing on AI and our elections is essential for drawing attention to the need for action, and I commend you and Ranking Member Fischer for doing just that.

In the meantime, I will continue working with Senators Rounds, Heinrich, and Young to host AI inside forums that focus on issues

like AI and democracy, to supplement the work of the Rules Committee and our other Committees, and I look forward to working with both Senators Klobuchar and Fischer, and all of the Rules Committee Members.

Thank you for being here, to Senators Welch, and Merkley, and Britt to develop bipartisan legislation that maximizes AI's benefits, and minimizes the risks.

Finally, the responsibility for protection—protecting our elections will not be Congress's alone. The Administration should continue leveraging the tools we have already provided them, and private companies must do their part to issue their own safeguards for how AI systems are used in the political arena.

It will take all of us, the Administration, the private sector, Congress working together to protect our democracy, ensure robust transparency and safeguards, and ultimately keep the vision of our founders alive in the 21st century.

Thank you again to the Members of this Committee. Thank you to Chairwoman Klobuchar, Ranking Member Fischer, for convening the hearing. I look forward to working with all of you on comprehensive AI legislation and learning from your ongoing work. Thank you.

Chairwoman KLOBUCHAR. Thank you very much, Senator Schumer. I will note it was this Committee, with your and Senator McConnell's support, that was able to pass the electoral reform bill, Electoral Count Reform Act, with near unanimous support and got it over the finish line on the floor.

We hope to do the same with some of these proposals. Thank you for your leadership and your willingness to work across the aisle to take on this important issue.

With that, Ms. Wiley, you are up next. Thanks.

OPENING STATEMENT OF MAYA WILEY, PRESIDENT AND CEO, THE LEADERSHIP CONFERENCE ON CIVIL AND HUMAN RIGHTS, WASHINGTON, DC

Ms. WILEY. Good afternoon, Chairwoman Klobuchar, Ranking Member Fischer, my own Senator, Majority Leader Schumer, Brooklyn, to be specific, and all the Members of this esteemed Committee. It is a great honor to be before you.

I do just want to correct the record, because I am no longer on the faculty at The New School, although I have joined the University of the District of Columbia School of Law as the Joseph Rao Professor.

I am going to be brief because so much of what has been said I agree with, but really to elevate three primary points that I think are critical to remember and that I hope we will discuss more deeply today and in the future.

One is that we know disinformation and misinformation is not new and it predates artificial intelligence. That is exactly why we should deepen our concern and why we need government action, because as has already been said, and we at the Leadership Conference have witnessed this growth already even in the last two election cycles, artificial intelligence is already expanding the opportunity and the depth of not only disinformation in the sense of

elevating falsehoods about where people vote, whether they can vote, how to vote.

That goes directly to the ability of voters to select candidates of their choice and exercise their franchise lawfully. We have seen that it disproportionately targets communities of color.

I mean, even the Senate Intelligence Committee noted that when it was looking at Russian interference in the 2016 election, that the African American community was really disproportionately targeted by that disinformation.

That the tools of artificial intelligence we are already seeing in the generative sense of artificial intelligence, the deepfakes already being utilized by some political action committees and political parties.

That is something that already tells us that it is already in our election cycle and that we must pay attention to whether or not people have clear information about what is and is not accurate, what a candidate did or did not say, in addition to the other things that we have talked about.

But I also want to talk about the conditions in which we have to consider this conversation about generative artificial intelligence and our election integrity. You know, we only have a democracy if we have trust in the integrity of our election systems.

A big part of the narrative we have been seeing driving disinformation in the last two cycles has been the narrative that our elections, in fact, are not trustworthy. This is something we are continuing to see increase.

We have also watched as social media platforms have turned back from policies, have gutted staffing to ensure that their public squares, essentially that they maintain as private companies, adhering to their user agreements and policies in ways that ensure that everyone online is safe from hatred, safe from harassment, but also is clear what is and is not factual information.

I say that because we cannot rely on social media companies to do that on their own. We have been spending much of our time over the past few years focused on trying to get social media companies both to improve their policies, as well as to ensure that they are policing them fairly and equally.

With regard to communities that are particularly targeted for mis and disinformation, I can tell you what you have seen in many news reports. In many instances we have seen a gutting of the staffing that has produced the ability to do some of that oversight. Even when they had that staffing, it was inadequate.

We as a civil rights community, as a coalition of over 240 national organizations are very, very, very much in favor, obviously, of the bipartisan processes that we are able to participate in. But also, to say, unless we start to recognize both how people are targeted, who is targeted, and its increase in violence in our election cycles—not just, it is not just theoretical, it is practical, it is documented, and we are seeing an increase.

FBI data shows that we are at risk, but that we can take action both in regulating artificial intelligence and ensuring the public knows what is artificially produced, and also ensuring that we have oversight of what social media companies are doing, and whether

they are complying with their own policies and ensuring that they are helping to keep us safe.

Thank you.

[The prepared statement of Ms. Wiley was submitted for the record.]

Chairwoman KLOBUCHAR. Very good. Thank you very much, Ms. Wiley. Mr. Chilson.

OPENING STATEMENT OF NEIL CHILSON, SENIOR RESEARCH FELLOW, CENTER FOR GROWTH AND OPPORTUNITY, UTAH STATE UNIVERSITY, LOGAN, UTAH

Mr. CHILSON. Good afternoon, Chairwoman Klobuchar, Ranking Member Fischer, esteemed Committee Members.

Thank you for inviting me to discuss the influence of artificial intelligence on elections.

Imagine a world where our most valuable resource, intelligence, is abundant to a degree we have never seen. A world where education, art, and scientific innovations are supercharged by tools that augment our cognitive abilities. Where high fidelity political speech can be created by voices that lack deep pockets. Where real time fact checking and inexpensive voter education are the norm. Where AI fortifies our democracy.

That is a promise of AI's future, and it seems plausible to me. But if you take one message from my comments, it should be this: artificial intelligence and political speech is not emerging, it is here and it has been for years. AI technologies are entangled in modern content creation. This is not just about futuristic tech or deepfakes. It is about the foundational technologies that we use to craft our political discourse today.

Let's follow a political ad from inception to distribution. Today, an ad campaign director does not just brainstorm ideas over coffee. She taps tools like ChatGPT to rapidly prototype variations on her core message.

When her media team gathers assets, automatic computer vision tagging makes it a breeze to sift through vast image databases. Her photographers' cameras use AI. The camera sensors adjust to capture images based on the lens attached or the lighting conditions. AI powered facial and eye detection ensures that subjects remain in focus.

Apple's newly announced iPhone takes this to the next level. Its dedicated neural nets powering its computational photography. It is no exaggeration to say that every photo taken on an iPhone will be generated in part by AI.

AI also powers post-production. Speech recognition tools make it easy to do text based video edits. Sophisticated software automatically joints multiple raw video streams into a polished final product. Blemishes disappear and backgrounds are beautified because of AI, and tools like HeyGen make it possible to adapt the audio and video of a final ad into an entirely different language seamlessly. These are just some of the AI tools that are involved in creating content today. Some are new, but many others have been here for years and in use.

AI is so intricately woven into the fabric of modern content creation that determining whether a particular ad contains AI-gen-

erated content is very difficult. I suspect each Senator here has used AI content in their ad campaigns, knowingly or not.

Here is why this matters: because AI is so pervasive in ad creation, requiring AI content disclosures could affect all campaign ads. Check-the-box disclosures will not aid transparency, they will only clutter everyone's political messages.

And to address what unique problems? AI will facilitate more political speech, but there is no reason to think that it will shift the ratio of truth to deception. Historically malicious actors do not use cutting edge tech. Cheap fakes, selective editing, overseas content farms, and plain old Photoshop are inexpensive and effective enough.

Distribution, not content generation, is the bottleneck for misinformation campaigns. Money and time spent creating content is money and time that they cannot spend spreading it.

This Committee should continue to investigate what new problems AI raises. It could review AI's effects on past elections and should obviously closely monitor its use and effects on the coming election cycle. More broadly, Congress should establish a permanent central hub of technical expertise on AI to advise the many federal agencies dealing with AI related issues.

Remember, AI is here now, already affecting and improving how we communicate, persuade, and engage. Imprecise legislative approaches could burden political speech today and prevent the promise of a better informed, more engaging political dialog tomorrow.

Thank you for your attention. I am eager to address any questions that you have.

[The prepared statement of Mr. Chilson was submitted for the record.]

Chairwoman KLOBUCHAR. Thank you, Mr. Chilson. Mr. Cohn.

OPENING STATEMENT OF ARI COHN, FREE SPEECH COUNSEL, TECHFREEDOM, WASHINGTON, DC

Mr. COHN. Chair Klobuchar, Ranking Member Fischer, Members of the Committee, thank you for inviting me to testify today. It is truly an honor. The preservation of our democratic processes is paramount. That word processes, I think, highlights a measure of agreement between all of us here.

False speech that misleads people on the electoral process, the mechanics of voting, where to vote, how to register to vote. Those statements are particularly damaging, and I think that the Government interest in preventing those specific process harms is where the Government's interest is at its most compelling.

But a fundamental prerequisite to our prized democratic self-governance is free and unfettered discourse, especially in political affairs. First Amendment protection is at its zenith for core political speech and has its fullest and most urgent application to speech uttered during a campaign for political office.

Even false speech is protected by the First Amendment. Indeed, the determination of truth and falsity in politics is properly the domain of the voters, and to avoid unjustified intrusion into that core civic right and duty, any restriction on political speech must satisfy the most rigorous Constitutional scrutiny, which requires us to ask a few questions.

First, is the restriction actually necessary to serve a compelling government interest? We are not standing here today on the precipice of calamity brought on by seismic shift. AI presents an incremental change in the way we communicate, much of it for the better, and a corresponding incremental change in human behavior that predates the concept of elections itself.

Surely deceptively edited media has played a role in political campaigns since well before the advent of modern AI technology. There is simply no evidence that AI poses a unique threat to our political discussion and conversation.

Despite breathless warnings, deepfakes appear to have played little, if any, role in the 2020 Presidential election. While the technology has become marginally better and more available in the intervening years, there is no indication that deepfakes pose a serious risk of materially misleading voters and changing their actual voting behavior.

In fact, one study of the effect of political deepfakes found that they are not uniquely credible or more emotionally manipulative relative to non-AI manipulated media. The few instances of AI use in current election cycle appear to back that up.

Even where not labeled, AI-generated media that has been used recently has been promptly identified and subject to immense scrutiny, even ridicule.

The second question is whether the law is narrowly tailored. It would be difficult to draft a narrowly tailored regulation in specifically at AI. Such a law would be inherently under inclusive, failing to regulate deceptively edited media that does not utilize AI—media, which not only poses the same purported threat, but also has a long and demonstrable history of use compared to the relatively speculative fears about AI.

A law prohibiting AI-generated political speech would also sweep an enormous amount of protected and even valuable political discourse under its ambit. Much like media manually spliced to create the impression of speech that did not in fact occur, AI-generated media can serve to characterize a candidate's position or highlight differences between two candidates' beliefs.

In fact, the ultimate gist of a message conveyed through technical falsity may even turn out to be true. To prohibit such expression, particularly in the political context, steps beyond what the First Amendment allows.

But even more obviously, prohibiting the use of political AI-generated media broadly by anyone, in any place, at any time, no matter how intimate the audience or how the low the risk of harm, clearly is not narrowly tailored to protect against any harm as the Government might claim it has the right to prevent.

The third question is whether there is a less restrictive alternative. When regulating speech on the basis of content, the Government must choose the least restrictive means by which to do so. Helpfully, the same study revealing that AI does not pose a unique risk also points to a less restrictive alternative. Digital literacy and political knowledge were factors that uniformly increased viewer's discernment when it comes to deepfakes.

Congress could focus on bolstering those traits in the polity instead of enacting broad prophylactics. Another more fundamental

alternative is also available, more speech. In over a decade as a First Amendment lawyer, I have rarely encountered a scenario where the exposition of truth could not serve as an effective countermeasure to falsity, and I do not think I find myself in such a position today.

Nowhere is the importance, potential, or efficacy of counter speech more important than in the context of political campaigns. That is the fundamental basis of our democracy, and we have already seen its effectiveness in rebutting deepfakes. We can expect more of that.

Campaign related speech is put under the most powerful microscope we have, and we should not presume that voters will be asleep at the wheel. Reflexive legislation, prompted by fear of the next technological boogeyman, will not safeguard us.

Free and unfettered discourse has been the lifeblood of our democracy, and it has kept us free. If we sacrifice that fundamental liberty and discard that tried and true wisdom, that the best remedy for false or bad speech is true or better a speech, no law will save our democratic institutions, they will already have been lost.

More detail on these issues can be found in my written testimony and thank you again for the opportunity to testify today. I look forward to your questions.

[The prepared statement of Mr. Cohn was submitted for the record.]

Chairwoman KLOBUCHAR. Thank you, Mr. Cohn. I am going to turn it over to Senator Merkley in the interest of our schedule here, but I wanted to just ask one question, then I will come back—a twofold question.

I want to make sure you all agree that there is a risk posed by the use of AI to deceive voters and undermine our elections. Do you all agree with that? There is at least a risk?

[Nods in the affirmative.]

Chairwoman KLOBUCHAR. Okay, great. Then secondly, last, do you believe that we should work, and I know we vary on how to do this, but do you believe that we should work to ensure guardrails are in place that protect voters from this threat?

[Nods in the affirmative.]

Chairwoman KLOBUCHAR. Okay, great. Well, that is a good way to begin. I am going to turn it over to Senator Merkley, and then we will go to Senator Fischer, and then I think Senator Warner, who just so kindly joined us—has a scheduling crunch as well. Senator Merkley.

Senator MERKLEY. I thank you so much, Madam Chairwoman. Really, this is such an important issue. I am struck by a conversation I had with a group of my wife's friends who said, "how do we know what is real in political discourse? Because we hear one thing from one cable television, another from another."

I said, well, one thing you can do is go to trusted sources and listen to the candidates themselves. But now we are talking about deepfakes where the candidates themselves might be profoundly misrepresented.

I wanted to start by turning to you, Mr. Potter, in your role as a former Chair of the Federal Election Commission. Currently, it is not uncommon in ads to distort a picture of an opponent. They

get warped, they get blurred. They are kind of maybe tweaked a little bit to look evil.

Is there anything about that right now that is a violation of federal election law?

Mr. POTTER. No, it is not.

Senator MERKLEY. Okay. Thank you. You have got your microphone on there. Okay. He said, no, it is not. What if an ad, an individual quotes their opponent, and the quote is false. Is that a violation of that?

Mr. POTTER. No, it is not a violation of law—well, wait a minute. If you had a candidate misrepresenting what their opponent had said, under the current FEC rules, if the candidate did it themselves and they were misrepresenting the speaker, then it possibly could be.

Senator MERKLEY. An advertisement in which one candidate says, hey, my opponent took this position and said such and such, and that is not true. That is not true. That is a violation?

Mr. POTTER. If you are characterizing what your opponent said, I think that would not be a violation. It would be perhaps a mischaracterization.

If you create a quote and put it in the mouth of your opponent, and those words are inaccurate, then the FEC would look at it and say, is that a misrepresentation of the other candidate?

But it would have to be a deliberate creation of something that the opponent had not said, quoting it, as opposed to the candidate's opinion of what they had said.

Senator MERKLEY. Would a candidate's use of a completely falsified digital image of the opponent saying something that the person had never said, would that be illegal under current election law?

Mr. POTTER. I think it would. That is what I have urged the FEC in my testimony to make clear. That if they use—if a candidate creates a completely false image and statement by an opponent through this artificial intelligence, which is what could be done, that would violate existing law.

Senator MERKLEY. Okay, great. Secretary Simon, you talked about a leaflet that told people if their name ends in I think M through Z, to vote on Wednesday.

I picture now with modern technology, having that message come from a trusted source, a community leader in the voice of or the sound of, you know, if they were not identified as whomever.

Suddenly Barack Obama is on the line telling you, you are supposed to vote on Wednesday. Is such a presentation today a violation of election law?

Mr. SIMON. Boy, that is a tough one, Senator. Thanks for the question. I am hung up on a couple of details of Minnesota law. I do not know if it came up in the federal context. I think Mr. Potter might have the answer to that one. But, you know, not—I would say arguably, yes, it would be. Maybe not election law, but other forms of law. I mean, it is perpetrating a fraud.

Senator MERKLEY. Okay. I recognize there is some uncertainty about exactly where the line is, and that is part of why this hearing is so important as we think about this elaboration. Mr. Cohn, you said that deepfakes are not credible.

There was a 2020 study that 85 percent of the folks who saw the deepfakes said, oh, these are credible. It has much improved since then. Isn't there—I am not sure why you feel that a deepfake done, you know a well done one, is somehow not credible when studies have shown that the vast majority of people that see them go, wow, I cannot believe that person said that. They believe the fake.

Mr. COHN. Thank you for the question, Senator. A study in 2021 that actually studied a deepfake of Senator Warren actually particularly said that they could test whether or not misogyny also played a role into it, found that in terms of identifying whether something is a deepfake or not—the road is pretty—it does not really—it is not really more likely that someone is going to be moved by a deepfake than another piece of non AI-generated, manipulated media.

Senator MERKLEY. Okay. Thank you. My time is up. I just want to summarize by saying I—my overall impression is the use of deepfakes in campaigns, whether by a candidate or by a third party, can be powerful and can have people, can you believe what so-and-so said or what position they took. Because our eyes see the real person as if they are real, and so I am really pleased that we are holding this hearing and wrestling with this challenge. I appreciate your all's testimony.

Chairwoman KLOBUCHAR. Very good. Thank you very much, Senator Merkley. Senator Fischer.

Senator FISCHER. Thank you, Madam Chair. Mr. Chilson, you mentioned that AI tools are already common in the creation and distribution of digital ads. Can you please talk about the practical implications of a law that would ban or severely restrict the use of AI, or that would require broad disclosure?

Mr. CHILSON. Thank you for the question. Laws like this would mean that requiring disclosures, for example, would sweep in a lot of advertising content.

Imagine you are a lawyer advising a candidate on an ad that they want to run. If having AI-generated content in the ad means that ad cannot be run or that it has to have a disclosure, the lawyer is going to try to figure out whether or not there is AI-generated content in the ad. As I pointed out in my testimony, AI-generated content is a very broad category of content.

I know we all use the term deepfake, but the line between deepfake and tweaks to make somebody look slightly younger in their ad is pretty blurry and drawing that line in legislation is very difficult.

I think that in ad campaigns, as a lawyer advising a candidate, one will tend to be conservative, especially if the penalty is a potential private defamation lawsuit, with damages, where the defamation is per se.

I think that if the consequences are high that lawyers will be conservative, and it will chill a lot of speech.

Senator FISCHER. It could add to increased cost of elections, couldn't it, because of the increased cost in ads where you would have to meet all those requirements in an ad for the time you are spending there?

Mr. CHILSON. Absolutely. Increased costs. Also, less effective ads in conveying your content. It crowds out the message you want to get across. It could raise a barrier, too for smaller campaigns.

Senator FISCHER. Right. You also advocated an approach to preventing potential election interference that judges outcomes instead of regulating tools. What would that look like in practice?

Mr. CHILSON. I am hearing a lot of concern about deceptive content in ads and in campaigns overall. The question is, if that is the concern, why are we limiting restrictions to only AI-generated content?

When I say an outcome neutral test, I mean test based on the content that we are worried about, not the tool that is used to create it. If the concern is with a certain type of outcome, let us focus on that outcome and not the tools used to create it.

Senator FISCHER. Okay. Mr. Cohn, I understand that while all paid political advertisements already require at least one disclaimer, the Supreme Court has long recognized that compelled disclaimers could infringe on First Amendment rights. In your view, would an additional AI specific disclaimer in political advertisements violate political speakers' First Amendment rights?

Mr. COHN. Thank you for the question, Senator. I think there are two things to be concerned about. First, the Government still has to have a Constitutionally sufficient interest.

When it comes to the kinds of disclaimers and disclosures that we see presently, the informational interest that we are protecting is the identification of the speaker, who is talking to us, who is giving us this ad, which helps us determine whether we credit that ad or view it with some kind of skepticism.

Now, it is one thing to further that informational interest, and certainly it can make a difference in how someone sees a message. But that ties into the second problem, which is that pretty much as Mr. Chilson said, everything uses AI these days. If the interest is in making people a little more circumspect about what they believe, that actually creates the same liar's dividend problem that Secretary Simon said.

If everything has a disclosure, nothing has a disclosure, and it gives cover for bad actors to put these advertisements out, and the deceptive ones are going to be viewed just as skeptically as the non-deceptive ones because everything has to have a disclosure on it.

I am not sure that the, you know, proposed disclosure would actually further the Government interest, unless it is much more narrowly drawn.

Senator FISCHER. Some people have proposed using a reasonable person standard to determine whether an AI-generated image is deceptive. You have used that word here. Can you tell us how this type of standard has been used to regulate speech and other content?

Mr. COHN. Well, that is a great question, because who knows what the reasonable person is. But, you know, generally speaking, I think that is a harder standard to impose when you are talking about something like political speech.

It ties in closely, I think, with materiality. What is material to any particular voters? What is material to a group of voter? How

does the reasonable standard person correspond with the digital literacy of a particular person?

A reasonable person of a high education level may be much less likely to have a fundamentally different view of what a piece of edited material says than the original version. Whereas a person with lower—a lower education level might be more susceptible to it.

It really defies a reasonable person standard, particularly with such sensitive and important speech.

Senator FISCHER. Thank you. Thank you, Madam Chair.

Chairwoman KLOBUCHAR. I have returned. Senator Warner, the Chair of the Intel Committee, and one of the esteemed Members of the Rules Committee.

Senator WARNER. Thank you, Madam Chairwoman. I was actually just at a hearing on the PRC's use of a lot of these disinformation and misinformation tools. Candidly, I am not going to debate with the panel. I completely disagree with them on a number of topics, and I would love them to get some of the classified briefings we receive.

I really appreciate the fact that you have taken, Madam Chair, a lead on AI regulations around elections. As I think about the exponentially greater power of AI in misinformation, disinformation, the level of bot usage, it is child's play to what happened in terms of Russia's 2016 interference, the tools that are existing now.

I think it would be naive to underestimate that that we are dealing with a threat of a different magnitude. I applaud what you are doing. I actually think if we look at this, where our existing AI tools right now with very little increase in power, where can they have the most immediate effect that could have huge negative consequences and does not have to be necessarily generated by a potential adversarial, a nation like China, but just generally.

I would say those are areas where public trust is the key glue that keeps an institution stuck together. You have identified one in the question of public elections, and we have seen how public trust has been eroded, again, using somewhat now, you know, tools, and in 2016.

While we thank goodness the FEC has finally required the fact that a political ad on Facebook has to have some level of disclosure, as you know, that was your legislation, we still have not passed law number one to equalize disclosure requirements on social media and to equalize with traditional TV and broadcast. I think that is a mistake.

The other area, I would argue for consideration for the panel, maybe for a later time, is the fact that the other institution that is as reliant on public faith as public elections, that we could have the same kind of devastating effect if AI tools immediately are used, are faith in our public markets.

You know, there has been one example so far where AI tool did a false depiction of the Pentagon burning, had a disruption in the market. Child's play, frankly, the level of what could take place, maybe not in Fortune 50 companies, but Fortune 100 to 500 companies.

The ability to not just simply use deepfakes, but to generate tools that would have massive false information about products. Across a whole series of other ways that the imagination is pretty wild.

Again, I would welcome my colleagues to come for a classified briefing on the tools that are already being deployed by our adversaries using AI. Somehow this notion that there is, you know, well, if it is already a law, why do we need anything else?

Well, there are plenty of examples, and I will cite two, where because the harm is potentially so great, we have decided either in a higher penalty level or certain times a lower threshold of proof or in more extreme cases, even a prohibition, if the harm is so great that we have to think twice as a society. I mean, murder is murder.

But if that murder is created by a terrorist, there is a higher and differential level of—society has implied a different level of heinousness of that. We have lots of rules—or tools of war, but we have decided that, you know, there may be some tools of war, chemical weapons, atomic weapons, that go beyond the pale.

I think it would be naive to make assumptions at this point that would the potential that AI has, that we should not at least consider if these tools are unleashed, and I again applaud the fact that we are starting to drill down this issue around public elections.

Obviously, there is First Amendment rights that have to be respected. Might even be easier on public markets because I could very easily see massive AI disruption tools being used to disrupt public markets that could have hugely catastrophic effects, and we might then overreact.

But I do want to make sure I get in a question. I will go to Ms. Wiley. You know, one of the things we found in the 2016 elections were Russia disproportionately targeted black community in this country with misinformation and disinformation.

We just came from the hearing I was referencing where the Freedom House indicated that PRC's current influence operations, some using AI tools, some not, are once again targeting the black communities in our country.

You know, don't you think if the tools that were used in 2016 are now 100x, 1,000x, 1 million-x because of the enormous power of large language models and generative AI, don't we need to take some precautions in this space?

Ms. WILEY. Thank you, Senator. We absolutely must. What you are quoting is extremely important. It is also important to note, and when we look at the research and the RAND study that came out just last year showed that a minimum of 30–33 to 50 percent of all people in their subject pool of over 2,500 people took the deepfake to be accurate.

What they found is increased exposure actually deepened the problem. You know, the notion that you see it over and over again from different sources actually can deepen the impact, the belief in the deepfake.

I am saying that because part of what we have seen, and it is not only foreign governments, but it certainly includes them, but also domestic hate groups utilizing social media and utilizing the opportunity.

We are starting to have a lot of concerns about some of the ways the technology, particularly with chat bots and text message, actually can vastly increase exponentially the reach. But targeting communities that are more easily made afraid or giving false information about where and how to vote.

But also, I want to make this clear, too. We are seeing it a lot with people who are lawfully allowed to vote, but for whom English is not their first language. They have also been targeted, particularly Spanish speakers, but not also—also in the Asian community. We know that there is, and a lot of social science shows that there is real targeting of communities of color.

It does go to the way that we see even with political parties and political advertising, the attack on the integrity of our election systems, and even whether voters are voting lawfully or fraudulently in ways that have made people more vulnerable to violence.

Chairwoman KLOBUCHAR. Very good. Thank you, Senator Warner. I know Senator Britt was here earlier, and we thank her for being here. Senator Hagerty.

Senator HAGERTY. Thank you, Senator Klobuchar, Ranking Member Fischer. Good to be with you both. Mr. Chilson, I would like to start with you. If I could, just engage in a thought experiment with you for a few minutes. Let's go back to early 2020 when the COVID–19 pandemic hit.

Many policymakers, experts are advocating for things like mask mandates, shutting down schools, and mandatory remote learning.

In many states, many localities adopted mandates of that nature at the outset. I think we know the result of those mandates had great economic damage, particularly to small businesses, children's learning was set back considerably, and loss of liberty. What I am concerned about is that Congress and the Biden Administration may be finding itself right at the same place again when we are looking at artificial intelligence, and I do not want to see us make the same set of mistakes.

I would like to start with a very basic question, if I might, and that is, is artificial intelligence a term with an agreed upon legal definition?

Mr. CHILSON. It is not. It does not have an even agreed upon technical definition. If you read one of the leading treatises that many computer scientists are trained on, the Russell and Norvig book, they describe four different categories of definitions, and underneath those, there are many different individual definitions.

Then, if you run through the list of things that have been considered AI in the past and which nobody really calls AI now, you have everything from edge detection, which is in everybody's cameras, to letter detection, to playing chess, to playing checkers, things that once it works, we kind of stop calling it AI.

That paraphrases computer scientist John McCarthy who actually coined the term AI. There is not an agreed upon legal definition, and it is quite difficult actually to define.

Senator HAGERTY. Yes. Using broadly how we think about AI and AI tools, do political candidates and others that engage in political speech use AI today for routine functions like taking and editing pictures like you just mentioned, or for speech recognition, or for processing audio and video content?

Mr. CHILSON. Absolutely. Ads are created using and all content is created using many different algorithms. My cell phone here has many, many different AI algorithms on it that are used to create content.

Senator HAGERTY. I would like to use this scenario to illustrate my concern. Madam Chair, I would like to introduce this article for the record. It is one of many that cites this particular——

Chairwoman KLOBUCHAR. You have it in the record.

[The information referred to was submitted for the record.]

Senator HAGERTY [continuing]. that I will come back to. One of the proposals that is under consideration now would prohibit entities from using, "deceptive AI-generated audio or video visual media in election related speech."

This would include altering an image in a way that makes it inauthentic or inaccurate. That is a pretty vague concept. For example, age may be a very relevant factor in the upcoming 2024 elections. You may recall recent media reports, again, this is one of them right here, describing how President Biden's appearance is being digitally altered in photographs to make him look younger.

My next question for you, Mr. Chilson, if the Biden campaign were to use photo editing software that utilizes AI to make Joe Biden look younger in pictures on his website, could that use of artificial intelligence software potentially violate such a law against inaccurate or inauthentic images?

Mr. CHILSON. Potentially, I believe it could. The question should be, why does the use of those tools violate it but not the use of makeup and use of lighting in order to make somebody look younger.

Senator HAGERTY. Is there a risk then, in your view, that hastily regulating in a very uncertain a rapidly growing concept like AI might actually chill political speech?

Mr. CHILSON. Absolutely.

Senator HAGERTY. That is my concern too. My point is that Congress and the Biden Administration should not engage in heavy handed regulation with uncertain impacts that I believe pose a great risk to limiting political speech.

We should immediately indulge the impulse for Government to just do something, as they say, before we fully understand the impacts of the emerging technology, especially when that something encroaches on political speech.

That is not to say there are not a significant number of issues with this new technology. But my concern is that the solution needs to be thoughtful and not be hastily implemented. Thank you.

Chairwoman KLOBUCHAR. Thank you very much, Senator Hagerty. I will start with you, Senator Simon, and get at some of—I am sorry, Secretary of State Simon, and get at some of the questions that Senator Hagerty was raising. Just first, just for now, because all my colleagues are here, and I have not asked questions yet. Which state has consistently had the highest voter turnout of all the States in America?

Mr. SIMON. Senator, that would be——

Chairwoman KLOBUCHAR. Okay. Thank you very much.

Mr. SIMON. Yes, that would be Minnesota.

Chairwoman KLOBUCHAR. Especially because Senator Bennet is here, and he is always in a close race with me for Colorado. I thought I would put that on the record. Okay.

Senator Hagerty has raised some issues, and I wanted to get at what we are doing here with a bill that Senator Hawley, certainly not a Member of the Biden Administration, that Senator Hawley and I have introduced with Senator Collins and Senator Ricketts, Senator Bennet, who has been such a leader on this, Senator Coons, and others will be getting on it as well.

This bill gets at not just any cosmetic changes to how someone—this gets at materially deceptive ads. This gets at the fake ad showing Donald Trump hugging Dr. Fauci, which was a lie. That is what it gets at.

It gets at the person that looks like Elizabeth Warren but isn't Elizabeth Warren claiming that Republicans should not be allowed to vote. It is of grave concern to people on both sides of the aisle. Can you talk about and help us with this kind of materially deceptive content has no place in our elections.

Mr. SIMON. Thank you, Senator, for the question. I think that is the key. The materiality test and courts, it seems, are well equipped to use that test in terms of drawing lines.

I do not pretend to say—and I think Senator Hagerty is correct and right to point out that this is difficult, and that Congress and any legislative body needs to get it right. But though the line drawing exercise might be difficult, courts are equipped under something like a materiality standard to draw that line.

I think that materiality, it really in the realm of elections is not so different from other realms of our national life. It is true, as Mr. Cohn and others have said, that the political speech, the bar for political speech is rightly high. It is, and it should be.

But in some senses, it is no different than if someone were to say something false in the healthcare field. If someone said something just totally false, a false positive or negative attribute—if someone said that breath mints cure cancer or breath mints cause cancer or something like that, I do not think we have quite the same hesitation.

Political speech, of course there is a high bar, but courts, given the right language such as a materiality test, could navigate through that.

Chairwoman KLOBUCHAR. Right. I am going to turn to Mr. Potter, but I note that even in a recent decision, Supreme Court decision by Justice Barrett, a 7 to 2 decision, the Supreme Court was joined by Justice Barrett, Justices Roberts, Thomas, Alito, Kagan, Gorsuch, and Kavanaugh stated that the First Amendment does not shield fraud.

The point is that we are getting at a very specific subset, not what Mr. Cohn was talking about with the broad use of some of the technology that we have on political ads. Mr. Potter, you would be a good person to talk to.

You were a Republican appointee, Chair of the FEC. Can you expand on how prohibiting materially deceptive AI-generated content in our election falls squarely within the framework of the Constitution?

Mr. POTTER. Thank you, Madam Chair. The court has repeatedly said that it is Constitutional to require certain disclosure so that voters have information about who is speaking. There, I think Justice Kennedy in Citizens United was very clear in saying that voters need to know who is speaking, to put it in context.

Who the speaker is informs the voters' decisions as to whether to believe them or not. In those circumstances where we are talking about disclosure, it seems to me particularly urgent to have voters know that the person who is allegedly speaking is fake. That the person who they think is speaking to them or doing an act is actually not that person.

There, it is the negative of, yes, who is paying for the ad, but is the speaker actually the speaker. That would fit within the disclosure framework. In terms of the prevention of fraud, I think that goes to the fact that the court has always recognized that the integrity of our election system and citizen faith in that system is what makes this democracy work.

To have a circumstance where we could have the deepfake and somebody is being alleged to say something they never said or engage in an act where they never did, is highly likely to create distrust. Where you have a situation where that occurs, the comment has been made, well, the solution is just more speech.

But I think we all know, and there is research showing this, but we intuitively know that, you know, I saw it with my own eyes is a very strong perspective. To see somebody, hear them engaging in surreptitiously recorded racist and misogynist comments, and then have the candidate whose words and image have been portrayed say, that is not me, I did not say that, that is all fake.

Are you going to believe what you saw, or are you going to believe a candidate who says that is not me? I think that is your first inherent problem.

Chairwoman KLOBUCHAR. Thank you for doing it, and also in neutral terms, because I think we know it could happen on either side and why we are working so hard to try to get this done.

I would also add in this was on the disclosure comment with Scalia, who said in 2010 for an opinion concurrence. "For my part, I do not look forward to a society which, thanks to the Supreme Court campaigns anonymously hidden from public scrutiny and protected from the accountability of criticism. This does not resemble the home of the brave."

There has been a clear indication and why Senator Hawley, and Collins, and Senator Bennet, and a number of the rest of us drafted a bill that had the ability to look at this in a very narrow fashion, but also allowed for satire and the like.

I did find, Mr. Cohn's—I went over and told Senator Warner, some of your points, I might have to turn it over here, interesting. When we get to beyond the ones that would be banned, of which ones the disclaimer applies to, and that we may, you know, want to look at that in a careful light so that we do not have every ad—it becomes meaningless, as you said. I really did appreciate those comments.

With that, I am going to—I think it is Senator—I think our order is Senator Ossoff, because he has to leave. Is this correct? Then we go to Senator Welch, who has been dutifully here for quite a while.

Then Senator Bennet and then Senator Padilla, even though he does represent the largest state in our Nation and is a former Secretary of State. Hopefully that order will work out. If you need to trade among each other, please do. Thank you.

Senator OSSOFF. Thank you, Madam Chair. I think you just got to the root of the matter very efficiently and elegantly. You know, Mr. Cohn, I appreciate your comments, but I think that the matter that is being discussed here is not subjective, complex, judgments about subtle mischaracterization in public discourse.

We are talking about, for example, Senator Fischer, one of your political adversaries, willfully, knowingly, and with extreme realism, falsely depicting you or any of us, or a candidate challenging us, making statements that we never made in a way that is indistinguishable to the consumer of the media from a realistic documentation of our speech.

That is the most significant threat that I think we are talking about here. Mr. Potter, in your opinion, isn't there a compelling public interest in ensuring that that kind of knowing—knowingly and willfully deceptive content whose purpose, again, is not to express an opinion, it is not to caricature, but it is to deceive the public about statements made by candidates for office—isn't there a compelling public interest in regulating that?

Mr. POTTER. I think absolutely there is and that the court would recognize that compelling interest. I also—I mean, there is no argument that there is a compelling interest in fraudulent speech, as the Chair noted.

I think what you would find here is that in a circumstance where we are talking about this sort of deepfake, as opposed to the conversations about did you use a computer to create the text, but where you are creating a completely false image, I think we would have a compelling public interest and no countervailing private interest.

Because the First Amendment goes to my right, our right, to say what we think, even about the Government and in campaigns, without being penalized. But the whole point of this conversation is you are falsifying the speaker. It is not what I think my First Amendment right.

It is creating this fake speech where the speaker never actually said it. That, I think, is where the court would come down and say, creating that is not a First Amendment right.

Senator OSSOFF. Indeed, as you point out, there is substantial jurisprudence that would support the regulation of speech in this extreme case, with knowing and willfully deceptive fabrication of statements made by candidates for office or public figures.

Mr. POTTER. Yes. I think the distinction I draw is that the court has protected a candidate saying I think this even if it is false, or my opponent supports or opposes abortion rights. That may be a mischaracterization. It may be deceptive.

But if it is what I am saying, engaging in my First Amendment speech, mischaracterizing an opponent's position, that is in the political give and take. But I think that is completely different from what we are talking about here, where you have an image, or a voice being created that is saying something it never said.

It is not me characterizing it. It is putting it in the image of this candidate.

Senator OSSOFF. Thank you. Mr. Cohn, since I invoked your name earlier, I will give you the chance to respond. But is it your position that broadcast advertisements, which knowingly and willfully mischaracterize a candidate for office, and I do not mean mischaracterize as in mischaracterized their position or give shaded opinions about what they believe stand for or may have said in the past, but depict them saying things they never said for the purpose of misleading the public about what they said, is it your position that that should be protected speech?

Mr. COHN. Well, thank you for the question, Senator. I think there is two things.

First of all, you know, it is one thing to say the word fraud, but fraud generally requires reliance and damages. Stripping those requirements out of here into and effectively presuming them takes us well outside of the conceptualization of fraud that we know.

I think there are circumstances in which I would probably agree with you that things cross the line. But take, for example, two—just two examples. First, in 2012, the Romney campaign cut some infamous lines out of President Obama's speech in the you did not build that campaign ad.

They made it seem like he was denigrating the hard work of business owners, but instead he was actually referring to the infrastructure that supported those businesses. Just in this last election, the Biden campaign was accused of cutting out about 19 sentences or so from a President Trump campaign rally that made it sound like he was calling COVID–19 a hoax.

My point is not that these are good or valuable and that we need people to say these. It is that this is already a problem, and by trying to legislate them with AI specifically, instead of addressing it as Mr. Chilson said, the broader effect causes a Constitutional concern that the government interest is not actually being advanced.

Senator OSSOFF. I see. If I understand correctly, and do not let me put words in your mouth, but you agree, broadly speaking, with the premise that certain forms of deceptive advertising in the political arena are subject to regulation on the basis there is a compelling public interest in preventing outright, willful, knowingly deception, such as putting words in Senator Fischer's mouth she never put on, in a highly realistic way.

Your argument is that the question is not the technology used to do so, the question is the materiality, the nature of the speech itself. Is that your position?

Mr. COHN. Yes. I think that drawing the statute narrowly enough is an exceedingly difficult task. I think in principle is a, you know, pie in the sky concept. I think I agree with you, I just am not sure how to get from point A to point B in a manner that will satisfy strict scrutiny.

Senator OSSOFF. Forgive me, Senator Fischer, for invoking your example in that hypothetical. Thank you all for your testimony.

Chairwoman KLOBUCHAR. Okay, very good. Thank you. I will point out that while network TVs have some requirements and they take ads down when they find them highly deceptive, that is not going to happen online.

That is one of our problems here, why we feel we have to act and why we have to make clear that the FEC has the power to act as well, because otherwise we are going to have the Wild West right now on the platforms where a lot of people, as we know, are getting their news and there's no rules. Senator Welch.

Senator WELCH. Yes and thank you. Kind of following up on Senator Ossoff and Senator Klobuchar, nobody wants to be censoring, so I get that. What that line is, is very porous. But the example that Senator Ossoff just gave was not about political speech, it was of flat out fraud, right.

Whether it was AI-generated or it was used with older technologies in broadcast, would you guys agree that there should be a remedy for that?

Mr. COHN. Well, thank you, Senator. I am not entirely sure that we can define it exclusively as——

Senator WELCH. All right. Let me stop for a second, because what I am hearing you say is, it is really, really difficult to define, which I think it is, but your conclusion is we cannot do anything. I mean, the issue with AI is not just AI, it is just the amplification of the deception.

You know, something that happened to Senator Fischer is so toxic to trust in the political system, and that is getting out of control as it is. You know, I will ask you, Mr. Potter, how do we define that line between where you are doing something that is totally false versus the very broad definition of political speech.

Then one other thing I want to ask, there has to be some expectation that the platforms like, say, Google, take some responsibility for what is on the platform. They have been laying off the folks whose job it is to monitor this and make a judgment about what is a flat out deception.

How do we deal with this? Then second, what is your observation about the platforms like Twitter, now X, Google, Facebook, essentially laying off all the folks whose job it was within those organizations to be reviewing this material that is so dangerous for democracy?

Mr. POTTER. Yes. Let me start with the first one, which is I think what you are hearing from all the panelists. It is important to have a carefully crafted, narrow statute to withstand Supreme Court scrutiny, but also to work. The language that gets used is going to be the key question.

Senator WELCH. All right. We all agree on that, but there is a real apprehension, understandably so, that this is going to be censoring speech. I do not know who is going to draft the statute.

We will let all of you do that. But it is a real problem. But what about the platforms laying people off so that we do not even get real time information? It gets out—the false, the deceitful advertising is out there, and we do not even know it, and cannot verify that it is false.

Mr. POTTER. Right. If I could, one more line on your first question and then I will jump to your second.

Senator WELCH. Okay.

Mr. POTTER. On the first one, I think the comment, the examples cited by Mr. Cohn in terms of snippets being taken from a Romney speech or snippets from a Trump speech and then

mischaracterized, that to me falls on the line of that is defensible, permissible political speech that falls into the arena where we argue with each other over whether it was right or wrong, because in his example, those people actually said that and it was their words, and you are interpreting them or misinterpreting them, but they said it.

That is where I draw the line. Say where you are creating words they did not say, the technology we have heard about, where my testimony today, because I have been talking enough, can be put into a computer and my voice pattern can be used, and it can create an entirely different thing, where I sat here and said, this is ridiculous.

You should not be holding this hearing, and you should not regulate any of this. That could be created and be false.

Senator WELCH. Would there be any problem banning that? I mean, why would that be legitimate in any campaign? I will ask you, Mr. Chilson or Mr. Potter.

Mr. CHILSON. Rearranging somebody's speech to say something truthful, even if it is a misrepresentation, I do not think you could ban that. If, you know, if I had this, your recording of this speech——

Senator WELCH. No, we are talking about using the—using whatever technology to have somebody, me, saying something I never said, at a place I never went. Yes, sorry. Thank you.

Mr. CHILSON. I think that it would really depend. If you have AI video of somebody saying something that they did not say in a place that they did not go, but it makes them look good, right? It is not defamatory in any way. It is truthful and it is positive on you. It would be hard to draw a line that would ban one of those and not the other.

Chairwoman KLOBUCHAR. Okay. Mister—Senator Bennet.

Senator BENNET. Thank you, Madam Chair. Thank you very much for holding this hearing and thank you for the bill that you have allowed me to co-sponsor as well. I think it is a good start in this area. Thank you, the witnesses, for being here.

You know, not everybody up here, and I think everybody on this panel, is grappling with the newness of AI. Disinformation itself, of course, is not something that is new. Ms. Wiley, this is a going to be a question for you once I get my—through it.

It was common in the 20th century for observers and journalists or maybe journalists themselves to say that if it bleeds, it leads. Digital platforms, which have in many cases, I think tragically replaced traditional news media, have turned this maxim into the center of their business model, creating algorithms that are stoked by outrage to addict humans, children in particular, but others to their platforms to sell advertising to generate profit.

That has then found its way into our political system, and not just our political system. In 2016, foreign autocrats exploited the platforms' algorithms to undermine Americans' trust in our institutions, our elections, and each other.

I remember as a Member of the Intelligence Committee just being horrified by not just the Russian attack on our elections, but also the fact that it took Facebook forever to even admit that it had happened—that they had sold ads to Russians that were then used

to anonymously attack our elections and spread falsehoods—in our democracy.

In 2017, you know, it was Meta—now Meta, where it was Facebook, now Meta's algorithms played what the United Nations described as a determining role in the Myanmar genocide. Facebook said that they, "lose some sleep over this." That was their response. Clearly not enough sleep, in my view.

Thousands of Rohingya were killed, tortured, and raped, and displaced as a result of what happened on their platform with no oversight and with no even attempt to try to deal with it. In 2018, false stories went viral on WhatsApp, warning about gangs of child abductors in India.

At least two dozen innocent people were killed, including a 65 year old woman who was stripped naked and beaten with iron rods, wooden sticks, bare hands and feet. Just last night, The Washington Post reported—by the way, these are not hypotheticals. Like this is actually happening in our world today.

Just last night, The Washington Post reported how Indian political parties have built a propaganda machine on WhatsApp with tens of thousands of activists spreading disinformation and inflammatory religious content. Last month, when the Maui wildfires hit, Chinese operatives capitalized on the death of our neighbors and the destruction of their homes, claiming that this was the result of a secret weather weapon being tested by the United States.

To bolster their claims, their post included what appear to be AI-generated photographs. Big tech has allowed this false content to course through our platforms for almost a decade. We have allowed it to course through these platforms.

I mean, I am meeting every single day, it is not the subject almost every day at home. I did, literally did on Monday with educators in the Cherry Creek School District, listening to them talk about the mental health effects of these algorithms. I know that is not the subject of today's hearing, but let me tell you something, our inability to deal with this is enormously costly.

I am a lawyer. I believe strongly in the First Amendment, and I think that is a critical part of our democracy, and a critical part of journalism and politics. We have to find a way to protect it. But it cannot be an excuse for not acting. The list of things that I am talking about here that I read today, these are foreign actors to begin with that are undermining our elections.

The idea that somehow we are going to throw up the First Amendment in their defense cannot be the answer. We have to have a debate about the First Amendment to be sure. We need to write legislation here that does not compromise or unconstitutionally impinge on the First Amendment.

I totally agree with that. We cannot go through another decade like the last decade. Ms. Wiley, I almost am out of time, but just in the last seconds that I have left, could you discuss the harm disinformation has played in our elections and the need for new regulation to grapple with traditional social media platforms, as well as the new AI models that we are talking about here today? I am sorry to leave you so little time.

Ms. WILEY. No, thank you. Just to be very brief and very explicit, we have been working as a civil rights community on these issues for a decade as well, Senator Bennet.

What we have seen, sadly, is even when the social media platforms have policies in place prohibiting conduct which they are Constitutionally allowed to do, to say you cannot come on and spew hate speech and disinformation without us either demoting it or labeling it or kicking you off the platform potentially, right, in the worst—for the worst offenders.

Yet, what we have seen is, sadly and frankly, not consistent enforcement of those policies and most recently actually pulling back from some of those policies that enable not only a safe space for people to interact—you know, we should just acknowledge that for 8 year olds and under, we have seen double the rate of 8 year olds on YouTube since 2017, double.

It really is significant what we have seen, both in terms of telling people they cannot vote or sending them to the wrong place. But it is even worse because as we saw with YouTube, a video that went viral out of Georgia, that gets to Arizona, and then we have an elected officials who call out vigilantes to go armed to mail drop boxes, intimidating vote, which essentially intimidates voters from dropping off their ballot.

Senator BENNET. My colleague from California has waited. I apologize.

Chairwoman KLOBUCHAR. Yes, I think we are going to let him go.

Senator BENNET. One observation, that that, Ms. Wiley, is such an important point. In 2016, the Russians were telling, the Russian Government was telling the American people that they could not go someplace to vote. It is the point you are making. They do not have a First Amendment right to do that and we need to stop it.

Chairwoman KLOBUCHAR. Okay. Thank you for your patience and your great leadership on elections. Senator Padilla.

Senator PADILLA. Thank you, Madam Chair. Want to just sort of associate myself with a lot of the concerns that have been raised by various Members of the Committee today. But as the Senate as a whole is having a more of a complete comprehensive conversation about AI, I think Leader Schumer and others have encouraged us to consider balanced thinking.

We want to minimize the risk, the negative impact of AI, but at the same time, be mindful of the potential upside and benefits AI, not just in elections, but across the board. While I share some of the concerns, I have a question relative to the potential benefits of AI. One example of the potential benefits is the identification of disinformation super spreaders, right.

We are all concerned about disinformation. There are some small players and big players. I am talking about super spreaders who are influencers, accounts, web pages, and other actors that are responsible for wide dissemination of disinformation.

AI can, if properly implemented, help scrape for these actors and identify them so that platforms and government entities can respond accordingly. I see some heads nodding, so I think the experts are familiar with what I am talking about.

Another example is, in the enforcement of AI rules and regulations. One example, Google just announced that it will require political ads that use synthetic content to include a disclosure to that effect.

Using AI to identify synthetic content will be an important tool for enforcing this rule and others like it. Question for Mr. Chilson. What—can you think of one or two other examples of benefits of AI in the election space?

Mr. CHILSON. Absolutely. As I said in my statement, it is already integrated deeply into how we create content, and it has made it much easier to produce content. One of the things that comes to mind immediately is a relatively recent tool that lets you upload, sample a video and then pick a language to translate it into.

That—and it translates not just the audio, but it also translates the image so that it looks like the person is speaking in that language. That type of tool to quickly be able to reach an audience that maybe was harder to reach for the campaign before, especially campaigns that do not have deep resources, I think that is a powerful, potential tool.

Senator PADILLA. Thank you. Question for a former colleague, Secretary Simon.

I think that one short term tool that could benefit both voters and election workers is the development of media literacy and disinformation toolkits that could then be branded and disseminated by state and local offices.

Do you think it would be helpful to have additional resources like this from the federal level to boost media literacy and counter disinformation?

Mr. SIMON. Thank you, Senator, and good to see you. We in the Secretary of State community miss you, but we are glad you are here as well. Thank you for the question. Yes, I think the answer to that is yes.

When it comes to disinformation or misinformation, I think you put your finger on it, media literacy really does matter. I mean, I know you are aware, and I alluded to earlier in my testimony, the Trusted Sources Initiative of the National Association of Secretaries of State.

The more we can do to channel people to trusted sources, however they may define that, I would like to think it is the Secretary of State's Office, but someone may not. Someone may think it's a county or a city or someone else, I think that would be quite helpful.

Senator PADILLA. Thank you. While we cannot combat disinformation—well, we cannot combat disinformation, whether it is AI disinformation or any other form, without fully understanding where disinformation comes from and how it impacts our elections.

We know there is numerous large nonpartisan organizations, I would emphasize that, nonpartisan groups that are dedicated to studying and tracking disinformation in order to help our democratic institutions combat it. But these organizations are now facing a calculated legal campaign from the far right under the guise of fighting censorship, to halt their research into and work to highlight disinformation. Just one example The Election Integrity Partnership, led jointly by Stanford Internet Observatory and the Uni-

versity of Washington Center for an Informed Public, tracks and analyzes disinformation in the election space and studies how bad actors can manipulate the information environment and distort outcomes.

In the face of a legal campaign by the far right, this work is now being chilled, and the researchers are being silenced. This is happening even as some platforms are getting their own trust and safety teams that previously helped guard against election hoaxes and disinformation on their platforms.

Ms. Wiley, what impact does the right wing campaign to chill disinformation researchers have on the health of our information ecosystems?

Ms. WILEY. Well, quite sadly and disturbingly, we are seeing the chilling effect take effect, meaning we are seeing research institutions changing what they are researching and how.

I think one thing I really appreciate about this panel is I think our shared belief, not just in the First Amendment, but in the importance of information and learning, and the importance of making sure we are disseminating it broadly.

There is nothing more important right now than understanding disinformation, its flow, and how better to identify it in the way I think everyone on the panel has named. I think we have to acknowledge that.

Certainly, there is enough indications from higher education in particular that it has had a devastating impact on our ability to understand what we desperately have to keep researching and learning about.

Senator PADILLA. Thank you. Thank you, Madam Chair.

Chairwoman KLOBUCHAR. Well, thank you very much. Thank you for your patience and that of your staffs. I want to thank everyone. We could not have had a more thorough hearing. I want to thank Senator Fischer and the Members of the Committee for the hearing.

I also want to thank the witnesses for sharing their testimony, the range of risks with this emerging technology, and going in deep with us about potential solutions and what would work. I appreciated that every witness acknowledged that this is a risk to our democracy, and every witness acknowledged that we need to put on some guardrails.

While we know we have to be thoughtful about it, I would emphasize the election is upon us. These things are happening now. I would just ask people who are watching this hearing, who are part of this, who are, you know, within—with the different candidates or on different sides, that we simply put some guardrails in place.

I personally think giving the FEC some clear authority is going to be helpful. I think doing—then, of course, doing some kind of ban for the most extreme fraud is going to be really, really important, and I am so glad to have a number of Senators joining me on this, including conservatives on the Republican side, and then figuring out disclaimer provisions that work. That has been the most eye opening to me as we have this hearing today about which things we should have them cover and how we should do that. That is where I am on this. I do not want that to replace the ability, and

this is what I am very concerned about, to actually take some of this stuff down that is just all out fraud in the candidates voices and pretending to be the candidate.

Clearly, the testimony underscored the importance of congressional action, and I look forward to working with my colleagues on this Committee in a bipartisan manner, as we did in the hardest of circumstances in the last Congress and last years, including, by the way, not just the Electoral Count Reform Act bill that we passed through this Committee with leadership in this Committee, but also the work that we did in investigating security changes that were needed at the Capitol, along with the Senator Peters and Portman at the time over in the Homeland Security Committee—the list of recommendations that Senator Blunt, the Ranking Member at the time, and I, and those two leaders came up with, most of which have been implemented with bipartisan support.

We just have a history of trying to do things on a bipartisan basis. That cries out right now for the Senate to take a lead, hopefully before the end of the year. We look forward to working on this as we approach the elections and certainly as soon as possible.

The hearing record will remain open for a week, only a week, because, like I said, we are trying to be speedy, and hope the Senate is not shut down at that time. We will find a way to get your stuff, even if it is.

But we are hopeful, given that nearly 80 percent of the Senate, actually 80 percent, the few people who worked on, that supported the bill last night that Senator McConnell and Senator Schumer put together to avoid a Government shutdown.

We go from there in that spirit, and this Committee is adjourned. Thank you.

[Whereupon, at 5:22 p.m., the hearing was adjourned.]

APPENDIX MATERIAL SUBMITTED

Testimony of The Honorable Steve Simon, Minnesota Secretary of State
U.S. Senate Committee on Rules & Administration
"AI and the Future of our Elections"
September 27, 2023

Chair Klobuchar, Ranking Member Fischer, and members of the committee:

Thank you for this opportunity. I am grateful for your willingness to engage on this important topic, and I am honored to be here.

Artificial Intelligence is not a threat to American democracy in and of itself, but it is an emerging and powerful amplifier of existing threats.

All of us who touch the election process must be watchful and proactive – especially as the 2024 presidential contest approaches.

A year ago, we were not talking that much about generative AI. The release of the newly accessible tools such as ChatGPT changed all that. In the hands of those who want to mislead, AI is a new and improved tool. Instead of stilted communications with poor grammar, generative AI can provide apparent precision and clarity.

There is a potential threat to the administration of elections. We are talking about an old problem (election misinformation and disinformation) that can now more easily be amplified. One possible danger could come from an innocent circumstance - AI software simply fails to grasp the nuances of our state-by-state election system. A prominent computer scientist in Minnesota named Max Hailperin made this point in an article several months ago. He asked ChatGPT questions about Minnesota election law – and the program gave the wrong answers to several questions. Intentional misdirection? Probably not. Still, it is a danger to voters who may get bad information about critical election rules.

In the wrong hands, AI could be used to misdirect intentionally – and in ways that are more advanced than ever. I remember seeing a paper leaflet from twenty or more years ago, distributed in a particular neighborhood, advising that in the coming election, voting would occur on Tuesday for those whose last names begin with the letters "A" through "L," while everyone else would vote on Wednesday. That was a paper leaflet. Now imagine a convincing-seeming email or deepfake conveying that kind of disinformation in 2024. The perpetrators could be domestic or foreign. In fact, the Department of Homeland Security has recently warned that our foreign adversaries may use AI to sharpen their attacks on our democracy.

One last point on potential consequences. The Brennan Center recently identified a so-called "liar's dividend" from the use of AI. Simply put, the mere existence of AI can lead to undeserved suspicion of messages that are actually true; a video, for example, that contradicts a person's preconceived ideas may now be simply dismissed as a deepfake.

The bottom line is that misdirection in elections can cause disruption. If AI misdirects, it could become an instrument of disruption.

So, what can be done about it?

In our office, we are trying to be proactive.

- First, we are leading with the truth – pushing out reliable and accurate information while also standing up to mis- and disinformation quickly.
- Second, we have been working with local and federal partners to monitor and respond to inaccuracies that could morph into conspiracy theories on election-related topics.
- Third, we have emphasized media literacy. The National Association of Secretaries of State has helped with its "trusted sources" initiative – urging Americans to seek out sources of election information from Secretaries of State and local election administrators.
- Fourth, our cyber defenses are strong. We have invested time and resources in guarding against intrusions that could introduce misleading information to voters.

As for possible legislation, I believe a federal approach would be helpful. The impacts of AI will be felt at a national level, so I applaud bipartisan efforts such as the Protect Elections from Deceptive AI Act and the Honest Ads Act.

There is a critical role for the private sector, too. Companies have a responsibility to the public to make sure their AI products are secure and trustworthy. I support the efforts already underway to encourage adherence to basic standards.

Let me end on a note of cautious optimism. AI is definitely a challenge, but in some ways, we have confronted similar challenges before. With each technological leap, we have generally been able to manage the potential disruptions to the way we receive and respond to information. The move to computerization, the arrival of the internet, and the emergence of social media all threatened to destabilize information pathways. In short order, the American people got smart about those things. They adapted.

AI may be qualitatively different from those other advances. But if we get better at identifying false information, and if we continue to rely on trusted sources for election information, we can overcome many of the threats that AI poses (while harnessing its benefits to efficiency and productivity).

Thank you for inviting me to testify today. I look forward to our continued partnership.

Statement of Trevor Potter

President, Campaign Legal Center

U.S. Senate Committee on Rules & Administration

Hearing on "AI and the Future of our Elections"

September 27, 2023

Chairwoman Klobuchar, Ranking Member Fischer, and Members of the Rules Committee, thank you for the opportunity to testify at this important hearing about the impact of artificial intelligence (AI) on our elections, as well as the urgent need for federal action to regulate this rapidly developing technology.

I am the founder and president of Campaign Legal Center (CLC), a nonpartisan, nonprofit organization that advances democracy through law at the federal, state, and local levels. Among our mission areas, CLC advocates for reforms to strengthen and ensure the consistent and robust enforcement of campaign laws in the United States. Prior to founding CLC, I served as Chairman of the Federal Election Commission (FEC), and have also been legal counsel to several presidential campaigns, and an advisor to the drafters of the Bipartisan Campaign Reform Act of 2002.

Today, my testimony describes how AI fits into the broader context of our campaign finance system. I will discuss how AI tools can be used, and increasingly are being used, to design and spread fraudulent or deceptive political communications that infringe on voters' fundamental right to make informed decisions at the ballot box. The issue is not automation itself, but the ways in which this technology, without the proper safeguards, can distort reality and undermine our democratic process of self-governance.

In addition to presenting these concerns, my testimony outlines several recommendations for federal regulation that could mitigate the harms of AI's usage in our elections. These proposals include action the FEC could take under current

law to address the dangers of fraudulent misrepresentation through AI-generated political advertisements. They also include measures that Congress could adopt to prohibit the most pernicious uses of AI to manipulate or disrupt elections, as well as to provide greater transparency for voters on the receiving end of communications made using this new technology.

The solutions described below are by no means mutually exclusive or exhaustive. AI technologies will continue to develop, and the challenges facing our democracy will evolve with them. As a result, I hope this discussion about AI and elections continues long after today's hearing.

AI and Our Campaign Finance System

To understand what is at issue when we talk about AI impacting our elections, a good place to start is how money influences our elections. Every election cycle, candidates, parties, and a wide array of outside groups spend billions of dollars to influence voters and sway election results, and these spending figures continue to grow – breaking new records cycle after cycle. The nonpartisan nonprofit OpenSecrets estimated that federal election spending in 2020 exceeded $14.4 billion, an unprecedented sum for a presidential election cycle. In 2022, spending on the federal midterms was estimated to be around $8.9 billion, likewise breaking the previous record for spending on a midterm election.

Much of this money is spent on creating, targeting, and distributing electoral communications, including ads advocating for or against or featuring candidates, and this is an area where AI has already begun to influence what voters are seeing. If AI use in ads becomes more commonplace, it could significantly impact our elections, especially when the technological power of AI is magnified by the spending power of billions of dollars.

AI has the power to manipulate what viewers are seeing and hearing in a way that is as convincing as it can be misleading, and that presents a unique challenge. For voters to decide how to vote, they have to parse through the many messages they are being bombarded with every election cycle and decide what to believe. They have to be able to evaluate the credibility of electoral messages and the underlying motivations of the people paying for them. AI has the potential to make that task much more difficult because the technology can be used to craft a very convincing and realistic misrepresentation of who is speaking, what is being communicated, or even whether something shown in an ad really happened.

Campaign finance laws allow voters to assess an ad's credibility and reliability by requiring transparency about who is paying for the ad and who is spending money to influence our elections. Certain electoral ads must make clear on their face—with a visual or audio statement in the ad itself—who is responsible for and authorized

the content the viewer is hearing or seeing. Candidates must "stand by" their electoral messages. Through mandatory disclosure reports and disclaimer requirements, Congress has taken steps to protect voters' rights to be informed about who is behind political communications that influence our elections. Voters, equipped with this information, still need to decide for themselves whether what they are seeing or hearing is credible and whether it will influence how they will vote. That is where AI presents a real problem that federal policymakers must address.

The ability of AI to create an extremely convincing yet imperceptibly false alternative reality poses a serious threat to the voting public's ability to properly evaluate political messages seeking to influence their voting decisions—a First Amendment interest recognized by the Supreme Court. AI could facilitate a political landscape where electoral ads are increasingly used for manipulation: to misrepresent who is speaking and what is being said. AI could thus interfere with voters' ability to meaningfully evaluate the candidates vying to represent them, while also impeding the ability of candidates and political parties to effectively communicate their messages to voters.

If voters are unable to trust that what they are seeing is real, then they could be easily misled about a candidate's positions or actions. At a time when the public's trust in our political process is already extraordinarily low—with a recent poll by the Brookings Institution showing that only 20% of Americans feel "very confident" in the integrity of the U.S. election system, while 56% have "little or no confidence" that our elections represent the will of the people—the potential uses of AI to mislead or defraud voters threatens to further erode confidence in our system of government. This could lead more voters to disengage from the political process, undermining our democracy.

Applications of AI in Political Ads

AI could affect electoral communications in a variety of ways—some relatively innocuous and others deeply concerning. There has been extensive reporting on the use of AI to make so-called "deepfakes," which mimic, distort, or fabricate the voice or appearance of a person and create the very realistic but false impression of a person saying or doing something they did not actually do or say.

On the relatively innocuous end, people may recall an image that circulated online earlier this year that depicted Pope Francis wearing a distinctive white puffy coat that was apparently the creation of a high-end fashion designer. In reality, the Pope never wore the coat. This same technology can be used in the context of elections to fraudulently (or at least misleadingly) create the appearance of a candidate doing or saying something they never did or said—with the underlying intention of manipulating or deceiving the public and influencing voters and elections.

For instance, a super PAC recently used AI to generate audio mimicking the voice of a candidate it opposed, speaking words that the candidate had posted on social media. The voice was created by AI; it was not actually the candidate speaking, although it sounded convincingly like his voice. If the super PAC had not publicly acknowledged using AI to recreate the sound of the candidate's voice, any listener reasonably could have believed it was actually the candidate speaking in the ad.

Another recent example went further, using AI to depict an event that never occurred: A presidential campaign used AI to create an ad with images depicting another candidate hugging Dr. Anthony Fauci. The images show something that had not occurred, but a reasonable person easily could have concluded that it had, thus intentionally and artificially interfering with voters' ability to decide whether they can trust what they are seeing.

To be sure, political ads can distort reality even without AI. Not every sales pitch or commercial we see is the unvarnished truth. Tools like photoshop have become staples of the media we consume, whether we are aware of it or not. A voice actor could imitate a candidate's voice, and an impersonator could pretend to be a candidate in an ad. But AI is an unprecedented game changer because of the technology's unique ability to easily create deceptively realistic false content. If left unregulated, AI could make it so common to see the false depicted as true that the public, and, in particular, voters, will be unable to know whether what they are seeing or hearing is real. The ease with which such false content could deceive voters threatens to undermine our democratic process.

To illustrate what is at stake, consider another recent example: An AI deepfake that was anonymously released on social media in the weeks before Chicago's recent mayoral election. This AI-generated ad depicted an image of a candidate and an imitation of that candidate's voice saying something to the effect of "back in my day, cops would kill 17 or 18 people and 'nobody would bat an eye.'" The complete lack of transparency about who paid for the ad, and the fact that it was not disclosed as an AI-generated fabrication, made it virtually impossible for viewers to know that this was an overt attempt at electoral manipulation. Concerningly, the video was reportedly viewed thousands of times before it was reported as false content and taken down.

We cannot say for certain how ads like those described above might impact a particular election, but the potential for manipulation has to deeply concern anyone who cares about the integrity of the democratic process. No less concerning would be AI deepfake ads that could undermine the administration of elections, such as by misrepresenting where and when people should go to vote, presenting false information about one's eligibility to vote, or other blatant attempts at electoral misinformation.

AI could also affect elections in less obvious ways. For example, AI could be used to determine who sees communications, which political consultants might refer to as the "targeting" of ads. With enough data "scraped" from various sources, AI has the potential to know what you like to eat, wear, and do—perhaps to know our preferences better than we know them ourselves. It could "categorize" voters based on their preferences to decipher their receptiveness to certain ideas or values over others, and then show them ads uniquely tailored to those preferences.

Again, targeting is not a new concept. Marketing experts and political consultants are in the business of trying to help their clients reach their desired audience, and polling and sample testing can help them determine not only what goes into an ad but who the ad is being shown to, and on what medium. Yet, an AI-powered advertising platform that shows different versions of ads to different viewers, based on their perceived group identity or their identifiable political views, opens new possibilities for problematic microtargeting.

Microtargeting political ads based on user data, a process that is invisible to voters, means that audiences have little understanding of what other voters are being shown or told. The ability to secretively direct a range of specially tailored and potentially conflicting messages to different audiences is incompatible with a transparent democratic process. By showing audiences more content that caters to their existing views and preferences, microtargeting can feed into echo chambers that exacerbate polarization. AI-powered microtargeting could make these problems even worse, as in theory, every person might see a slightly different version of an ad, recalibrated constantly based on viewer interest and engagement.

To summarize the concerns created by AI's usage in our elections, AI has the potential to sow doubt and mistrust among voters trying to evaluate the credibility of election ads they view, by making it easy and inexpensive to fabricate manipulative and false content that can be incredibly convincing. There may be relatively innocuous uses of AI in elections, but there are also deeply concerning applications. While AI did not invent the danger of electoral manipulation, it has the capacity to make existing problems much worse because of the power of the technology to easily make the false look real, and to micro-target segments of voters in ways not previously possible. Regulators must recognize and proactively address these risks immediately.

Recommendations for Federal Action

There are many ways that AI could impact elections, and policymakers must act now to prevent some of the most dangerous outcomes, as well as to ensure that even the more innocuous applications of AI are handled responsibly.

Before detailing my recommendations for federal regulation, it is important to note that the application of AI to political communications is clearly in its infancy, and we do not yet know what may be possible with AI in the future. Nevertheless, we have already seen enough to be concerned about the ramifications for the integrity of the democratic process.

It is an understatement to observe that technology develops more quickly than the law. It took more than a decade for the FEC to update the rules for disclaimers on digital electoral ads, during which time digital advertising evolved and expanded to new mediums, including streaming media platforms. Technology evolved rapidly and political advertisers changed their practices to better reach their audience. Trailing years behind was the agency responsible for ensuring voters would know who paid for these ads.

Congress must consider this dynamic when addressing the current threats posed by AI. The technology will continue to advance, and political advertisers—as well as those that employ them—will respond swiftly. New rules will need to be flexible enough to remain relevant over time, yet pointed enough to ensure that voters are protected.

With that in mind, I believe there are three concurrent paths for addressing these issues under the law through our campaign finance system. The first is to use and enhance the FEC's existing authority to protect elections against fraud. The second is to pass new legislation prohibiting the most pernicious uses of AI to influence elections. The third is to expand existing statutory disclosure and disclaimer requirements to ensure voters know when AI is used in election-related communications.

On the first point, the FEC currently has authority, under the Federal Election Campaign Act, to prohibit "fraudulent misrepresentation" for federal candidates. The non-profit organization Public Citizen has petitioned the agency to clarify explicitly, through the rulemaking process, that fraudulent uses of AI are included in this existing prohibition. **I believe the FEC ought to take this action**. This relatively simple regulatory action would clarify the application of long-established rules that are animated by the same underlying concerns that apply to AI: the law must ensure that candidates and voters alike are protected from fraud.

However, while this reform is relatively straightforward and simple, it is also narrow. More still needs to be done. The existing law only prohibits federal candidates, along with their employees and agents, from misrepresenting themselves as speaking for another candidate or party "on a matter which is damaging" to that candidate or party. It also prohibits any person from fraudulently misrepresenting their authority to solicit funds on behalf of any candidate or party.

The statute, at present, is not broad enough to address the wide variety of ways AI could be used to manipulate voters and undermine elections. **As a result, Congress should expand this provision**, and in fact, the FEC has long asked for such an expansion. Since 2004, the FEC's annual legislative recommendations have asked Congress to expand the relevant provision to "encompass all persons purporting to act on behalf of candidates and real or fictitious political committees and political organizations." Moreover, the agency has asked Congress to "remove the requirement that the fraudulent misrepresentation must pertain to a matter that is "damaging" to another candidate or political party." As modified in this way, the law would ban fraudulent electoral ads using AI to speak or act on a candidate's behalf—regardless of the person behind the ad, or whether the ad was "damaging" to the candidate or party it fraudulently depicted.

These relatively simple amendments to the federal fraudulent misrepresentation law would empower the FEC with greater authority to combat election fraud—authority the agency has sought for nearly two decades. While those changes would be important, they would not be enough to protect against some of the most pernicious uses of AI, which may not be limited to depicting a candidate speaking or acting on their own behalf.

As a second action, Congress should also pass a law that specifically prohibits uses of AI for the purpose of engaging in electoral fraud or manipulation—an area where the government has a clear, compelling interest in protecting voters and the integrity of the electoral process.

In recent years, we have seen efforts to undermine our elections, some of which were engineered by foreign governments, through systematic efforts to mislead voters. A similar future effort using AI to create and distribute manipulative or fraudulent content could be even more damaging, heightening the threat to our elections and, in a real sense, to our national security. Alarmingly, examples of this new avenue for election interference have begun to appear, with reports indicating that suspected Chinese operatives have already used AI-generated images to spread disinformation and create controversy along America's socioeconomic and political fault lines. The risk extends to bad actors at home and abroad. To mitigate this concern, Congress should enact a law that prohibits the use of AI to manipulate and deceive voters or disrupt the administration of elections.

It is important to note here that voters have a well-recognized First Amendment interest in being able to assess the political messages they see; there is no countervailing First Amendment right to intentionally defraud or deceive voters or interfere in elections. A narrow law prohibiting the use of AI to engage in fraud designed to undermine the electoral process would rest on firm constitutional footing.

Third and finally, Congress should expand existing disclosure and disclaimer requirements to ensure voters know when the electoral content they are receiving has been materially altered, created, or otherwise influenced by the use of AI. If the fundamental concern with AI is that voters may not even know when AI was used to make the ads they are seeing or hearing, a legal mandate that the use of AI to materially influence electoral content be disclosed would at least ensure that voters can treat such content with the appropriate level of skepticism. Voters have a right to know who is speaking through an election ad, which is why existing disclaimer laws require candidates to "stand by" their message. Similarly, an AI disclaimer requirement would require an ad's sponsors to "stand by" their use of AI, heightening the public's ability to decide for themselves whether the ad can be relied on to influence their decision-making.

The private sector is already taking steps in this direction: Alphabet, the parent company of Google and YouTube, recently announced a new policy that paid political advertisements must explicitly disclose when ads contain "synthetic content that inauthentically depicts real or realistic-looking people or events." However, the rules of the road for AI cannot be left up to private companies. For starters, such an approach is neither uniform nor comprehensive; Alphabet's policies have no bearing on other social media platforms like Facebook or X (formerly known as Twitter), and even Alphabet's new policy apparently does not apply to unpaid content uploaded to YouTube. It also goes without saying that private companies may have their own interests; Alphabet itself has a subsidiary, DeepMind, working on generative AI. A company's voluntary policies regarding the disclosure of AI could change at any time, regardless of whether the underlying threat to our elections remains. Federal policymakers need to act.

An AI disclaimer law could be tailored to the level of concern regarding the technology's use. A basic disclaimer could simply state that AI had been used in the ad, whereas a more detailed disclaimer might require additional information. This could include how AI was used, what it was used to depict, alter, or imitate, and more. The scope of the required disclosure could be tailored to best address varied concerns by giving voters the necessary information to evaluate an ad's credibility and reliability.

These approaches are not mutually exclusive. Congress should consider each and might very well decide to use a combination of tools at its disposal. For the most pernicious forms of AI-based fraud and manipulation, a prohibition would seem most appropriate. Whereas for other material forms of AI use, a disclaimer would be sufficient to put voters on notice that the content before them is AI-based.

In addition, Congress should carefully consider how each of these policies could be enforced. The FEC's well-established procedures for the civil enforcement of campaign finance laws, though in need of fundamental reform, could be applied to

the use of AI in elections. The unique challenges of AI might also require additional enforcement mechanisms, including establishing criminal penalties enforceable by the Department of Justice or creating a private right of action for candidates targeted by AI to seek rapid relief in federal court.

These recommendations are also not meant to be exhaustive. In light of the rapid acceleration and evolution of AI, Congress should continue studying how these technologies are used in the coming years, particularly in our elections. A one-and-done solution to the problems AI presents is unlikely to remain relevant for long, and at stake is the very fabric of our democracy. This problem merits continued vigilance.

Conclusion

Today's hearing is an encouraging sign that Congress is proactively working to address the impact of AI on our electoral process. The concerns raised by this technology are real and growing, and I strongly urge this Committee to approach these challenges without regard for partisanship or political gain. If left unregulated, AI will increase the risk of misinformation, deceptive advertisements, and distrust for candidates and voters on both sides of the aisle. If appropriately safeguarded, we may yet enjoy the benefits of this technology alongside a stronger democracy.

Thank you again for the invitation to testify. I look forward to answering your questions.

The Leadership Conference
on Civil and Human Rights

1620 L Street, NW
Suite 1100
Washington, DC
20036

202.466.3311 voice
202.466.3435 fax
www.civilrights.org

STATEMENT OF MAYA WILEY, PRESIDENT AND CEO
THE LEADERSHIP CONFERENCE ON CIVIL AND HUMAN RIGHTS

UNITED STATES SENATE
COMMITTEE ON RULES AND ADMINISTRATION

HEARING ON "AI AND THE FUTURE OF OUR ELECTIONS"

September 27, 2023

I. Introduction

Chair Klobuchar, Ranking Member Fischer, and members of the committee: Thank you for the opportunity to testify today. My name is Maya Wiley, and I am the president and CEO of The Leadership Conference on Civil and Human Rights, the nation's oldest and largest civil rights coalition with a diverse membership of more than 240 national organizations working to build an America as good as its ideals. Our coalition is dedicated to democracy, which requires promoting and protecting the civil and human rights of every person in the United States. Our work is nonpartisan, from the fight for voting rights and every bipartisan reauthorization of the Voting Rights Act of 1965 to the continued efforts to ensure voters have access to the ballot and the ability to cast their ballots without interference.

To support fair and non-discriminatory elections, we also work to ensure voters know how to register, where to cast their ballots, and how engage in the civic process as poll workers. With our members and partners, we monitor election cycles for misinformation and disinformation to ensure that voters are not misled, deceived, or threatened by others. We actively engage social media platforms in the types of policies and practices that prevent and reduce hate, bias, and deception online to ensure that our democracy works for all of us.

We have been working at the intersection of civil rights and technology for more than a decade, and we have long been leading the fight against online hate, bias, and mis- and disinformation. We recently announced the establishment of a first-of-its-kind Center for Civil Rights and Technology. This Center will expand and deepen our long-standing work on media and technology, and it will actively work with our coalition, advocates, academics, policymakers, and the private sector on the positive possibilities of artificial intelligence (AI). This work will help to ensure a regulatory structure, transparency, and accountability along with other guardrails necessary to ensure that AI supports a country where all people and communities reap the benefits of, rather than suffer harm from, rapidly transforming technology.

Today, I will share two observations from our coalitional work on free and fair elections: First, technology is a tool, but we must have checks on its intentional misuse, consistent with Constitutional

protections; and second, we must ensure fair elections, free from bias and discrimination, which reinforces what we know to be true: We are fortunate to have our democracy.

Sixty years ago, in August of 1963, more than 200,000 people gathered in Washington, D.C., for the March on Washington for Jobs and Freedom. The Leadership Conference was there with our founders. From the steps of the Lincoln Memorial, one of our founders — union leader A. Philip Randolph, president of the Brotherhood of Sleeping Car Porters — said that "The March on Washington is not the climax of our struggle, but a new beginning not only for the Negro but for all Americans who thirst for a better life." This remains true today. And now, it is technology that poses a potential opportunity, but also a rapidly evolving set of risks. This includes the poisonous bloom of disinformation and hate that can threaten free and fair elections.

Our democracy depends on trust — trust in the integrity of our elections, trust in the information we receive about our elections, and trust that when we participate, we can do so safely. Democracy also depends on trust in our leaders — trust that they will be truthful about the facts, even when rational minds can differ about how to interpret those facts. Mis- and disinformation are clear threats, and their spread has already threatened the peaceful transfer of power after our 2020 presidential election.

In 2018, many of us watched a YouTube video, purportedly of Barack Obama, discussing the ability of an "enemy" using technology to make it look like someone was saying something they never said.[1] About halfway through the video, the screen splits and we see that the voice is not from the former president but from the actor, comedian, and filmmaker Jordan Peele imitating Barack Obama. This educational video gave us a glimpse of the future of mis- and disinformation in the form of deepfakes and their ability to deceive our people and undermine our national security. Artificial intelligence is not new, but it is rapidly developing thanks to breakthroughs in generative technology. In the 2020 election, we witnessed little in the way of deepfakes when compared to today — but it was a risk even then.

Even before the rapid advances we are seeing in AI and the tools they enable for the rapid spread of disinformation, we had disinformation intentionally targeting communities of color.[2] As the Associated Press reported, in 2020, Facebook ads targeting Latino and Asian communities in Chicago falsely claimed candidate Joe Biden was a communist. A doctored photo was altered from the true image of a dog urinating on a Biden campaign poster to a fake image of the dog urinating on a Trump campaign poster.[3] The Senate Intelligence Committee found that Russia-backed operatives aggressively targeted disinformation toward African Americans during the 2016 election to suppress votes, stating, "[N]o single

---

[1] BuzzFeed News, "You Won't Believe What Obama Says In This Video!," YouTube (2018), https://www.youtube.com/watch?v=cQ54GDm1eL0.
[2] Christine Fernando, "Election Disinformation Targeted Voters of Color in 2020. Experts Expect 2024 to be Worse.," Associated Press (July 29, 2023), https://apnews.com/article/elections-voting-misinformation-race-immigration-712a5c5a9b72c1668b8c9b1eb6e0038a.
[3] Jordan Liles, "Photo Shows Dog Peeing on Sign for Biden, Not Trump," Snopes (Oct. 21, 2022), https://www.snopes.com/fact-check/dog-peeing-trump-sign/.

group of Americans was targeted by IRA information operatives more than African Americans." We have seen enough to know we must be prepared for what is upon us.[4]

II. The Origins of Artificial Intelligence on Our Democracy

During the 2020 and 2022 election cycles, The Leadership Conference and our partner Common Cause monitored, analyzed, and responded to mis- and disinformation. Most of the content that we flagged falls into the narratives and trends that we have seen over the last two years, including the Big Lie — the phrase used to describe false claims that Joe Biden did not win the 2020 election — and false information about elections processes, such as mail-in ballots and the use of ballot drop boxes. We used the flagged content as evidence to inform the platforms and the government of election disinformation trends/issues in private conversations and in public letters and statements. What we have found is also supported by academic research, including:

- Online mis- and disinformation continues to confuse, intimidate, and harass people, including voters; suppress the right to vote; spread hate speech; and otherwise disrupt our democracy.

- Vulnerable communities and communities of color are disproportionately threatened. They are often the intended recipients of mis- and disinformation intended to suppress the vote and drive wedges betwteen communities of color. They are also impacted online and offline by mis- and disinformation about their communities, including content that stirs up hate and distrust of these communities within other demographics.

- One example of the ramifications of online disinformation includes:

    - After an election official in Colorado (who was an outspoken election denier) allegedly assisted in the posting of a video showing passwords used to access the county's voting system that was secretly recorded during a security update, election workers statewide started receiving threats and resorted to wearing bulletproof vests and undergoing active shooter preparedness training in response to the possibility of right-wing violence.[5]

Fast-evolving technologies literally enable users to put words and actions into the mouths and bodies of real people that they never uttered or made.[6] Since 2017,[7] the use of "deepfakes" have become all too

---

[4] Report of the Select Committee on Intelligence of the United States Senate on Russian Active Measures Campaigns and Interference in the 2016 U.S. Election, Volume 2: Russia's Use of Social Media with Additional Views, 116th Cong. (2020) https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf.
[5] Laura Romero, "Colorado Voting Officials Adopt Safety Measures as State Becomes Target for Election Conspiracists," ABC News (May 9, 2022), https://abcnews.go.com/US/colorado-voting-officials-adopt-safety-measures-state-target/story?id=84546801.
[6] Britt Paris and Joan Donovan, *Deepfakes and Cheap Fakes: The Manipulation of Audio and Visual Evidence*, Data & Society (2019) https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1.pdf (Deepfakes and Cheap Fakes).
[7] *Id* at 2.

common, where the faces of people, largely women,[8] have been attached to the bodies of others in pornographic material, destroying reputations and even being used for blackmail and human trafficking.[9] While "cheapfakes," such as photoshopping images and other forms of visual manipulation, are not new, AI is advancing the opportunities for intentional manipulation. Data and Society, an independent nonprofit research institute, recently stated in a report on deepfakes: "With thousands of images of many of us online, in the cloud, and on our devices, anyone with a public social media profile is fair game to be faked."[10]

Deepfakes are already being used in national, state, and local races by supporters of candidates and political action committees, according to some researchers.[11] Just this past summer, Never Back Down, a group that supports Florida Governor Ron DeSantis' candidacy for president, ran an ad in Iowa with a voice, generated by AI, which sounded like Donald Trump's voice saying words he wrote on social media but never spoke.[12]

This is just one example, and there are many other national and local examples that have been widely reported, including:

- In a Chicago mayoral race, an X account called "Chicago Lakefront News" posted a video that appeared to show candidate Paul Vallas saying, "in my day," a police officer could kill as many as 17-18 civilians and "no one would bat an eye." This video was entirely AI-generated, although it appeared authentic.[13]
- A deepfake also circulated depicting Sen. Elizabeth Warren (D. Mass.) insisting that Republicans should be barred from voting in 2024.[14]
- Earlier this year, an AI-generated video showing President Biden declaring a national draft to aid Ukraine's war effort — initially acknowledged as a deepfake but later stripped of that context — led to a misleading tweet that garnered more than 8 million views.[15]

---

[8] Suzie Dunn, "Women, Not Politicians, Are Most Often Target By Deepfake Videos," Centre for International Governance Innovation (Mar. 3, 2021), https://www.cigionline.org/articles/women-not-politicians-are-targeted-most-often-deepfake-videos/.
[9] Public Service Announcement, "Malicious Actors Manipulating Photos and Videos to Create Explicit Content and Sextoration Schemes," Federal Bureau of Investigation (June 5, 2023), https://www.ic3.gov/Media/Y2023/PSA230605.
[10] Deepfakes and Cheap Fakes at 7.
[11] Ayesha Rascoe, "How Real is the Threat of AI Deepfakes in the 2024 Election," NPR (July 30, 2023), https://www.npr.org/2023/07/30/1190970436/how-real-is-the-threat-of-ai-deepfakes-in-the-2024-election.
[12] Louis Jacobson and Loreben Tuquero, "A Pro-Ron DeSantis Ad Used AI to Recreate Donald Trump's Voice. Experts Say It Won't Be The Last.," Poynter (July 20, 2023). https://www.poynter.org/fact-checking/2023/desantis-never-back-down-pac-ai-ad-trump-voice/.
[13] Megan Hickey, "Vallas Campaign Condemns Deepfake Video Posted to Twitter," CBS News (Feb. 27, 2023), https://www.cbsnews.com/chicago/news/vallas-campaign-deepfake-video/.
[14] Rob Lever, "Deepfake Video of Elizabeth Warren Spreads on TikTok," AFP Factcheck (Mar. 6, 2023), https://factcheck.afp.com/doc.afp.com.33AE6KT.
[15] Nur Ibrahim, "Did Biden Call for a Military National Draft?," Snopes (Mar. 2, 2023), https://www.snopes.com/fact-check/biden-military-national-draft/.

The rapid growth of AI since the 2022 elections has the potential to significantly increase the volume and speed of disinformation in the 2024 election cycle. The false narratives we have seen in prior elections are certain to come up again in 2024 in new forms with even more intensity. AI has the potential to turbocharge the volume and the spread of disinformation. Generative AI can increase the ability and scale to spread false information and propaganda, leaving voters confused and further questioning what they see and hear. Our coalition member, The Brennan Center, has pointed out that elections are particularly vulnerable to AI disinformation. Generative AI tools are most effective when they produce content similar to their current databases. Since the same false election narratives will likely come up again in future elections, there is an abundance of past election disinformation in the training data underlying generative AI tools that can make them a time bomb for future election disinformation.[16] The proliferation of AI-generated content could accelerate the loss of trust in the integrity and security of our overall election system and dramatically interfere with the right to vote.[17]

AI can spread disinformation, particularly toward communities of color, in a number of ways, including deepfaked audio and visuals. Chatboxes (e.g., ChatGPT) can spread false narratives further. AI could send false or deceptive comments from fake constituents or advocates or set up fake news sites. Chatboxes and deepfakes could threaten election systems through targeting and phishing efforts that are more targeted and personalized to users, making them potentially more effective.[18] Disinformation spreaders can plant false information on websites that generative AI can pick up and spread. While a benefit of AI is its support for non-English speakers to have translation tools, non-English speakers have been targeted by disinformation in elections.

Unfounded conspiracy theories are causing some states to withdraw from the Election Registration Information Center despite its bipartisan origins and the bipartisan praise it has received for supporting election integrity, including the ability to better identify deceased voters.[19]

Arizona, for example, attempted to provide more access to voters' personal information, making targeting with powerful AI more effective and raising the threat that information will be abused through AI targeting.

This is just one of many examples of where comprehensive reform and regulation could allow for effective AI use, but underscoring a need for regulation so that use is not abused.

III. Eroding Voting Rights Creates Conditions for Disinformation to Thrive

---

[16] Mekela Panditharatne and Noah Giansiracusa, "How AI Puts Elections at Risk - And the Safeguards Needed," Brennan Center for Justice (July 21, 2023), https://www.brennancenter.org/our-work/analysis-opinion/how-ai-puts-elections-risk-and-needed-safeguards (Panditharatne and Giansciracusa).
[17] *Id.*
[18] *Id.*
[19] Ben Paviour and Miles Parks, "Virginia Becomes the Latest GOP-Governed State to Quit a Voter Data Partnership," NPR (May 11, 2023), https://www.npr.org/2023/05/11/1175662382/virginia-eric-withdrawal.

Since our independence from England, the nation has spent more than 200 years, through protest and too often in the face of threats and violence, to make our democracy real for all citizens — from white men who did not have the wealth to own land until the 1820s, to the Suffragette movement and Native American enfranchisement in 1920, to Bloody Sunday and the Voting Rights Act of 1965 that, after 100 years of blatantly undermining the Black vote, finally provided meaningful protections for lawful voting without discrimination. The attacks on those hard-won protections have created the conditions in which mis- and disinformation are more easily cooked up and consumed by the American public.

Critical to those protections was the bipartisan Voting Rights Act and all its reauthorizations from 1965 through 2006, which included the clear support of four Republican presidents. The Voting Rights Act built trust in our democracy. After successive demands for a more inclusive democracy, it was not until the second half of the 20th century that we finally achieved the right to democratic inclusion, which did not come easily. When we did achieve democratic inclusion, it was bipartisan. Until 2010, we had roughly four decades of a relatively trustworthy and trusted democratic process. It has never been perfect, but the integrity of our voting systems, the trust we shared across our beliefs and our diversity, and the fact that — for the most part — ours is a trustworthy system, has been the envy of much of the world.

That began to change in 2010.[20] The historic turnout of voters, including Black voters and young voters of all races in the 2008 presidential race, highlighted the significant power that coalitions of voters brought to the polls and our democracy. Backed by deep pockets including many corporations, state legislation was introduced across the country that would create barriers to the ballot box for lawful voters.[21] Voter identification laws were becoming the most prevalent form of barriers to the ballot. The alleged justification for these restrictive laws was voter fraud,[22] often wrapped in fear mongering about undocumented immigrants. In some instances, party operatives openly stated that the laws were designed to make it harder for voters who were likely to vote for the opposing party.[23] Black voters in particular were targeted.[24] The claims were not only unsupported, but research over the past two decades makes clear that voter fraud has not been a significant problem.[25]

As noted in a recent report from The Leadership Conference Education Fund, in the 10 years since the U.S. Supreme Court's *Shelby County v. Holder* ruling, state and local jurisdictions have used the absence of federal voting protections to try to take us backward by creating barriers to the ballot for Black, Brown,

---

[20] Ethan Magoc, "Flurry of Voter ID Laws Tied to Conservative Group ALEC," NBC News (Aug. 21, 2012), https://www.nbcnews.com/news/investigations/flurry-voter-id-laws-tied-conservative-group-alec-flna955652.
[21] Molly Jackman, "ALEC's Influence Over Lawmaking in State Legislatures," Brookings (Dec. 6, 2013), https://www.brookings.edu/articles/alecs-influence-over-lawmaking-in-state-legislatures/.
[22] Sarah Childress, "'Unprecedented' Number of Restrictive Voting Laws Being Introduced," PBS (May 31, 2012), https://www.pbs.org/wgbh/frontline/article/unprecedented-number-of-restrictive-voting-laws-being-introduced/.
[23] Andrew Cohen, "No More Pretending: Republicans Admit Vote Restrictions Are All About Winning," Washington Monthly (Mar. 26, 2021), https://washingtonmonthly.com/2021/03/26/no-more-pretending-republicans-admit-vote-restrictions-are-all-about-winning/.
[24] Arturo Garcia, "Did North Carolina Admit to Target Black Voters with a 'Voter ID Law?,'" Snopes (Oct. 22, 2018), https://www.snopes.com/fact-check/north-carolina-voter-id/.
[25] "What Research Tells Us About Voter Fraud," Cornell University (accessed Sept. 25, 2023), https://evidencebasedliving.human.cornell.edu/blog/what-research-tells-us-about-voter-fraud.

and Native voters; people with disabilities; young and older people; and new Americans.[26] This has created more fuel to fire mis- and disinformation. Well-funded state-by-state strategies — erected to create unnecessary barriers to the ballot — have created a tidal erosion of trust in the integrity of our elections and have done so by driving divisions and stoking fear in some instances that Mexican immigrants in particular — and Latino voters in general — were voting unlawfully. One of the central legal protections used to protect voters from partisan power grabs was the Voting Rights Act of 1965 and all its bipartisan reauthorizations. Years after the 2008 election, fake news about millions of undocumented immigrants voting in 2008 appeared on a Fox and Friends tweet.[27]

These unfounded statements supported attacks on vote by mail and stricter voter identification laws, voter roll purges, and the creation of voting police — intimidating marginalized voters or making it less likely they would vote. Older Americans, students, and people with disabilities saw improvement in their ability to have access to voting with mail-in ballots, early voting, and more polling sites, which were crucial for their ability to participate in our democracy.

The frontal and explicit attacks on the free and fair election of 2020 — and the false claims that President Biden is not a duly elected leader of this free republic — have resulted in unprecedented violence, and they have only served to deepen divisions and stoke everything from threats to harassment to outright violence against public servants and ordinary people, as well as elected leaders.[28]

State laws that suppress the vote and fuel distrust and division have continued to grow, resulting in an astounding number of state voter restrictions in recent years. State legislatures in 20 states enacted at least 33 new restrictions in effect for the 2022 midterms.[29]And already this year, at least 11 states have enacted 13 more restrictive voting laws.[30] As in 2010, this push to restrict voting access gained new intensity after voters of color made their voices heard through robust turnout in 2020. In a dangerous affront to basic concepts of a constitutional democracy, 26 states created or expanded existing criminal penalties related to voting. In total, 60 new felonies and 50 new misdemeanors were enacted after the 2020 election, including criminal penalties for assisting voters in some instances.[31] It amounts to state-driven voter intimidation, the likes of which we have not seen since the crush of state violence to prevent newly recognized Black citizens from voting and taking up public office after the end of the Civil War.

---

[26] *Ten Years After Shelby County v. Holder: Charting the Path Forward for Our Democracy*, The Leadership Conference Education Fund (June 30, 2023) https://civilrights.org/resource/ten-years-after-shelby-county-v-holder-charting-the-path-forward-for-our-democracy/.
[27] Alex Kasprak, "Did 5.7 Million 'Illigal Immigrants' Vote in the 2008 U.S. Election?," Snopes (Jun. 23, 2017), https://www.snopes.com/fact-check/illegal-immigrants-2008-election/.
[28] Reuters Fact Check, "Fact Check-Re-examining How and Why Voter Fraud is Exceedingly Rare in the U.S. Ahead of the 2022 Midterms," Reuters (June 2, 2022), https://www.reuters.com/article/factcheck-fraud-elections-idUSL1N2XP2AI.
[29] *Voting Laws Roundup: October 2022*, Brennan Center for Justice (Oct. 2022), https://www.brennancenter.org/our-work/research-reports/voting-laws-roundup-october-2022.
[30] *Voting Laws Roundup: June 2023*, Brennan Center for Justice (June 2023), https://www.brennancenter.org/our-work/research-reports/voting-laws-roundup-october-2022.
[31] Kira Lerner, "Criminalizing the Vote: GOP-Led States Enacted 102 New Election Penalties After 2020," Governing (July 17, 2022), https://www.governing.com/now/criminalizing-the-vote-gop-led-states-enacted-102-new-election-penalties-after-2020.

IV. Social Media Platforms Must Do More

Foreign governments, crime rings, organizations, and malicious individuals have actively created and disseminated baseless conspiracy theories about the integrity of our elections. They have spread dangerous, debunked, and factually unsupported claims about one of the deadliest pandemics in our history. This toxic content has caused immense harm — including, in some instances, acts of violence.

Social media platforms have not done their part. Unfounded claims of election fraud spread by social media platforms, according to researchers at the Stern School of Business at New York University,[32] amplified false claims about election denial, including on Meta, YouTube, X, and TikTok. Despite violations of their policies, these platforms either failed to take sufficient action to protect users of their platforms through potential corrective actions — such as increased labeling, demotion of content, or limiting the sharing of content — or allowed proponents of disinformation to exploit long-form videos.

As we have seen, this disinformation propelled the horrific acts of violence on our Capitol and our constitutional process for certifying Electoral College votes on January 6, 2021. But it didn't stop there. Both organized extremists and distrusting and dangerous individuals have been empowered and incited by unfounded claims that some elected officials, business leaders, and celebrities have elevated and endorsed, which has deepened distrust and stoked the spread of hate, harassment, and harm — both in and between election cycles.[33]

There have been online and real-life attacks on election officials and poll workers, which threaten the operations and functioning of our election systems and intimidate the very people to whom we should extend our gratitude and our protection.[34] Disinformation from the "Big Lie" about voter fraud and intentional efforts to mislead voters on where and how to vote continue as drivers of threats and undermine our democratic practice of lawful voting. False information about the "Big Lie" is still spreading rapidly on social media, and it is the basis for the forthcoming spread of disinformation about the 2024 elections by election deniers and high-profile users.

The Leadership Conference and our coalition have repeatedly urged social media platforms to take immediate steps to curb the spread of voting disinformation and hate speech in the midterms and future elections to protect the health of our democracy. In May and October of 2022, our coalition sent letters that called on major platforms to take several affirmative steps well in advance of the midterm elections

---

[32] Paul M. Barrett, *Spreading the Big Lie: How Social Media Sites Have Amplified False Claims of U.S. Election Fraud,* NYU Stern Center for Business and Human Rights (Sept. 2022), https://bhr.stern.nyu.edu/tech-big-lie.
[33] Lindsay Whitehurst and Christina Cassidy, "Election Workers Are Being Bombarded With Death Threats, the U.S. Government Says," PBS (Aug. 31, 2023), https://www.pbs.org/newshour/politics/election-workers-are-being-bombarded-with-death-threats-the-u-s-government-says.
[34] Jacob Knutson, "Over 140 Threats Against Election Workers Recorded in Maricopa County," Axios (Nov. 6, 2022), https://www.axios.com/2022/11/06/arizona-maricopa-county-threats-against-election-workers.

to combat election disinformation.[35] These include consistently enforcing existing policies, addressing election disinformation continuously, and taking action against non-English disinformation.

We also recently urged YouTube to reverse its recent decision to allow false election claims about the outcome of the 2020 election on its platform and reinstate the policy that prohibits 2020 election denial content on the platform.[36]

However, platforms did not and have not taken meaningful steps on the actions we have requested. We have seen:

- Lack of enforcement of current voting disinformation policies
- Reduction or elimination of trust and safety teams charged with monitoring content
    - YouTube shed two of its five policy experts who worked on hate speech and harassment issues, leaving only one person in charge of misinformation worldwide.[37]
    - According to an interview with the BBC in April, Elon Musk claimed that X had only about 1,500 employees, about an 80 percent drop from prior to his takeover.[38]
    - X completely eliminated the team that oversaw disinformation and trust and safety issues.[39] Sadly, X is not alone. Meta has cut 21,000 jobs over the last nine months, including deep cuts to trust and safety teams.[40]

Most major platforms' enforcement of their own disinformation policies continues to be erratic and inconsistent, particularly against high-profile users, leading to false claims about voting processes and attempts to overturn the certified results of our elections. Platforms are actually loosening restrictions,

---

[35] Leadership Conference and Democracy Groups Urge Social Media Platforms to Address Voter Disinformation Ahead of Midterms, The Leadership Conference on Civil and Human Rights (Nov. 9, 2022), https://civilrights.org/resource/leadership-conference-and-democracy-groups-urge-social-media-platforms-to-address-voter-disinformation-ahead-of-midterms-2/; The Leadership Conference and 120 Civil Rights & Democracy Groups Urge Social Media Platforms to Take Meaningful Steps to Address Election Disinformation, The Leadership Conference on Civil and Human Rights (May 12, 2022), https://civilrights.org/resource/the-leadership-conference-and-120-civil-rights-democracy-groups-urge-social-media-platforms-to-take-meaningful-steps-to-address-election-disinformation/.

[36] Civil Rights Concerns Over YouTube Decision to Allow Election Disinformation, The Leadership Conference on Civil and Human Rights (June 26, 2023), https://civilrights.org/resource/civil-rights-concerns-over-youtube-decision-to-allow-election-disinformation/

[37] https://www.nytimes.com/2023/02/14/technology/disinformation-moderation-social-media.html

[38] Mike Wendling, "Musk Claims Zapping Bots Will Stop False Information," BBC (Apr. 12, 2023), https://www.bbc.com/news/live/world-us-canada-65247272/page/2.

[39] Ben Collins, Brandy Zadronzy, and David Ingram, "Days Before the Midterm, Twitter Lays Off Employees Who Fight Misinformation," NBC News (Nov. 4, 2022), https://www.nbcnews.com/tech/misinformation/twitter-fires-employees-fight-misinformation-midterm-elections-rcna55750.

[40] Hayden Field and Jonathan Vanian, "Tech Layoffs Ravage the Teams That Fight Online Misinformation and Hate Speech," CNBC (May 27, 2023), https://www.cnbc.com/2023/05/26/tech-companies-are-laying-off-their-ethics-and-safety-teams-.html.

such as YouTube now allowing content about the Big Lie to be posted at a time when platforms should be stepping up to detect and address AI-generated disinformation about voting.[41]

Despite the policies these platforms claim to have, we have tracked a systematic failure to enforce them. Despite these platforms' immense profits and user bases — nearly 3 billion users for Facebook, more than 2.5 billion users for YouTube, and nearly 1.5 billion users for Instagram — public reporting shows that the platforms are cutting staff whose job is to help protect the public from dangerous content. A major social media platform simply cannot responsibly apply its policies with only one person responsible for political misinformation and two for medical misinformation.

The threat of election disinformation is still prevalent and far from over. There were disinformation narratives that took hold in 2022, particularly in contested states and where the counting of results took several days. We particularly saw this in Arizona and Nevada: While there was progress in addressing voting disinformation in 2022, election denialism will likely continue to be a persistent fixture in future campaigns. Disinformation around certification, counting results, and harassment of election officials will likely continue.

V. Potential Steps

The Leadership Conference and our coalition have been actively advocating for Congress and the administration to take action to ensure sufficient safety, transparency, accountability, and protections to ensure AI delivers on the possibilities it can offer to promote more fairness and equity while protecting privacy, civil rights, and our democracy. Addressing more significant issues of discrimination and bias through an accountability framework will go a long way toward addressing disinformation. In addition to addressing mis- and disinformation, we believe Congress should enact broader AI protections that ensure:

- Safety and effective systems are developed with diverse stakeholders to ensure that risks, concerns, and impacts are considered and addressed.
- Algorithms must be free of discrimination with proactive and continuous measures — that are transparent and trustworthy — to protect the public from discrimination.
- Consumers should be protected from privacy violations in the design and based on defaults in systems.
- The public must know AI is being used and understand its impacts.
- People should be able to opt out of AI and get access to a human being.

The REAL Political Advertisements Act and the Protect Elections from Deceptive AI Act are essential steps in the right direction. Transparency into whether AI has been used in ads will help to keep voters informed, and a prohibition on the distribution of materially deceptive audio or visual media by political campaigns will help ensure the integrity of our democracy. The REAL Political Advertisements Act is a commonsense step forward in addressing some of the problems caused by using AI in political ads and the risks that come with it. We also view the recently introduced bipartisan legislation to ban deceptive

---

[41] Blog, "An Update to Our Approach to US Election Misinformation," YouTube (June 2, 2023), https://blog.youtube/inside-youtube/us-election-misinformation-update-2023/.

AI-generated content in elections as a promising start to ensuring our democracy is protected against tools promoting false or fraudulent information.

As we enter the 2024 election cycle, it is paramount to keep pressure on the platforms to better address voting disinformation and to push for additional, continued oversight of the platforms:

- Besides consistent enforcement of their rules, platforms must do a better job of addressing non-English speaking disinformation, mainly Spanish and Asian language disinformation, particularly on video-focused platforms such as YouTube and TikTok.
- We have continually pushed the platforms to provide more data on questionable content regarding voting/elections (as well as civil rights and hate and bias issues) so that we can work together with the platforms to utilize that data and find more solutions to address disinformation.
- As generative AI advances, the potential for rampant false content and the rapid spread of disinformation is alarming. Platforms and tech companies must have policies, systems, and guardrails to stem the disinformation resulting from generative AI.
    - The Brennan Center has noted that platforms can devote more resources to identifying and removing coordinated bots and labeling deepfakes that could influence elections.[42]
- The Brennan Center has also suggested that the Cybersecurity and Infrastructure Security Agency should create and share resources to help election offices address disinformation campaigns that exploit deepfake tools and language models to undermine election processes;[43] and, to reduce the risk of AI misuse by political campaigns, the Federal Election Commission should ensure that its political ad disclosure requirements cover the full range of online communications currently permitted under federal law.[44]
- As Public Knowledge has noted, a whole of society approach is needed to address the issues presented by generative AI and restore trust in our information environment. This can include policymakers creating incentives for the technology platforms to change their policies and product design, and they should foster more competition and choice among media outlets. Civil society should convene stakeholders, including from the communities most impacted by misinformation, to research and design while protecting privacy and freedom of expression.[45]

Congress must use its oversight powers and require the platforms to provide more data and solutions on voting/election disinformation and AI. Disinformation, sometimes driven intentionally by foreign governments in our election cycles, often targets Black and Latino communities and poses significant risks to our society.

To that end, Congress should:

---

[42] Panditharatne and Giansciracusa.
[43] *Id.*
[44] *Id.*
[45] Lisa Macpherson, "Lies, Damn Lies, and Generative Artifical Intelligence: How GAI Automates Disinformation and What We Should Do About It," Public Knowledge (Aug. 7, 2023), https://publicknowledge.org/lies-damn-lies-and-generative-artificial-intelligence-how-gai-automates-disinformation-and-what-we-should-do-about-it/.

- Ensure coordinated follow-through by federal agencies using congressional oversight.
- Launch sustained public engagement with diverse stakeholders through hearings and other fora.
- Continue oversight hearings and deliberations on AI legislation with a focus on addressing the issues surrounding election disinformation and hate speech.
- Along with state legislatures, regulate AI to help identify AI-generated content and limit interference with elections.
- Direct DHS to develop training, tools, and resources to ensure election officials and administrators can detect, address, and prevent abuses of AI in the election context.
- Press DOJ to continue to aggressively enforce federal voting rights laws and ensure that voters of color are not targeted with disinformation through the use of AI to suppress their vote.

VI. Closing

If we poison our democratic soil with false statements about our own elections and fear that they are rigged, we create ground that truth dies in — but where propaganda, planted by those who seek to deceive, can become invasive using the tools technology provides. That means we must consider whether we are farmers or flame throwers when it comes to our democracy.

We stand ready to work to find solutions that will keep our democracy safe and stop the suppressive effect that AI-generated disinformation can have on civil rights and racial justice. Our civil rights and the integrity of our democracy are at stake.

Thank you for inviting me to testify today. I am pleased to answer any questions you may have.

TESTIMONY

The Center for Growth and Opportunity at Utah State University

The Integral Role of AI Tools in Modern Political Discourse

September 27, 2023

Testimony Before the United States Senate Committee on Rules and Administration
Hearing: AI and the Future of our Elections

Neil Chilson, Senior Research Fellow at The Center for Growth and Opportunity at Utah State University

Thank you, Chairwoman Klobuchar, Ranking Member Fischer, and committee members for the chance to speak today on artificial intelligence's influence on elections. I'm Neil Chilson, a senior fellow at the Center for Growth and Opportunity at Utah State University (CGO), a former Chief Technologist at the Federal Trade Commission, and a past advisor to acting FTC Chair Maureen K. Ohlhausen.

At CGO I study and promote the conditions that best enable technology to create widespread human prosperity and abundance.

Now is a crucial time for this mission, as the many technologies of artificial intelligence launch us toward prosperity and abundance.

Intelligence is our most vital resource. It is the tool with which humanity has overcome countless challenges. Every breakthrough technology, scientific discovery, business success, art piece, and literary work stems from human intellect. Often, this involves massive collaboration and coordination, generating intelligent outcomes beyond any one person's capabilities.

Tools amplifying human intelligence, therefore, have vast potential. AI promises to help humanity create a healthier, more productive, more artistic, more interesting, and more enjoyable world.

While powerful tools can be misused, generative AI tools seem unlikely to materially affect election results because political speech already uses AI tools and has for years. Generative AI *will* lower the cost—in time and money—to generate high-quality creative content. When costs decrease, but demand persists, these are conditions of abundance. We'll see more speech, including political speech. But we shouldn't expect a shift in the truth-to-lies ratio.

In fact, if a lie is halfway around the world before the truth gets its shoes on, generative AI is a rocket-powered running shoe. AI tools will enable real-time fact-checking, cheaper voter education, and messages tailored to voter needs. These tools can strengthen democracy.

I recommend this committee focus on the following as it explores how to respond to the use of computation and artificial intelligence in the creation of political speech.

- First, defining "AI" is hard. It encompasses many commonly used technologies and most political ads contain content "generated in whole or part by the use of artificial intelligence."[1]
- Second, existing technology-neutral laws and emerging private norms can tackle potential election manipulation.
- Third, generative AI's threat to elections may be less than other technologies.

AI can improve political discourse. There are risks, but they are not novel. AI should not be used as a pretext for limiting political speech. Congress should build on existing expertise and capabilities, not create new regulatory regimes.

AI-generated political speech is already here—and has been, for years

Defining AI is hard. Experts have debated the term "artificial intelligence" for decades. The widely-used AI textbook by Peter Norvig and Stuart Russell dedicates a whole section to defining AI.[2] They describe how the scope of AI is fluid, and has included different types of software and algorithms over the decades. Indeed, John McCarthy, who coined the term, remarked that once an AI algorithm works, "we stop calling it AI."[3]

Norvig and Russell identify four historical approaches to defining AI: acting humanly, thinking humanly, thinking rationally, and acting rationally. They emphasize "acting rationally" as the prevailing model, defining AI as the study and construction of agents that "do the right thing."[4] In other words, after much discussion, Norvig and Russell advocate a functional approach: categorize something as AI or not based, not on its design or nature, but on its uses and actions.

So what counts as AI today? There is no definitive answer, but it is an expansive category. Norvig and Russell provide an incomplete catalog of example AI applications, including such varying technologies as robotic vehicles, machine translation, speech recognition, recommendation algorithms, image understanding, game playing, and medical diagnosis.[5]

This hearing is focused on AI and the future of our elections. Let's first take a walk through the present production and delivery of a digital ad for a political campaign, and explore where artificial intelligence might intersect with ad creation.

Sarah, an ad account manager, has an idea for a new political ad. She fires up ChatGPT, asks it for ten variations on that idea, and further iterates on phrases for ad copy and script.[6] As she edits the script in Microsoft Word, text predictions speed her work. Once the script is finalized, she trans-

---

1 REAL Political Advertisements Act, S.1596, 118th Cong. §4(a) (adding 52 U.S.C. 30120(e)(1)).
2 Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. (London: Pearson 2021).
3 Moshe Vardi, "Artificial Intelligence: Past and Future," *Communications of the ACM* 55, no.1 (Jan. 2012): 5.
4 Russell and Norvig, *Artificial intelligence*, 3–4. The authors further note that this "right action" should align with human benefits; Vardi, "Artificial Intelligence: Past and Future," 5.
5 Russell and Norvig, *Artificial intelligence*, 28–30.
6 See ChatGPT, http://chat.openai.com; See also, "ChatGPT-4 - Brainstorming," W3 Schools, https://www.w3schools.com/gen_ai/chatgpt-4/chatgpt-4_brainstorming.php; "AI Advertising Script Generator," Taskade, https://www.taskade.com/generate/marketing/advertising-script.

lates it into Spanish and Hmong using Google Translate.[7] Her media team starts gathering assets. One team member easily searches a large database of stock photos, thanks to automatic tagging using a computer vision algorithm.[8] Another heads out to get some establishing shots in still and video. Her camera's physical light sensor has sophisticated embedded algorithms that adjust how the picture is turned into data.[9] Such algorithms even adapt depending on which lens she uses.[10] This helps the camera capture the most information from the scene and correct for known lens defects. Facial detection and eye detection keep the photo subjects in focus and properly exposed.[11] Electronic image stabilization algorithms constantly revise the raw data from the sensor to eliminate shaky video without expensive rigs.[12] Similarly, as she records audio, algorithms suppress background noise.[13]

The production team uses editing tools to refine and reassemble the content. Speech recognition tools sync transcripts to videos, enabling the editors to make text-based video edits.[14] One editor uploads multiple raw video and audio feeds into software that can identify who is speaking at any point and produce a final cut with appropriate closeups and wide-angle shots.[15] A producer removes background noise using algorithms that isolate and remove specific noises, such as filler words, traffic honking or crowds cheering, and then inserts an instant voiceover to one of the ads with text-to-speech.[16] The photos and videos are edited for sharpness, contrast, cropped, and cleaned up. The editors automatically replace overhead power lines in one shot with blue sky and to erase unplanned bystanders from a video.[17] They use similar tools to enhance actor appearances and to change the lighting in one scene.[18] These types of tools generate new content to replace the original content. Some long precede the existing wave of generative AI.

In post-production, the team realizes that Hindi speakers could be a key audience. Rather than film an entirely new version of the ad, they upload the finished ad to HeyGen, which translates the audio seamlessly with one click and modifies the video so that the actors appear to be speaking Hindi in "an authentic speaking style."[19]

---

7 "Translation AI," Google Cloud | Cloud Translation, https://cloud.google.com/translate.

8 iStock Staff, "How iStock Search Helps You Find the Best Possible Image," iStock, February 19, 2020, https://marketing.istockphoto.com/blog/how-istock-search-helps-you-find-the-best-possible-image/; "Visual Search Powered by Shutterstock.AI," Shutterstock, https://www.shutterstock.com/developers/solutions/computer-vision.

9 "Understanding Color Interpolation," Teledyne Flir, May 11, 2017, https://www.flir.com/support-center/iis/machine-vision/application-note/understanding-color-interpolation/; Ron Lowman, "How Cameras Use AI and Neural Network Image Processing," Synopsys, June 29, 2022, https://www.synopsys.com/blogs/chip-design/how-cameras-use-ai-neural-network-image-processing.html.

10 "In-camera Lens Corrections," Canon, https://www.canon-europe.com/pro/infobank/in-camera-lens-corrections/.

11 Matthew Saville, "Autofocus Technology is Changing, Here's Why It's Not Just Bells & Whistles Anymore," SLR Lounge, https://www.slrlounge.com/autofocus-technology-is-changing-heres-why-its-not-bells-whistles-anymore/. Some recent cameras even have animal and bird tracking. See, "Everything You Wanted to Know about Autofocus (AF)," Canon, https://www.canon-europe.com/pro/infobank/autofocus/.

12 TDK, "Electronic Image Stabilization," https://invensense.tdk.com/solutions/electronic-image-stabilization/. Apple's recent iPhone 15 announcement takes computational photography to a whole new level: the iPhone 15 uses multiple custom neural nets to process images. It's not an exaggeration to say that every photo taken by an iPhone 15 will be AI generated in part. Jaron Schneider, "Apple Explains What the iPhone 15 Camera Can and Can't Do – and Why," PetaPixel, https://petapixel.com/2023/09/18/apple-explains-what-the-iphone-can-and-cant-do-and-why/.

13 Christina Gwira, "10 Best AI Audio Tools in 2023 (For Podcasts, Music & More)," August 23, 2023, https://www.elegantthemes.com/blog/business/best-ai-audio-tools.

14 "Text-Based Editing in Premiere Pro," Adobe, September 13, 2023, https://helpx.adobe.com/premiere-pro/using/text-based-editing.html; Probhas Pokharel, "Video Text Editing: Easily and Accurately Edit Videos by Editing the Text," Reduct, July 2021, https://reduct.video/blog/video-text-editing.

15 NapSaga, "Revolutionize Your Video Podcast Editing with AutoPod AI," *Artificial Intelligence in Plain English*, Medium, April 27, 2023, https://ai.plainenglish.io/revolutionize-your-video-podcast-editing-with-autopod-ai-d961c71df602.

16 "Remove Background Noise from Video," VEED.IO, https://www.veed.io/tools/remove-background-noise-from-video.

17 "Remove Objects from Your Photos with Content-Aware Fill," Adobe, July 25, 2023, https://helpx.adobe.com/photoshop/using/content-aware-fill.html.

18 Joe Fedewa, "How to Adjust Photo Lighting with Google Photos on Pixel," How To Geek, October 22, 2023, https://www.howtogeek.com/696698/how-to-adjust-photo-lighting-with-google-photos/.

19 "Video Translate," HeyGen Labs, https://labs.heygen.com/video-translate.

But the ad's creation is only part of the journey. The team has sophisticated plans for delivering the ad to viewers. Personalized advertising algorithms help deliver the ads to the desired audience. As viewers interact with the ad, these algorithms adjust the delivery and even the content of the ads, dynamically testing out options and going with what achieves the most email sign-ups. As the ad is uploaded to different platforms, the team uses filters and other automated tools to easily tweak the content for each platform.

To sum up: AI is everywhere. Today's content creation, editing, and distribution rely heavily on AI technologies. Likely, every senator has employed one or more of these technologies in their ad campaigns.

In other words, AI-generated ads already exist. There are no principled distinctions between traditional technologies and newer "generative AI." The boundary is so indistinct that lawyers advising clients will struggle to definitively say that an ad lacks AI-generated content.

Given this ambiguity, any law mandating AI-content disclosures could mean *all* ads carry such notices, regardless of the tools used to create the ad and with no connection to the ad's truthfulness. Such ubiquitous disclosures would be useless to viewers. Worse, they would displace speakers' political speech, particularly in online ads, which are brief and small. Notably, the FEC took twelve years, up until December 2022, to adapt its own existing disclosures for paid online public communications due to these challenges.[20]

Technology-neutral laws and targeted private norms already exist

AI-generated political advertising is already here. So too are governance mechanisms. Existing technology-neutral laws and emerging norms in the private sector can help address concerns.

Any attempts to limit deception in political advertising should learn from the Federal Trade Commission's experience defending consumers from fraudulent business practices, including deceptive commercial advertising.[21] The FTC applies a straightforward standard for deceptive commercial advertising: an ad is deceptive if it contains a "representation, omission or practice that is likely to mislead the consumer acting reasonably in the circumstances, to the consumer's detriment."[22]

This standard is technology neutral, remaining consistent regardless of the tools used to create the deceptive content. The FTC evaluates outcomes rather than methods, keeping the standard relevant even as technologies evolve. This consistent standard has allowed the FTC to bring deception cases spanning from the internet's infancy to a recent 2023 case about misleading AI claims.[23]

Of course, the FTC polices commercial speech, which receives less First Amendment protection than political speech. The FTC's deception standard would be unconstitutional if applied to political advertising. However, it does have the virtue of being technology neutral. Similarly, speech that receives full First Amendment protection deserves that protection regardless of the technology used to create it.

---

20 "Commission adopts final rule on internet communications disclaimers and the definition of public communication," Federal Elections Commission | FEC Record Regulations, December 19, 2022, https://www.fec.gov/updates/commission-adopts-final-rule-internet-communications-disclaimers-and-definition-public-communication/.

21 "About the FTC," Federal Trade Commission, https://www.ftc.gov/about-ftc.

22 "Policy Statement on Deception," Federal Trade Commission, October 14, 1983, 2, https://www.ftc.gov/legal-library/browse/ftc-policy-statement-deception.

23 Compare, "Internet Site Agrees to Settle FTC Charges of Deceptively Collecting Personal Information in Agency's First Internet Privacy Case," Federal Trade Commission, Press Release, August 13, 1998, https://www.ftc.gov/news-events/news/press-releases/1998/08/internet-site-agrees-settle-ftc-charges-deceptively-collecting-personal-information-agencys-first; with, "FTC Action Stops Business Opportunity Scheme That Promised Its AI-Boosted Tools Would Power High Earnings Through Online Stores," Federal Trade Commission, Press Release, August 22, 2023, https://www.ftc.gov/news-events/news/press-releases/2023/08/ftc-action-stops-business-opportunity-scheme-promised-its-ai-boosted-tools-would-power-high-earnings.

Within the constraints of the Constitution, defamation law can help check deceptive political speech.[24] Malicious lies about others can incur severe financial consequences, irrespective of the creation or distribution method.[25] Like the FTC's approach, defamation law is technology neutral.

Private firms distributing ads are setting norms for generative AI in political advertising. For example, Google now requires advertisers to "prominently disclose when their ads contain synthetic content that inauthentically depicts real or realistic-looking people or events."[26] The rule targets deceptive uses, providing exemptions for content additions or manipulations that are inconsequential to the claims made in the ads.

The threat from generative AI tools to elections may be small relative to other threats

Assessing a new technology's potential impact on elections requires more than viewing it in isolation. One must consider alternatives that malicious actors could exploit for the same goals, as they always aim for the most impact at the least cost. The real question is whether the technology offers unique cost or efficiency advantages over other methods.[27]

People worry that generative AI will be used to manipulate elections. Yet voice imitation and media editing aren't new. Past misinformation and voter intimidation tactics have used simpler, more cost-effective means to pursue their goals. They do not typically involve candidate advertising at all. One doesn't need AI to create a deceptive text message or email. The real challenge often lies in distribution rather than content creation, and generative AI doesn't significantly alter this cost dynamic.

Brief thoughts on specific legislative proposals

S.1596 - REAL Political Advertisement Act. This bill isn't primarily about AI. Its main shift is amending the Federal Election Campaign Act of 1971 to increase regulation of internet communications. It wraps the Honest Ads Act in an AI cloak. The operative language would newly apply FEC regulation to online ads whether or not they contain AI-generated content, expanding the scope of FEC regulation.[28] This approach seems provoked by Russian-sponsored online ads during the 2016 presidential campaign, even though recent evidence suggests those ads did not sway the election.[29]

In addition, the bill would require disclaimers on ads containing "image or video footage that was generated in whole or in part with the use of artificial intelligence (generative AI)."[30] Given AI's widespread use in content creation, this could sweep in every political advertisement.

---

24 David Bauder, Randall Chase, and Geoff Mulvihill, "Fox, Dominion Reach $787M Settlement Over Election Claims," *AP News*, April 18, 2023, https://apnews.com/article/fox-news-dominion-lawsuit-trial-trump-2020-0ac71f75acfacc52ea80b3e747fb0afe.

25 Tom Hals and Jonathan Stempel, "Alex Jones Files for Bankruptcy Following $1.5 Billion Sandy Hook Verdicts," Reuters, December 2, 2022, https://www.reuters.com/world/us/alex-jones-files-bankruptcy-following-sandy-hook-verdict-court-filing-2022-12-02/.

26 "Updates to Political Content Policy," Google | Advertising Policies Help, September 2023, https://support.google.com/adspolicy/answer/13755910.

27 James R. Ostrowski, "Shallowfakes: The Danger of Exaggerating the AI Disinfo Threat," *The New Atlantis*, Spring 2023, https://www.thenewatlantis.com/publications/shallowfakes. "Chroniclers of disinformation often assume that because a tactic is hypothetically available to an attacker, the attacker is using it. But state-backed actors assigned to carry out influence operations face budgetary and time constraints like everyone else, and must maximize the influence they get for every dollar spent."

28 Compare, REAL Political Advertisements Act, S.1596, 118th Cong. §3 ("Expansion of Definition of Electioneering Communications"); with, Honest Ads Act, S.486, 118th Cong. §6 ("Expansion of Definition of Electioneering Communications").

29 Gregory Eady, et al., "Exposure to Russian Internet Research Agency Foreign Influence Campaign on Twitter in the 2016 Election and Its Relationship to Attitudes and Voting Behavior," *Nature Communications* 14, no. 62 (2023), https://www.nature.com/articles/s41467-022-35576-9.

30 S.1596, 118th Cong. §4(a) (adding 52 U.S.C. 30120(e)(1)).

S.2770 - Protect Elections from Deceptive AI Act. This proposed bill could chill every US individual's creation and sharing of AI-generated content on political issues. In one problematic provision, it labels as "deceptive" any AI-modified ad content that gives a reasonable person "a fundamentally different understanding or impression of the appearance, speech, or expressive conduct" as compared to the original content—even if the final product is truthful.[31]

Let me say that again: if a truthful ad gives a "different understanding or impression" than the ad's source content, this bill labels it deceptive, and potentially *per se* defamatory.

Given AI's prevalent use, most campaign ads could be subject to this bill. Even simply using modern video editing tools to quote an opponent might risk a defamation lawsuit under this bill.

There is also a jurisdictional issue. Defamation law is state law. Using a federal law to impose *per se* liability under state defamation law is a mechanism unknown to the American legal system.

As for the effect on the use of newer generative AI tools, this isn't an election protection act; it's an incumbent protection act. New generative AI tools allow smaller teams to produce high-quality content affordably. There's no proof such tools are more misleading than other content creation tools. This bill restricts affordable content creation, disadvantaging new candidates without significant funds. Costly traditional hand-editing would still be allowed.

If the bill's backers truly worry about altered political communications about candidates, why limit this to AI tools?

What Congress can do

Given the rapid evolution and broad scope of AI technologies and their uses, Congress needs to increase its ability to understand and respond. Hearings like today's are a first step in building the institutional knowledge necessary to create effective legislation. In addition, this committee should work to prepare the Federal Election Commission and other relevant agencies to monitor the use of new technologies throughout the election cycle and to assess the relevant effects.

More broadly, Congress should establish a permanent central source of advisory technical expertise on AI and algorithmic issues. The National Institute of Standards and Technology is likely a good fit. This expert body could provide technical education and advice to the many federal agencies that are grappling with applications of AI technology within their specific sectors.[32] While some have called for creating a new overarching AI regulator, supporting the emerging sector-specific approach to AI governance by offering a shared hub of technical expertise has many practical and legal benefits.[33]

---

31 S.2770, 118th Cong. §2(a) (adding to 52 U.S.C. 30101 *et seq.*, a new Sect. 325 (a)(2)).

32 See Adam Thierer, "Is AI Really an Unregulated Wild West?" Technology Liberation Front, https://techliberation.com/2023/06/22/is-ai-really-an-unregulated-wild-west/.

33 See, Matthew Mittelsteadt and Brent Skorup, "Comments Urging a Sectoral Approach to AI Accountability," Mercatus Center at George Mason University, https://www.mercatus.org/research/public-interest-comments/comments-urging-sectoral-approach-ai-accountability; Alex Engler, "A Comprehensive and Distributed Approach to AI Regulation," Brookings Institute, https://www.brookings.edu/articles/a-comprehensive-and-distributed-approach-to-ai-regulation/; Neil Chilson, "Does Big Tech Need Its Own Regulator?" The Global Antitrust Institute Report on the Digital Economy 21, SSRN, November 19, 2020, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3733726.

Conclusion

AI is hard to define. But AI technologies are already ubiquitous in digital advertising, not to mention daily life. Every industry relies on algorithms and AI. Its use seldom poses new problems. To the extent there are issues, they are generally handled by existing laws and norms.

We should be vigilant for novel issues raised by AI. But the problem of deceptive and misleading political ads is not novel, nor is it particularly connected to AI. Repackaging past proposals to control and censor political speech as "AI regulation" will not solve misinformation in ads and will chill political speech.[34]

AI technology can enhance human intellectual endeavors, including speech about important topics such as politics. If we merely use this new technology as an excuse to intervene, we will squander a substantial opportunity to strengthen democratic values, expand human prosperity, and move toward a world of abundance.

Thank you again for allowing me to share my views. I look forward to your questions.

34 Unfortunately, using the AI moment to advance unrelated agendas is a growing trend. See Neil Chilson and Adam Thierer, "The Coming Onslaught of 'Algorithmic Fairness' Regulations," The Regulatory Transparency Project of the Federalist Society, SSRN, November 2, 2022, https://ssrn.com/abstract=4267101.

TECHFREEDOM
LAW FOR A DYNAMIC FUTURE

Testimony of

Ari Cohn[i]

TechFreedom

Before the Senate Committee on Rules & Administration

AI and the Future of our Elections

September 27, 2023

[i] Ari Cohn is Free Speech Counsel at TechFreedom, a nonprofit, nonpartisan technology policy think tank. He can be reached at acohn@techfreedom.org.

TABLE OF CONTENTS

TESTIMONY OF ARI COHN

Chair Klobuchar, Ranking Member Fischer, and Honorable Members of the Committee, thank you for the opportunity to testify today. My name is Ari Cohn, and I am Free Speech Counsel at TechFreedom, a nonpartisan, nonprofit organization devoted to technology law and policy, the protection of civil liberties and the rule of law in the digital age, and the enabling of innovation that drives technological advancement to the benefit of society. I have previously served as the director of the Individual Rights Defense Program at the Foundation for Individual Rights in Education (now the Foundation for Individual Rights and Expression), and for many years I have maintained a private practice for the purpose of defending individuals against abusive defamation and other litigation intended to silence them.

In over a decade of working as a First Amendment lawyer, I have defended expressive rights of speakers without regard for whether I agree with the substance their expression. I have defended the rights of those who would use their voice for good, and the rights of those would wield their expression as weapons of hatred and discord. I have done so not out of any desire to see noxious speech gain acceptance, but rather out of twin convictions: that engaging with ideas that we find disagreeable or contrary to our deeply held beliefs ultimately benefits both individuals and society, and that constitutional limitations on speech regulations must protect all if they are to protect any.

That we are "seeing AI used as a tool to influence our democracy"[1] is no surprise. The very purpose of the AI technology that prompted this hearing, as with any advancement in communications technology, is to allow us to better, more easily, and more efficiently communicate with one another. In other words, AI promises to enhance the activity most fundamental to our democracy and liberty.

New forms of communication—new media—have always been accompanied by concerns about their impact on the political atmosphere. The invention of the printing press posed a threat to the power of secular and ecclesiastical authorities, prompting a campaign of repression.[2] Yet ultimately, the age of mass communication dawned. When people began to flock to the Internet to express themselves, fears of unlimited and unregulated spending on political speech prompted attempts to regulate even the personal political blogger sharing

---

[1] Press Release, Klobuchar, Hawley, Coons, Collins Introduce Bipartisan Legislation to Ban the Use of Materially Deceptive AI-Generated Content in Elections (Sept. 12, 2023), https://www.klobuchar.senate.gov/public/index.cfm/2023/9/klobuchar-hawley-coons-collins-introduce-bipartisan-legislation-to-ban-the-use-of-materially-deceptive-ai-generated-content-in-elections

[2] Jonny Thomson, *People destroyed the printing presses out of* fear, BIG THINK (Apr. 6, 2023), https://bigthink.com/the-past/printing-press-ai/.

his views for free with the world.[3] But in the end, we settled on a scheme that allows citizens to share their political views on the low-cost, democratized forum of the Internet free from onerous regulatory burdens. This approach has unleashed boundless new possibilities for civic engagement.[4] Advances in communications technology will ultimately enhance our ability to express ourselves and engage in civic discourse. We should be excited for, not fearful of, these expanding horizons.

That isn't to say that the good comes without any bad. Technology, like the expression it enables, can be used for nefarious ends. But fear of that possibility should not lead us to reflexively stifle innovation and new avenues for speech.

It won't surprise any of you to hear that falsehoods occasionally arise in the context of political campaigns. Our nation's earliest elections were marked by crude lies and insults that would draw gasps even today.[5]

So too, deceptive editing of media has been a regular source of political controversy since long before the existence of today's AI tools. To provide only a few recent examples:

- During the 2012 presidential campaign, the Romney campaign was accused of deceptively editing audio of a speech by then-President Obama in the infamous "you didn't build that" campaign advertisement. A portion of the speech referring to the support that infrastructure provides was spliced out, creating the impression that Obama impugned the hard work of business owners.[6]

---

[3] *See* Robert Cwiklik, *Running Into Old-Media Rules, Web Soapboxes Find Trouble*, WALL STREET JOURNAL (Nov. 11, 1999), https://www.wsj.com/articles/SB942271586660571929; Press Release, ACLU, Mr. Smith Goes to Washington.com (Oct. 13, 1999), https://www.aclu.org/press-releases/mr-smith-goes-washingtoncom-how-small-town-internet-speaker-tripped-over-campaign.

[4] *See* End of Year Statement from Chairman Lee E. Goodman, Federal Elections Commission 2 (Dec. 2014), https://www.fec.gov/resources/about-fec/commissioners/goodman/statements/LEG_Closing_Statement_Dec_2014.pdf; Ajit Pai & Lee Goodman, *Internet Freedom Works*, POLITICO (Feb. 23, 2015), https://www.politico.com/magazine/story/2015/02/fcc-internet-regulations-ajit-pai-115399/ ("The freedom protected by the 2006 rule fostered a robust national forum for political discussion."); Uncompensated Internet activity by individuals that is not a contribution, 11 C.F.R. § 100.94 (2016).

[5] In the presidential election of 1800, supporters of Thomas Jefferson called John Adams a "hideous hermaphroditical character which has neither the force or firmness of a man, nor the gentleness and sensibility of a woman," and supporters of Adams referred to Jefferson as "the son of a half-breed Indian squaw, sired by a Virginia mulatto father." Jed Shugerman, *The Golden or Bronze Age of Judicial Selection?*, 100 IOWA L. REV. BULL. 69, 74 (2015).

[6] Rachel Weiner, *Romney releases 'You didn't build that' ad*, WASH. POST (July 20, 2012, 8:59 AM), https://www.washingtonpost.com/blogs/the-fix/post/romney-releases-you-didnt-build-that-ad/2012/07/20/gJQAbGMrxW_blog.html; Stephanie Condon, *Obama responds to "you didn't build that" attack in new ad*, CBS NEWS (July 24, 2012), https://www.cbsnews.com/news/obama-responds-to-you-didnt-build-that-attack-in-new-ad/ (Obama campaign spokesperson Jen Psaki told reporters today, "We are not going to

- In 2020, Congressman Steve Scalise was accused of deceptively editing a video of a conversation between an activist and President Biden by appending the words "for police" from another portion of the video to the end of a question about redirecting funding. Scalise spokesperson Lauren Fine claimed that such edits "for clarity" are "common practice."[7]
- In 2020, the Biden campaign was accused of deceptively editing a video of then-President Trump for a campaign ad, splicing out more than a dozen sentences from a campaign rally speech to make it appear that Trump called COVID-19 a "hoax."[8]

None of these deceptive edits required AI. While once they may have required sophisticated or expensive audio-visual equipment, now anyone with a computer can perform them with free, easy-to-use, non-AI editing software.

Put simply, AI has not created a new problem. Rather, it's just the latest iteration of a longstanding political reality. Because AI presents a difference (however little) in degree, rather than kind, it is worth carefully examining whether special treatment is warranted—and whether such treatment might pose particular constitutional problems.

The preservation of our democratic processes is surely paramount. Free and fair elections, and the peaceful transition of power, have long set the United States apart from much of the world. But a fundamental prerequisite to our prized democratic self-governance is free and unfettered discourse,[9] especially regarding political affairs and campaigns for public office. First Amendment protection is "at its zenith" for such core political speech,[10] and has its "fullest and most urgent application to speech uttered during a campaign for political office."[11]

---

stand by while Mitt Romney slices and dices and deliberately takes out of context the president's remarks on businesses.").

[7] *See* David Weigel, *Twitter flags GOP video after activist's computerized voice was manipulated*, WASH. POST (Aug. 30, 2020, 9:21 PM), https://www.washingtonpost.com/politics/2020/08/30/ady-barkan-scalise-twitter-video/; Tyler Olson, *House Dems campaign arm files ethics complaint against Scalise after controversy over Biden video edit*, FOX NEWS (Sept. 3, 2020, 1:29 PM), https://www.foxnews.com/politics/house-dems-campaign-arm-files-ethics-complaint-against-scalise-after-controversy-over-video-edit.

[8] Meg Kelly, *Biden ad manipulates video to slam Trump*, WASH. POST (Mar. 14, 2020, 3:00 AM), https://www.washingtonpost.com/politics/2020/03/13/biden-ad-manipulates-video-slam-trump/

[9] Citizens United v. FEC, 558 U.S. 310, 339 (2010) ("Speech is an essential mechanism of democracy, for it is the means to hold officials accountable to the people.").

[10] Meyer v. Grant, 486 U.S. 414, 425 (1988).

[11] Eu v. S.F. Cnty. Democratic Cent. Comm., 489 U.S. 214, 223 (1989) (quoting Monitor Patriot Co. v. Roy, 401 U.S. 265, 272 (1971)) (internal quotation marks omitted).

Of course, speech does not become any less speech simply because it is created with the assistance of AI. Accordingly, any regulation of its use in election-related speech will face the same constitutional hurdles as regulations of non-AI speech. Because concerns about the use of AI pertain largely to the potential for the dissemination of false and misleading information, I begin with a brief overview of the jurisprudence of false speech generally and within the context of elections, and then highlight the constitutional challenges of regulating election-related AI speech specifically.

I. Constitutional Protection for False Speech

"[A]s a general matter, the First Amendment means that the government has no power to restrict expression because of its message, its ideas, its subject matter, or its content."[12] Only narrow categories of speech have historically been considered outside the protection of the First Amendment, such as defamation, fraud, obscenity, incitement, and true threats.[13]

The government asked the Supreme Court to add false speech to that list in *United States v. Alvarez*.[14] At issue was the Stolen Valor Act, which criminalized false representations of having been awarded any military decoration or medal.

The Court held that false speech is *not* categorically unprotected by the First Amendment: "This comports with the common understanding that some false statements are inevitable if there is to be an open and vigorous expression of views in public and private conversation, expression the First Amendment seeks to guarantee."[15]

In so holding, the Court rejected the government's argument that existing prohibitions of false speech indicated its exclusion from First Amendment protection. While no opinion in *Alvarez* commanded a majority, both Justice Kennedy's plurality opinion and Justice Breyer's concurring opinion rebutted this proposition in similar fashion.

Both opinions noted that such prohibitions are tied to specific, legally cognizable harms to identifiable victims.[16] Fraud, for example, requires not only a showing of a knowing falsehood, but that the falsehood is material, relied upon by the victim, and causes "actual injury."[17] And notably, both opinions distinguished statutes that punish perjury, false statements to government officials, and false representations of government authority.

---

[12] Ashcroft v. American Civil Liberties Union, 535 U.S. 564, 573 (2002) (internal quotation marks omitted).

[13] *See* United States v. Stevens, 559 U.S. 460, 468 (2010).

[14] 567 U.S. 709 (2012).

[15] *Id.* at 718.

[16] *Id.* at 719 (plurality); *id.* at 734 (Breyer, J., concurring).

[17] *Id.* at 734 (Breyer, J., concurring).

Those statutes, wrote Justice Kennedy and Justice Breyer, "protect the integrity of Government processes"[18] and prevent the "particular and specific harm [of] interfering with the functioning of [government]."[19]

Thus, the Court's animating concern was that the Stolen Valor Act constituted a broad ban on false speech irrespective of a tangible, cognizable harm. "Were the Court to hold that the interest in truthful discourse alone is sufficient to sustain a ban on speech, absent any evidence that the speech was used to gain a material advantage, it would give government broad censorial power unprecedented in this Court's cases or in our constitutional tradition."[20]

Having decided that false speech is not categorically excluded from First Amendment protection, the *Alvarez* Court split on the level of scrutiny that should apply to regulations of false speech generally.

The plurality applied strict scrutiny as in any other case involving content-based speech regulation.[21] While acknowledging that the government had a compelling interest in protecting the Medal of Honor, the plurality found that the government had not shown that the law was "actually necessary" by way of a "direct causal link" between the harm and the proposed restriction, and that the law was not narrowly tailored: the government had not shown why counterspeech would not sufficiently address the government's interest.[22]

Justice Breyer, joined by Justice Kagan, also acknowledged the significance of the government's interest.[23] But they would have applied intermediate scrutiny. In their view, because the Stolen Valor Act regulated "false statements about easily verifiable facts" not related to "philosophy, religion, history, the social sciences, the arts, and the like," it posed a lower risk of inhibiting valuable contributions to the marketplace of ideas and thus deserved a lower level of scrutiny.[24] Nevertheless, Justices Breyer and Kagan, too, found that the law's breadth and applicability in instances posing little threat of harm rendered it unconstitutional.[25]

---

[18] *Id.* at 721 (plurality).

[19] *Id.* at 734–35 (Breyer, J., concurring).

[20] *Id.* at 723 (plurality).

[21] *Id.* at 724.

[22] *Id.* at 725–28.

[23] *Id.* at 737 (Breyer, J., concurring).

[24] *Id.* at 731–32.

[25] *Id.* at 737–38.

II. False Election Speech

While *Alvarez* did not involve election-related speech, in distinguishing permissible prohibitions on perjury and other speech that impairs government functions, the Court highlighted an important distinction between two types of election-related harms: harms to the electoral *process*, and harms to the electoral *conversation* caused by the substance of political speech (what I will refer to as "electoral substance").

A. Harms to the Electoral Process

Regulations are on firm constitutional ground when they safeguard the integrity of the electoral process and access to the ballot.[26] There is "indisputably . . . a compelling interest in preserving the integrity of [the] election process."[27] Accordingly, to protect voters against deception, confusion, intimidation, and fraud in the voting *process*, courts have upheld campaign-free zones around polling places,[28] restrictions on who may appear on the ballot,[29] bans on write-in voting during primary elections,[30] and other reasonable regulations.

Similarly, laws prohibiting voter intimidation, the purchase or sale of votes, and the like do not offend the First Amendment. These laws properly reach knowingly false statements about the electoral process itself, such as disinformation about voting procedures, places, and times.[31] Thus, a criminal prosecution for social media posts intended to mislead voters into believing they could vote by text message was held constitutional.[32] Likewise, prosecutions and a civil lawsuit for false robocalls to black neighborhoods stating that police and debt collectors would use personal information from mail-in voters were likewise held not to offend the First Amendment.[33] Indeed, because such statements involved "easily verifiable facts" and posed little risk of impacting substantive political speech, they were analyzed under the intermediate scrutiny applied by the *Alvarez* concurrence.[34]

---

[26] Burson v. Freeman, 504 U.S. 191, 199 (1992).

[27] *Id.*

[28] *Id.*

[29] American Party of Texas v. White, 415 U.S. 767 (1974); Bullock v. Carter, 405 U.S. 134 (1972); Jenness v. Fortson, 403 U.S. 431 (1971).

[30] Burdick v. Takushi, 504 U.S. 428 (1992).

[31] *See* Minn. Voters All. v. Mansky, 138 S. Ct. 1876, 1889 n.4 (2018) ("We do not doubt that the State may prohibit messages intended to mislead voters about voting requirements and procedures.").

[32] United States v. Mackey, No. 21-CR-80 (NGG) (E.D.N.Y. Jan. 23, 2023).

[33] Nat'l Coal. on Black Civil Participation v. Wohl, No. 20 Civ. 8668 (VM) (S.D.N.Y. Mar. 8, 2023); People v. Burkman, No. 356600 (Mich. Ct. App. Jun. 2, 2022).

[34] Wohl, No. 20 Civ. 8668 (VM) at 78–79; Mackey, No. 21-CR-80 (NGG) at 47.

Clearly, the government has significant latitude to regulate false statements of fact regarding the electoral process itself. However, it is unlikely that AI-specific regulations are needed here; existing laws are technology-agnostic and would appear to cover any such activity, whether AI-generated or not.

B. Laws Targeting Harms to the Electoral Conversation Will Fail Strict Scrutiny

Conversely, speech restrictions are perhaps most suspect when they regulate the content of electoral substance and attempt to determine for the polity what political information is true or false. Regulation of "what is said or distributed during an election . . . goes beyond an attempt to control the process to enhance the fairness overall so as to carefully protect the right to vote."[35] Because the right, and civic duty, to evaluate such speech rests with the electorate,[36] government attempts to "enhance[e] the ability of its citizenry to make wise decisions by restricting the flow of information to them must be viewed with some skepticism."[37]

Following *Alvarez*, courts have indeed cast a wary eye on laws prohibiting false statements about candidates and ballot issues.

Such laws are generally drawn too broadly to be upheld as regulating only categories of unprotected speech. For instance, Massachusetts defended its law prohibiting false statements made about a candidate "designed or that tend[] to aid or to injure or defeat such candidate" by asserting that such statements constituted either fraud or defamation, obviating the need for First Amendment scrutiny.[38] But the Supreme Judicial Court of Massachusetts found that the law reached much further, encompassing speech that constituted neither fraud (because the law did not require a showing of reliance or damage) nor defamation (because it reached "statements regarding ballot questions and statements by a candidate about himself designed to enhance his own candidacy, i.e., statements that clearly are not defamatory.").[39] Accordingly, such laws will be subject to First Amendment scrutiny.

Because political speech receives the highest protection and because of the dangers of allowing the government to operate as a political Ministry of Truth, courts have

---

[35] 281 Care Comm. v. Arneson, 766 F.3d 774, 787 (8th Cir. 2014).

[36] United States v. Alvarez, 567 U.S. 709, 728 (2012).

[37] Anderson v. Celebrezze, 460 U.S. 780, 798 (1983).

[38] Commonwealth v. Lucas, 34 N.E.3d 1242 (Mass. 2015).

[39] *Id.* at 1249–50.

overwhelmingly held that laws regulating electoral substance must satisfy strict scrutiny.[40] The government therefore bears an exceptionally heavy burden: it must prove that its restriction is necessary to serve a compelling government interest, is narrowly tailored to serve that interest, and is the least restrictive means of achieving its stated goal.[41]

Even on the threshold question, whether there is a legitimate government interest in protecting the public from false speech during campaigns, there is debate. To one judge, such an interest "is patronizing and paternalistic . . . . It assumes the people of this state are too ignorant or disinterested to investigate, learn, and to determine for themselves the truth or falsity in political debate . . . ."[42] On the other hand, many courts have either held or assumed that the government *does* have a compelling interest in assuring that the electorate is not led astray by false electoral substance[43]—yet have struck down such laws anyway with relative ease after finding that they are not actually necessary or narrowly tailored. A few themes, echoing the plurality opinion in *Alvarez*, recur in these courts' analyses.

1. Is the Restriction Actually Necessary?

Content-based speech restrictions must not only serve a compelling government interest, but must also be "actually necessary" to achieve that interest.[44] "The state must specifically identify an actual problem in need of solving," and demonstrate a "direct causal link" between the harm to be prevented and the restriction chosen by the government.[45]

Where the government has not demonstrated that there is an existing threat of false statements that will cause harm, such laws have been invalidated. For example, in *281 Care Committee v. Arneson*, the U.S. Court of Appeals for the Eighth Circuit considered a challenge

---

[40] *See* 281 Care Comm., 766 F.3d at 784; *see also* Alvarez, 567 U.S. at 738 (Breyer, J., concurring) ("I recognize that in some contexts, particularly political contexts, such a narrowing will not always be easy to achieve. In the political arena . . . the statute may have to be significantly narrowed in its applications."); Ex parte Stafford, No. 05-22-00396 at 8 (Tex. App. May 1, 2023).

[41] Burson v. Freeman, 504 U.S. 191, 198 (1992).

[42] State v. 119 Vote No! Committee, 135 Wn. 2d 618, 631–32 (Wash. 1998).

[43] *See, e.g.*, Susan B. Anthony List v. Driehaus, 814 F.3d 466, 473 (6th Cir. 2016) ("Here, Ohio's interests in preserving the integrity of its elections, protecting voters from confusion and undue influence, and ensuring that an individual's right to vote is not undermined by fraud in the election process are compelling.") (internal quotation marks and citations omitted); 281 Care Comm., 766 F.3d at 787 ("Today we need not determine whether, on these facts, preserving fair an honest elections and preventing fraud on the electorate comprise a compelling state interest because the narrow tailoring that must juxtapose that interest is absent here."); Ex parte Stafford, No. 05-22-00396 at 8 (Tex. App. May 1, 2023) ("Stafford does not, and reasonably could not, dispute that promoting honest discourse and preventing misinformation in the political arena are compelling state interests.").

[44] United States v. Alvarez, 567 U.S. 709, 725 (2012).

[45] Brown v. Entertainment Merchants Assn., 564 U.S. 786, 799 (2011).

to a Minnesota law prohibiting the knowing dissemination of false information about a ballot question in any paid political advertising or campaign material.[46] Instead of presenting empirical evidence for the law's necessity, Minnesota simply "assert[ed] 'that common sense dictates that political advertising aimed at voters and intentionally designed to induce a particular vote through the use of false facts impacts voters' understanding and perceptions; can influence their vote; and ultimately change an election.'"[47] While such an inference might be justified in protecting consumers from fraudulent *commercial* advertising under intermediate scrutiny,[48] it cannot be applied to non-commercial, political speech under strict scrutiny. The court found that the state had not met its burden to prove that the law was actually necessary: "We have never accepted mere conjecture as adequate to carry a First Amendment burden .... Such conjecture about the effects and dangers of false statements equates to implausibility ... because, when the statute infringes core political speech, we tend not to take chances."[49]

2. Is the Restriction Overinclusive?

A law is not narrowly tailored if it "sweep[s] too broadly" and restricts more speech than necessary to achieve the government's interest.[50] Laws regulating false electoral substance have been struck down because of several variations of over-inclusivity.

First, some courts have held that a law sweeps too broadly when it prohibits even non-material falsehoods: "Thus, influencing an election by lying about a political candidate's shoe size or vote on whether to continue a congressional debate is just as actionable as lying about a candidate's party affiliation or vote on an important policy issue ... Penalizing non-material statements, particularly those made outside the political arena, is not narrowly tailored to preserve fair elections."[51] Materiality is essential, and unlike consumer protection law, which involves commercial speech,[52] the government cannot infer the materiality of claims about elections, which are non-commercial speech.

---

[46] 766 F.3d 774 (8th Cir. 2014).

[47] *Id.* at 787.

[48] Hudson Gas Elec. v. Public Serv. Comm'n, 447 U.S. 557, 567-68 (1980) ("In the absence of factors that would distort the decision to advertise, we may assume that the willingness of a business to promote its products reflects a belief that consumers are interested in the advertising."). Thus, the Federal Trade Commission has long presumed that "that express claims are material" when it polices advertising. Letter from the FTC to the Committee on Energy & Commerce, appended to Cliffdale Assocs., Inc., 103 F.T.C. 110, 174 (1984), https://www.ftc.gov/system/files/documents/public_statements/410531/831014deceptionstmt.pdf.

[49] *281 Care Comm.*, 766 F.3d at 790-91.

[50] *Id.* at 787.

[51] Susan B. Anthony List, 814 F.3d at 475.

[52] *See supra* note 48.

Second, some courts—consistent with Justice Breyer's concern about the Stolen Valor Act's breadth—have found that the law at issue restricts speech that causes minimal risk of harm: "[The law] reaches not only those statements that are widely disseminated through commercial advertisement, but also those exchanged between two friends engaged in a spirited political discussion . . . ."[53]

Finally, courts have held that false electoral substance laws are not narrowly tailored when they restrict speech over a long period of time, rather than close in time to an election.[54] The further removed from the date of an election, the more effective counterspeech can be, and the greater the availability of regulatory alternatives, as discussed below.

3. Is It the Least Restrictive Means?

Whenever the government imposes a content-based speech restriction, it must choose "the least restrictive means among available, effective alternatives."[55] When a law restricts statements of false electoral substance, that analysis will nearly always begin and end with counterspeech. "The remedy for speech that is false is speech that is true," ruled the *Alvarez* Court, adding: "That is the ordinary course in a free society."[56]

The government's burden to disprove the efficacy of counterspeech is appropriately high where it seeks to regulate core political speech. "Statutes broadly suppressing false statements about candidates or ballot questions cannot withstand strict scrutiny for the simple reason that [o]ur constitutional election system already contains the solution to the problem . . . That solution is counterspeech."[57] Indeed, courts have recognized that "there is no greater arena wherein counterspeech is at its most effective" than in the political

---

[53] Commonwealth v. Lucas, 34 N.E.3d at 1255.

[54] *See* Susan B. Anthony List, 814 F.3d at 476; Commonwealth v. Lucas, 34 N.E.3d at 1254–55.

[55] Ashcroft v. American Civil Liberties Union, 542 U.S. 656, 666 (2004).

[56] United States v. Alvarez, 567 U.S. 709, 727 (2012). *See also* Whitney v. California, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring) ("To courageous, self-reliant men, with confidence in the power of free and fearless reasoning applied through the processes of popular government, no danger flowing from speech can be deemed clear and present, unless the incidence of the evil apprehended is so imminent that it may befall before there is opportunity for full discussion. If there be time to expose through discussion the falsehood and fallacies, to avert the evil by the processes of education, the remedy to be applied is more speech, not enforced silence. Only an emergency can justify repression.").

[57] Commonwealth v. Lucas, 34 N.E.3d at 1253.

context.[58] There are precious few circumstances where this "tried and true buffer and elixir" will be considered inadequate.[59]

Perhaps the only such hint of such circumstance in election speech jurisprudence is where a speech restriction operates in a limited time frame close to an election, where there may not be adequate time to uncover and respond to falsehoods.[60] Just how close to the election such a time frame must be is unclear, but that window is likely shrinking rather than growing: while virality may cause falsehoods to spread rapidly, it can just as easily cause counterspeech to do the same.

III. Application to AI-Created Election Speech

Nothing about speech created using AI would remove it from the constitutional analysis outlined above. As such, the clearest way to illustrate the First Amendment challenges of regulating AI election speech may be simply to analyze an existing legislative proposal.

A. The Protect Elections from Deceptive AI Act (S. 2770)

Introduced on September 12, 2023, the Protect Elections from Deceptive AI Act would ban "materially deceptive AI-generated audio or visual media" from election-related speech and political advertisements.[61]

Its prohibition reads as follows: "[A] person, political committee, or other entity may not knowingly distribute materially deceptive AI-generated audio or visual media of a [candidate

---

[58] 281 Care Comm. v. Arneson, 766 F.3d 774, 793 (8th Cir. 2014). See also Alvarez, 567 U.S. at 738 (Breyer, J., concurring) ("I would also note, like the plurality, that in [the political arena] more accurate information will normally counteract the lie."); Grimmett v. Freeman, 59 F.4th 689, 695 (4th Cir. 2023) ("Public officials and public figures . . . have a more realistic opportunity to counteract false statements.") (quoting Gertz v. Robert Welch, Inc., 418 U.S. 323, 344 (1974)); Ex parte Stafford, No. 05-22-00396-CR at 15 (Tex. App. May 1, 2023) ("Our constitutional tradition is deeply rooted in the notion that the best test of truth is the power of the thought to get itself accepted in the competition of the marketplace) (quoting Abrams v. United States, 250 U.S. 616, 630 (1919) (Holmes, J., dissenting)).

[59] 281 Care Comm., 766 F.3d at 793.

[60] *See, e.g.*, Susan B. Anthony List, 814 F.3d at 476 (noting that the statute was constitutionally flawed because it included statements "whether made on the eve of an election, when the opportunity to reply is limited, or months in advance.") (quoting McIntyre v. Ohio Elections Comm'n, 514 U.S. 334, 351–52 (1995)) (internal quotation marks omitted).

[61] Press Release, Klobuchar, Hawley, Coons, Collins Introduce Bipartisan Legislation to Ban the Use of Materially Deceptive AI-Generated Content in Elections (Sept. 12, 2023), https://www.klobuchar.senate.gov/public/index.cfm/2023/9/klobuchar-hawley-coons-collins-introduce-bipartisan-legislation-to-ban-the-use-of-materially-deceptive-ai-generated-content-in-elections

for federal office], or in carrying out a Federal election activity, with the intent to—(1) influence an election; or (2) solicit funds."[62]

"Deceptive AI-generated audio or visual media" ("DAAV") is defined as:

> an image, audio, or video that—
>
> (A) is the product of artificial intelligence or machine learning, including deep learning techniques, that—
>
> > (i) merges, combines, replaces, or superimposes content onto an image, audio, or video, creating an image, audio, or video that appears authentic; or
> >
> > (ii) generates an inauthentic image, audio, or video that appears authentic; and
>
> (B) a reasonable person, having considered the qualities of the image, audio, or video and the nature of the distribution channel in which the image, audio, or video appears—
>
> > (i) would have a fundamentally different understanding or impression of the appearance, speech, or expressive conduct exhibited in the image, audio, or video than that person would have if that person were hearing or seeing the unaltered, original version of the image, audio, or video; or
> >
> > (ii) would believe that the image, audio, or video accurately exhibits any appearance, speech, or expressive conduct of a person who did not actually exhibit such appearance, speech, or expressive conduct.[63]

S. 2770 permits a candidate "whose voice or likeness appears in, or who is the subject of" a prohibited piece of media to file suit seeking injunctive relief prohibiting the distribution of the offending media, as well as for general or special damages.[64]

---

[62] Protect Elections from Deceptive Ads Act, S. 2770, 118th Cong. §2(a) (adding to 52 U.S.C. 30101 et seq., a new Sect. 325(b)).

[63] *Id.* (adding to 52 U.S.C. 30101 et seq., a new Sect. 325(a)(2)).

[64] *Id.* (adding to 52 U.S.C. 30101 et seq., a new Sect. 325(d)).

Excluded from the bill's prohibitions are media outlets that issue disclosures regarding the veracity of the media,[65] and content that constitutes parody or satire.[66]

B. The Constitutionality of S. 2770

The principles discussed above reveal several constitutional infirmities in this bill. While some of these might be present in any attempt to regulate false election speech, others are caused specifically by the specific targeting of AI-created speech.

1. Does the Prohibition Implicate Protected Speech?

Only if S. 2770 regulates only unprotected speech will it escape First Amendment scrutiny altogether. But because false speech is *not* categorically excluded from constitutional protection, the speech prohibited by the bill must fall into one of the traditional categories of unprotected speech. The relevant categories here are fraud and defamation.

Fraud. Clearly, some speech prohibited by S. 2770 might constitute fraud. Perhaps most obviously, DAAV might be used falsely to portray a solicitation as coming from a candidate. But as in *Lucas*, "the fact that [S. 2770] may reach fraudulent speech is not dispositive, because it also reaches speech that is not fraudulent."[67]

Indeed, the vast majority of speech prohibited by S. 2770 is likely to *not* be fraudulent. This is particularly so because the bill does not require reliance (or intent to induce reliance) or harm to a recipient for DAAV to be prohibited. Reliance is an essential element of any common law fraud tort.[68]

Defamation. On its face, S. 2770's prohibition goes beyond only defamatory speech. Certainly, a substantial amount of DAAV may constitute actionable defamation. But much of it may not: S. 2770 does not require that DAAV be injurious to a candidate's reputation—a basic element of any defamation claim.[69]

---

[65] *Id.* (adding to 52 U.S.C. 30101 et seq., a new Sect. 325(c)(1)–(2)).

[66] *Id.* (adding to 52 U.S.C. 30101 et seq., a new Sect. 325(c)(3).

[67] Commonwealth v. Lucas, 34 N.E.3d at 1249.

[68] *See, e.g.*, RESTATEMENT (SECOND) OF TORTS § 525 (1977) ("One who fraudulently makes a misrepresentation of fact . . . for the purpose of inducing another to act or to refrain from action in reliance upon it, is subject to liability to the other in deceit for pecuniary loss caused to him by his justifiable reliance upon the misrepresentation.")

[69] Gertz v. Robert Welch, Inc., 418 U.S. 323, 346 (1974); *see also* Andrew Sebbins, Buckingham, Doolittle & Burroughs, LLC, *Elements of Defamation*, JD SUPRA (Aug. 7, 2023), https://www.jdsupra.com/legalnews/elements-of-defamation-2431458/.

But S. 2770(b) creates a wrinkle: any violation of the bill constitutes defamation per se—establishing any material DAAV as injurious.[70]

As a threshold matter, defamation is a creature of state law; there is no federal defamation law. Congress's authority to add a new category of defamation per se to each state's law is questionable at best.

In addition, this provision results in absurdity. Suppose a supporter of a candidate distributes a *flattering* yet material DAAV featuring the candidate's voice or likeness. No actual injury to the candidate's reputation has been done; to the contrary, their reputation has been *bolstered* (falsely). Nevertheless, the DAAV is considered defamatory per se. Contrast this with common law: certain statements are considered per se defamatory only because they are so *obviously and materially harmful* that injury can be presumed.[71]

In any event, S. 2770 plainly prohibits other speech that is not defamatory. For instance, a candidate is prohibited from distributing material DAAV about *himself* to influence an election or solicit funds.

Clearly, S. 2770 prohibits far more than only unprotected speech and will be subject to First Amendment scrutiny.

2. What Level of Scrutiny Applies?

Because S. 2770 broadly regulates content of electoral substance (and is therefore content-based), it is virtually certain that courts would subject it to strict scrutiny.[72] One might argue that material DAAV is closer to "easily verifiable facts," regulation of which generally poses less risk of curtailing valuable contributions to the marketplace of ideas—and that therefore S. 2770 should therefore receive only intermediate scrutiny.[73]

But that is not so for several reasons. First, the apparent premise behind the need for regulation is that material DAAV are in fact difficult to detect, which necessarily makes them more difficult to verify.

---

[70] Protect Elections from Deceptive Ads Act, S. 2770, 118th Cong. §2(b).

[71] *See* Bryson v. News America Publications, 174 Ill. 2d 77, 87 (1996).

[72] *See supra* notes 40–41 and accompanying text.

[73] *See supra* notes 23–24 and accompanying text.

Second, the premise that AI-generated media does not implicate valuable political speech is false. S. 2770's expansive definition of DAAV sweeps up a broad array of potentially valuable speech and declares it off-limits.[74]

3. Is There a Compelling Government Interest?

Whether courts will find a compelling interest ensuring only truthful speech related to elections is unclear.[75] "Society has the right and civic duty to engage in open, dynamic, national discourse,"[76] and "it is the citizenry that can discern for themselves what the truth is, not [government]."[77] But assume for the sake of argument that a court analyzing S. 2770 would at least presume such an interest to determine whether the law is narrowly tailored.

4. Is the Speech Restriction Actually Necessary

S. 2770's prohibition must be actually necessary to serve that interest, and it must be justified by evidence that there is a concrete problem that needs to be addressed.[78]

It is unclear that evidence of such a problem currently exists. Despite breathless warnings, deepfakes did not appear to play any meaningful role in the 2020 election.[79] We are now hearing the same warnings about the impeding election. And while it is true that generative AI is more accessible now than it was in 2020, but projections based on availability of the technology alone does not rise above the level of conjecture, which is insufficient to justify a content-based speech restriction.[80]

It is also unclear as a general matter whether there is any evidence that material DAAV actually influences voters' beliefs, and ultimate voting decisions. Indeed, it seems likely that the risk of such influence is at least somewhat overstated.[81] Our political culture is increasingly polarized, and voters tend to be skeptical of content that casts a negative light

---

[74] *See infra* notes 80–82 and accompanying text.

[75] See *supra* note 42 and accompanying text.

[76] United States v. Alvarez, 567 U.S. 709, 728 (2012) (plurality).

[77] 281 Care Comm. v. Arneson, 766 F.3d 774, 793 (8th Cir. 2014).

[78] *See supra* notes 44–44 and accompanying text.

[79] Tom Simonite, *What Happened to the Deepfake Threat to the Election?*, WIRED (Nov. 16, 2020, 7:00 AM), https://www.wired.com/story/what-happened-deepfake-threat-election/.

[80] *See supra* note 49.

[81] *How worried should you be about AI disrupting elections?*, POLITICO (Aug. 31, 2023), https://www.economist.com/leaders/2023/08/31/how-artificial-intelligence-will-affect-the-elections-of-2024.

on their candidates of choice. They are unlikely to change their vote on account of such content.

Given the paucity of data indicating that DAAV presents an actual, rather than hypothetical, problem, the government would struggle to prove the necessity of S. 2770.

5. Is the Law Narrowly Tailored?

Even if the government could satisfy its burden of showing that S. 2770 is actually necessary, the bill fails the narrow tailoring prong of strict scrutiny several times over.

Overinclusiveness. As with most non-AI related electoral substance regulations, S. 2770 is staggeringly overinclusive.

Perhaps most glaringly, S. 2770 apparently deems it DAAV even when the ultimate message produced is *true*. There is no limiting factor requiring falsity of the overall message; all that is required is that AI-produced media in any part of a message or advertisement give "a fundamentally different understanding or impression of the appearance, speech, or expressive conduct" than the original version.[82] The prohibition of true messages alone renders S. 2770 unconstitutional.[83]

In a similar vein, S. 2770 would also prohibit protected political opinion. DAAV might not always be used to convey the false impression that a candidate literally did or said what is portrayed. Rather, such media might be used as a means by which to *characterize* the positions of a candidate. To illustrate, consider the non-AI deceptively edited media examples discussed above. Both the Romney campaign advertisement and Congressman Scalise's video[84] were edited in ways that would seemingly qualify them as DAAV if done with AI. But both the Romney campaign and Rep. Scalise's office replied to criticism with the retort that, despite edits that rendered the media clips technically false, they accurately conveyed their view of the target's beliefs. While Barack Obama and Joe Biden would disagree with that assertion, government prohibition of such devices would inappropriately intrude on prototypical political discourse.

S. 2770 is also overinclusive because it prohibits much speech that poses little risk of harm. *All* individuals are prohibited from disseminating material DAAV to *anyone at all*. Thus a

---

[82] Protect Elections from Deceptive Ads Act, S. 2770, 118th Cong. §2(a) (adding to 52 U.S.C. 30101 et seq., a new Sect. 325(a)(2)).

[83] *See* Grimmett v. Freeman, 59 F.4th 689, 694 (4th Cir. 2003) (invalidating North Carolina's criminal election libel statute in part because it permitted punishment of true statements made with reckless disregard for their truth or falsity).

[84] *See supra* notes 6-7 and accompanying text.

DAAV sent to a family member or small group of friends is just as prohibited as a campaign advertisement broadcast to millions. Whatever justification the government may assert for widespread dissemination of election falsehoods, prohibiting communications between friends or relatives that pose little if any risk of harm clearly sweeps too broadly. Whether or not enforcement is likely, the prohibition alone would unacceptably chill protected speech.

Finally, S. 2770 is overinclusive in its temporal breadth—it prohibits dissemination of DAAV at *any* time if intended to influence an election. "It may be invoked as soon as one announces his or her candidacy—not merely on the eve of the election."[85] Such a wide temporal prohibition is a hallmark of over-inclusivity in election falsehood regulations.

Underinclusivity. At the same time, S. 2770 is also *underinclusive*. "A statute can also fail strict scrutiny if it covers *too little speech*. Underinclusivity creates a First Amendment concern when the State regulates one aspect of a problem while declining to regulate a different aspect of the problem that affects its stated interest in a comparable way."[86] Underinclusivity may indicate that "a law does not actually advance a compelling government interest."[87]

S. 2770 is underinclusive in two ways.

First, the law necessarily only reaches those within the reach of United States jurisdiction; foreign actors would remain free to disseminate all of the worst kinds of DAAV unchecked.[88] While one might reasonably argue that failure to regulate worldwide should not ordinarily be an underinclusivity concern, circumstances may dictate otherwise here. Much of the concern about deceptive election speech has been precisely that it is coming *from* foreign actors seeking to interfere with our democratic institutions. Moreover, prohibiting AI-produced media domestically while leaving foreign actors free to disseminate it may in fact aggravate precisely the harm that S. 2770 seeks to prevent. It is quite possible that, believing

---

[85] Commonwealth v. Lucas, 34 N.E.3d at 1254–55.

[86] Mont. Citizens for Right to Work v. Mangan, 580 F. Supp. 3d 911, 920–21 (D. Mont. 2022) (internal quotation marks and citations omitted).

[87] Williams-Yulee v. Fla. Bar, 575 U.S. 433, 449 (2015).

[88] *See* Free Speech Coalition, Inc., v. Colmenero, No. 1:23-cv-00917, Order Granting Plaintiffs' Motion for a Preliminary Injunction, ECF No. 36 at 27, available at https://storage.courtlistener.com/recap/gov.uscourts.txwd.1172751222/gov.uscourts.txwd.1172751222.36.0.pdf (finding a Texas law requiring age verification for pornographic websites underinclusive because it left "minors able to access any pornography as long as its hosted by foreign websites with no ties to the United States.").

such media are prohibited, voters will become *more* susceptible to foreign DAAV because they anticipate it less.

More fundamentally, S. 2770 is clearly underinclusive because it fails to regulate deceptively edited media produced without the use of AI. Deceptively edited media produced manually does not merely affect the purported government interest in a comparable way; it affects it in the *same* way. Moreover, in contrast to DAAV, there is a long, proven history of the use of manually edited deceptive media in campaigns—a history as long as that of the United States. That S. 2770 ignores this well-documented problem while attempting to regulate the smaller and more speculative problem of DAAV casts serious doubt on whether the bill advances any compelling interest.

6. Is S. 2770's Prohibition the Least Restrictive Means?

Even if the litany of constitutional failures above did not exist, S. 2770 would still likely be unconstitutional because counterspeech provides a less restrictive alternative. There is no obvious reason why DAAV could not be countered with true speech, in the way that manually edited deceptive media is. Indeed, the staggering breadth and temporal scope of S. 2770 indicates that no consideration has been given at all to the types of prohibited speech that might effectively be countered with more speech. The freedom to engage with a diverse array of political speech and determine for ourselves what is true and what is false is sacrosanct. The government is not free to discard this principle so easily to assuage moral panic.

7. A Final Consideration

In addition to the constitutional infirmities detailed above, S. 2770 presents an exceptionally high risk of chilling and censoring protected political expression due to its enforcement mechanism. At present, no reliable technology exists to detect whether media has been produced by AI. Not only does that put *any* edited media at risk of an unwarranted lawsuit, it also provides a weapon with which to silence critics or opponents: the impossibility of accurate AI detection would permit a candidate to bring suit against another person or entity for any media that they simply do not like. While a prevailing defendant in an action for damages may recover their attorney's fees from an abusive plaintiff,[89] the provision for injunctive actions does not include recovery of fees by the prevailing party.[90] To cast a serious chill over the electoral discourse, a motivated candidate need only file scattershot lawsuits seeking to enjoin critical expression.

---

[89] Protect Elections from Deceptive Ads Act, S. 2770, 118th Cong. §2(a) (adding to 52 U.S.C. 30101 et seq., a new Sect. 325(d)(2)).

[90] *Id.* (adding to 52 U.S.C. 30101 et seq., a new Sect. 325(d)(1)).

The same First Amendment concerns that sharply limit prohibition of general false election speech pose a substantial obstacle to the prohibition of AI-produced media. And rather than resolving those concerns, attempts to narrowly target AI instead end up creating additional constitutional concerns. This is not to say that regulation is impossible. But in crafting such laws, Congress must take extreme care to provide the breathing room necessary for a diverse, free, and vibrant political discourse.

IV. Mandating Disclosure of AI-Produced Advertisements

Aside from regulating the use of AI, legislation has been proposed mandating disclosures that an advertisement has been produced with the use of AI. The Require the Exposure of AI-Led Political Advertisements Act (S. 1596) would require (among other things) any advertisement placed or promoted for a fee online to disclose if it "contains an image or video footage that was generated in whole or in part with the use of artificial intelligence (generative AI)."[91]

While mandatory disclosures may work less First Amendment harm than an outright prohibition on speech, they are not immune from constitutional scrutiny. Because disclosure requirements do not prevent speech, they are subjected to "exacting scrutiny, which requires a substantial relation between the disclosure requirement and a sufficient government interest."[92]

While the Supreme Court has held that "provid[ing] the electorate with information"[93] is a sufficient government interest, that informational interest has not been a general one. Rather, it has been limited to campaign contribution disclosures that assist in the prevention of corruption and the appearance of corruption that undermines public confidence in the electoral process,[94] and the disclosure of the source of paid advertisements to provide the electorate with the information they need to evaluate the source and veracity of the message.[95] A requirement to disclose the use of AI in advertisement production fits neither category.

It is unclear what government interest exists in disclosure of the means of an advertisement's production. The "simple interest in providing voters with additional relevant information does not justify a state requirement that a writer make statements or disclosures that she

---

[91] Require the Exposure of AI-Led Political Advertisements Act, S. 1596, 118th Cong. §4 (adding to 52 U.S.C. 30120 et seq., a new Sect. 325(e)(1)).

[92] Citizens United v. Fed. Election Comm'n, 558 U.S. 310, 366 (2010) (internal quotation marks omitted).

[93] *Id.* at 367 (quoting Buckley v. Valeo, 424 U.S. 1, 66 (1976).

[94] Fed. Election Comm'n v. National Right to Work Comm., 459 U.S. 197, 208 (1982).

[95] First National Bank of Boston v. Bellotti, 425 U.S. 765, 790 (1978).

would otherwise omit."[96] That an advertisement was produced using AI says nothing more relevant than the brand of computer or video camera used. The underlying presumption seems to be that AI-produced advertisements are somehow more likely to be deceptive or harmful in some manner—but that is hardly true. Various forms of AI are used in an extraordinary percentage of media productions; color correction, noise reduction, background object removal, caption text, and a large variety of other basic production tasks enlist the use of AI. The required disclosure of any use of AI in production would destroy the value of the disclosure—if everything has a disclosure, nothing has a disclosure—and the efficacy of achieving whatever government interest is asserted along with it.

Worse yet, flooding the marketplace with such disclosures gives cover to bad actors. If virtually all advertisements end up carrying AI disclosures, purveyors of falsehoods and harmful materials will be more difficult to detect. If there is a sufficient government interest to be achieved, disclosure requirements must be drafted with the utmost care and precision to avoid perpetuating the harms sought to be protected against.

CONCLUSION

Concern for our democratic processes and institutions is well-placed. But reflexive legislation prompted by fear of the next technological boogeyman will not safeguard our democratic values. Instead, intrusions on the free and unfettered political discourse that has been the lifeblood of our democracy will ultimately subvert it. Conversely, resisting the urge to legislate speculative problems out of existence before they arise will strengthen our resiliency, safeguard our fundamental liberties, and allow innovation to flourish and take us to new heights.

Thank you again for inviting me to testify before you today, and I look forward to your questions.

---

[96] McIntyre v. Ohio Elections Comm'n, 514 U.S. 334, 348 (1995).

Center for AI and Digital Policy

September 27, 2023

Chairwoman Amy Klobuchar
Ranking Member Deb Fischer
Rules and Administration Committee
U.S. Senate, Washington, DC

Re: Statement for the record: Committee Hearing on "*AI AND THE FUTURE OF OUR ELECTIONS*"

Dear Chairwoman Klobuchar, Ranking Member Fischer, and Committee Members,

We write to you regarding the upcoming hearing, "AI and the Future of Our Elections."[1] With the 2024 Presidential election approaching, we commend the Committee's oversight for AI-generated political content.

The Center for AI and Digital Policy (CAIDP) is an independent research organization, based in Washington, DC.[2] We advise national governments and international organizations regarding artificial intelligence policy. We previously testified on AI policy before the House Oversight Committee.[3] Our landmark report *AI and Democratic Values*[4] assesses AI policies and practices around the world and emphasizes the importance of transparency in national AI strategies.[5] (See attached image for digital download of the CAIDP report).

In a complaint filed earlier this year with the Federal Trade Commission, CAIDP warned of the specific Risks to Democracy of generative AI products such as ChatGPT.[6] We urged the FTC to act quickly, stating "The Federal Trade Commission may be the only federal agency with

[1] Senate Rules and Administration Committee, *AI and the Future of Our Elections*, Sept. 27, 2023, https://www.rules.senate.gov/hearings/ai-and-the-future-of-our-elections.
[2] CAIDP, https://www.caidp.org.
[3] Testimony and statement for the record of CAIDP President Merve Hickok, *Advances in AI: Are We Ready For a Tech Revolution?*, House Committee on Oversight and Accountability: Subcommittee on Cybersecurity, Information Technology, and Government Innovation (Mar. 8, 2023), https://oversight.house.gov/wp-content/uploads/2023/03/Merve-Hickok_testimony_March-8th-2023.pdf.
[4] CAIDP, *AI and Democratic Values* (2023), https://www.caidp.org/reports/aidv-2022/.
[5] *Id.* at 1132-33.
[6] CAIDP, *In the Matter of OpenAi* (2023), https://www.caidp.org/cases/openai/.

the opportunity and authority at this time to regulate ChatGPT so as to diminish threats to the 2024 election."[7]

We set out below brief observations and recommendations to address the threat of false and deceptive AI-generated political content in elections:

1) Move forward with federal AI legislation based on established governance frameworks

We need federal AI legislation which would mandate transparency, fairness, and accountability of AI systems, including those that provide information to voters and may impact election outcomes. There are well-established governance frameworks aimed at ensuring safety and scientific validity set out in the Universal Guidelines for AI[8] and the OECD AI Principles.[9]

The Universal Guidelines on AI (UGAI) were adopted in 2018 and over 330 leading experts and 60 associations (including the AAAS, the ACM, and the IEEE) have endorsed the UGAI. It sets out 12 principles that are foundational for the governance of AI systems. According to the UGAI Right to Transparency, "All individuals have the right to know the basis of an AI decision that concerns them. This includes access to the factors, the logic, and techniques that produced the outcome."[10] In these AI systems, "Both technical and institutional designs should ensure auditability and traceability of (the working of) AI systems to address any conflicts with human rights norms and standards and threats to environmental and ecosystem well-being."

The OECD has highlighted that the combination of AI language models and disinformation can lead to large-scale and damage public trust in democratic institutions.[11] A study found that Google's Bard AI tool generates persuasive misinformation content on 78 out of 100 tested narratives. [12] A recent study found that ChatGPT4 is more likely than its predecessor to generate

---

[7] CAIDP, *Supplement to the Original Complaint, In the Matter of Open AI* (2023), pp. 35, para. 138, https://files.constantcontact.com/dfc91b20901/72cccde7-44a7-44e4-bfee-d6801b3891d2.pdf.

[8] The Public Voice, *Universal Guidelines for Artificial Intelligence*, Guideline 5, https://thepublicvoice.org/ai-universal-guidelines/.

[9] Recommendation of the Council on Artificial Intelligence, OECD (May 21, 2019), https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

[10] UGAI Guideline 1.

[11] *supra* at note 7.

[12] Center for Countering Digital Hate (CCDH), Misinformation on Bard, Google's New AI Chat, Apr. 5. 2023, https://counterhate.com/research/misinformation-on-bard-google-ai-chat/#:%7E:text=Google%E2%80%99s%20new%20%E2%80%98Bard%E2%80%99%20AI%2

misinformation when prompted, including false narratives concerning vaccines, conspiracy theories, and propaganda.[13]

"OpenAI notes that even a powerful model like GPT-4 "is not fully reliable" and "great care should be taken when using language model outputs, particularly in high-stakes contexts, with the exact protocol (such as human review, grounding with additional context, or avoiding high-stakes uses altogether) matching the needs of a specific use-case." Additionally, because the models are generally trained on large amounts of data scraped from the internet, they can incorporate, reflect, and potentially amplify biases in such data."[14]

OpenAI has acknowledged the specific danger of Disinformation and influence operations. As we explained in our complaint to the FTC, OpenAI has warned that "'AI systems will have even greater potential to reinforce entire ideologies, worldviews, truths and untruths, and to cement them or lock them in, foreclosing future contestation, reflection, and improvement.' The company already disclaims liability for the consequences that may follow."[15] (*emphasis added)*

In May 2023, the American Association of Political Consultants (AAPC), issued a statement explaining that its board of directors had unanimously "condemn[ed] use of deceptive generative AI content in political campaigns" and noted that such communications were inconsistent with the organization's code of ethics.[16] However voluntary code of ethics will not suffice to address the serious risks generative AI systems pose to election processes.

Election interference can arise from private or public or even foreign actors. Laws must be enacted to penalize malicious uses of AI, thereby discouraging acts of misinformation and

---

0generates%20false%20and%20harmful%20narratives%20on%2078%20out%20of%20100%20topics.

[13] Axios, Exclusive: GPT-4 readily spouts misinformation, study finds, Mar. 21, 2023, https://www.axios.com/2023/03/21/gpt4-misinformation-newsguard-study

[14] Congressional Research Services (CRS), *Generative Artificial Intelligence: Overview, Issues, and Questions for Congress,* In Focus, June 9, 2023, https://crsreports.congress.gov/product/pdf/IF/IF12426

[15] CAIDP, *In Re OpenAI,* Complaint to the Federal Trade Commission (FTC), Apr. 30, 2023, pp. 1, https://www.caidp.org/app/download/8450269463/CAIDP-FTC-Complaint-OpenAI-GPT-033023.pdf

[16] *Id.* at note 14.

impersonation.[17] Federal AI legislation based on established governance frameworks would address risks arising from AI systems and preserve the integrity of the democratic processes. To be clear, there must be legal liability and accountability mechanisms for developers, deployers and users of AI systems. This cannot be achieved without federal AI legislation.

2) Mandate disclosures obligations for the use of AI in campaigns

The Committee must mandate transparency and accountability for AI-generated content to maintain democracy and informed votership in the United States. Generative AI systems can produce false information and spread a bias or opinions that do not represent the public sentiment.[18] "By intensifying the barrage of untrustworthy information, AI will presumably make voters more mistrustful, cynical and intransigent."[19]

As we also explained to the FTC, Jen Easterly, CISA Director called for stronger guardrails around new AI technologies such as ChatGPT stating that, "Countering disinformation is about to get much harder: In the near term, ChatGPT and similar chatbots powered by large language models, or LLMs, will let threat actors master a range of malicious activities, including manufacturing more believable lies at scale."[20]

Just last week, OpenAI launched the third version of DALL-E, a software for AI-generated art.[21] Now, DALLE-3 incorporates ChatGPT, another widely known and widely used OpenAI product.[22] OpenAI has also announced voice and image based verbal conversational capabilities

---

[17] CAIDP, Comments to the President's Council of Advisors on Science and Technology (PCAST) Working Group on Generative AI, Aug. 3, 2023, pp. 8, https://www.caidp.org/statements/

[18] European Parliamentary Research Services (EPRS), Artificial Intelligence, Democracy, and Elections, Briefing, Sept.2023, pp. 1, https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/751478/EPRS_BRI(2023)751478_EN.pdf.

[19] The Economist, *AI will change American elections, but not in the obvious way,* Aug. 31, 2023, https://www.economist.com/united-states/2023/08/31/ai-will-change-american-elections-but-not-in-the-obvious-way

[20] CAIDP, *In Re OpenAI,* Supplement to Original Complaint, July 10, 2023, pp. 18-19, https://www.caidp.org/app/download/8466615863/CAIDP-FTC-Supplement-OpenAI-07102023.pdf.

[21] Emilia David, *OpenAI releases third version of DALL-E*, The Verge (Sept. 20, 2023) https://www.theverge.com/2023/9/20/23882009/class-action-lawsuit-openai-privacy-dropped.

[22] Will Knight, *OpenAI's Dall-E 3 is an Art Generator Powered by ChatGPT*, Wired (Sept. 20, 2023), https://www.wired.com/story/dall-e-3-open-ai-chat-gpt/.

of ChatGPT.[23] OpenAI has announced that its new technology will be "capable of crafting realistic synthetic voices from just a few seconds of real speech."[24] In its own system card "OpenAI has acknowledged that GPT-4 will generate targeted conflict 'intended to mislead.'" In a section describing disinformation, Open AI has stated that "[G]PT-4 can generate plausibly realistic and targeted content, including news articles, tweets, dialogue, and emails."[25] *(emphasis added)*

As the European Parliamentary Research Service has warned, "The network analysis capabilities of AI can also be used to better target an audience and establish the profile of voters, in what is known as political micro-targeting. AI can dramatically increase the speed at which content is made, while also offering access to a wealth of resources. Consequently, this could give rise to entire fake-news websites posing as news outlets."[26]

False AI-generated political content also has a greater chance of harming vulnerable populations– those of low-income, low formal education, and the elderly.[27] Vulnerable populations are already being targeted during elections. For example, during the 2020 elections, conservative provocateurs targeted racial minorities and low-income groups with robocalls.[28] These robocalls contained threats to deter these groups from voting.[29] Political actors have already begun to use generative AI for their agendas.[30]The photos are surprisingly realistic. As Presidential elections become closer, false AI generation will only increase.[31]

---

[23] OpenAI Blog, *ChatGPT can now see, hear and speak,* Sept. 25, 2023, https://openai.com/blog/chatgpt-can-now-see-hear-and-speak.
[24] TechCrunch, *OpenAI gives ChatGPT a voice for verbal conversations,* Sept. 25, 2023, https://techcrunch.com/2023/09/25/openai-chatgpt-voice/.
[25] Open AI, *The GPT-4 System Card*, Mar.15 2023. https://cdn.openai.com/papers/gpt-4-system-card.pdf.
[26] *Id.* at note 18, pp. 3.
[27] Brian Kennedy et al., *Public Awareness of Artificial Intelligence in Everyday Activities*, Pew Research Center (Feb. 15, 2023), https://www.pewresearch.org/science/2023/02/15/public-awareness-of-artificial-intelligence-in-everyday-activities/.
[28] Charlene Richards, *Robocalls to voters before 2020 election result in $5 million fine*, NBC News (June 8, 2023), https://www.nbcnews.com/politics/elections/robocalls-voters-2020-election-result-5-million-fine-rcna88391.
[29] *Id.* at note 28.
[30] Tiffany Hsu and Steven Lee Myers, *A.I.'s Use in Elections Sets Off a Scramble for Guardrails*, N.Y. TIMES (June 25, 2023), https://www.nytimes.com/2023/06/25/technology/ai-elections-disinformation-guardrails.html.
[31] *Id.*

AI-generated content can be created at high volumes for a cheap price and can be convincing to the electorate. According to Georgetown University Center for Security and Emerging Technology, in experiments with AI-generated news content and authentic news content, humans were able to guess which content was authentic "at a rate *only slightly better than random chance*."[32] Researchers also found that AI-generated news articles "were nearly as persuasive as articles from real world covert propaganda campaigns. Language models could also be used to generate summary text of other articles, inflected for *ideological alignments*".[33] (*emphasis added*)

We urge the Committee to mandate disclosures when election campaigns publish and distribute AI-generated content. "Federal campaign finance law does not specifically regulate the use of artificial intelligence (AI) in political campaign advertising."[34] The current provisions of the Federal Election Campaign Act (FECA)[35] do not require the disclaimers and/or disclosures where an advertisement is created by or with AI.

Federal AI legislation (or amendments to FECA) should ban targeting and amplification techniques that involve the processing of sensitive personal data such as political opinions.

Candidates, campaign staff and political committees should be required to:

a. file ex-ante disclosures if they intend to deploy or use AI systems for creating, disseminating, or otherwise communicating with voters
b. watermark/label or implement provenance mechanisms to identify AI generated content

We further recommend amendments to Section 230 of the Communications Decency Act[36] to impose liability on communications service providers for amplifying political micro-targeting using sensitive personal data, like political opinions/beliefs, carried out through their services.

---

[32] Josh A. Goldstein et al., *Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations*, GEORGETOWN UNIVERSITY'S CENTER FOR SECURITY AND EMERGING TECHNOLOGY (Jan. 10, 2023), https://arxiv.org/pdf/2301.04246.pdf.
[33] *Id.*
[34] Congressional Research Service, *Artificial Intelligence (AI) in Federal Election Campaigns: Legal Background and Constitutional Considerations for Legislation,* In Focus, Aug. 17, 2023, https://crsreports.congress.gov/product/pdf/IF/IF12468
[35] 52 U.S.C. §§ 30101–46.
[36] 47 U.S. C. § 230.

3) Implement and accelerate public literacy and fact-checking mechanisms[37]

This Committee can also take guidance from the UNESCO Recommendation on the Ethics of Artificial Intelligence[18] on increasing media and public literacy:

> The development of AI technologies necessitates a commensurate increase in data, media, and information literacy as well as access to independent, pluralistic, trusted sources of information, including as part of efforts to mitigate risks of misinformation, disinformation and hate speech, and harm caused through the misuse of personal data.

The companies that deploy Generative AI systems must establish fact-checking mechanisms and promote the use of trusted sources. To augment the trustworthy design of generative AI, there should be a collaboration with media organizations and companies to establish rigorous fact-checking processes, ensuring the dissemination of accurate information and comprehensive AI education and media literacy programs as recognized by UNESCO. In addition to accountability practices, the public must be able to recognize and evaluate AI-generated content effectively, fostering resilience against manipulation.

4) Urge the Federal Trade Commission to Move Forward the OpenAI investigation as expeditiously as possible.

The FTC has opened the investigation of OpenAI we requested.[38] This is clearly a positive development, but the FTC needs to prioritize this investigation. It took two years from the time we filed similar complaints with the FTC concerning Google and Facebook before there was a settlement.[39] We can't wait that long this time. AI products are evolving rapidly and being

---

[37] CAIDP, *Comments to the President's Council of Advisors on Science and Technology (PCAST) Working Group on Generative AI*, Aug. 3, 2023, pp. 5, https://www.caidp.org/statements/.

[38] Cecilia Kang and Cade Metz, *F.T.C. Opens Investigation Into ChatGPT Maker Over Technology's Potential Harms*, N.Y. TIMES, July 13, 2023, https://www.nytimes.com/2023/07/13/technology/chatgpt- investigation-ftc-openai.html; John D. McKinnon and Ryan Tracy, *ChatGPT Comes Under Investigation by Federal Trade Commission,* Wall Street Journal, July 13, 2023, https://www.wsj.com/articles/chatgpt-under-investigation-by- ftc-21e4b3ef.

[39] Federal Trade Commission, Facebook Settles FTC Charges That It Deceived Consumers By Failing To Keep Privacy Promises, Nov. 29, 2011, https://www.ftc.gov/news-events/news/press-releases/2011/11/facebook-settles-ftc-charges-it-deceived-consumers-failing-keep-privacy-promises.

Center for AI and Digital Policy

deployed downstream in consumer facing services. AI products will certainly have an impact on the 2024 elections in the United States and around the world.

The Federal Trade Commission must act on our pending OpenAI complaint.

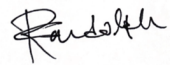
We ask that this statement be included in the hearing record.

Thank you for your attention to our statement. We would welcome the opportunity to speak with you further or to testify in a future hearing.

Sincerely,

Merve Hickok
CAIDP President

Marc Rotenberg
CAIDP Executive Director



Christabel Randolph
CAIDP Law Fellow

Brianna Rodriguez
CAIDP Law Fellow

Open Source Election Technology Institute Inc.
530 Lytton Avenue, 2nd Floor
Palo Alto, California 94301 USA
+1.650.600.1450
hello@osetinstitute.org

OSET INSTITUTE

TRUST THE VOTE PROJECT

Wednesday, October 4, 2023

Hon. Senator Klobuchar
Chairwoman
c/o Elizabeth Farrar
Majority Staff Director

Hon. Senator Fischer
Ranking Member
c/o Jackie Barber
Minority Staff Director

U.S. Senate Committee on Rules & Administration
United States Senate
Washington, DC 20510

via: eMail

RE: Comments Submission on Sep 27, 2023 Hearing: AI and the Future of Our Elections

May it please the Chair, Ranking Member & the Committee Members:

My name is Gregory A. Miller, and I have been authorized by our Board of Directors to submit the attached written comments regarding your important Hearing a week ago on AI and the Future of Our elections on behalf of the OSET Institute, Inc.—a 501(c)(3) non-profit non-partisan election technology research organization headquartered in Palo Alto, CA with over seventeen years of experience at the intersection of election technology design and cyber-security.

I, together with review by our Chief Legal Officer, also undersigned on this transmittal letter, hereby deliver the comments on the following pages to the U.S. Senate Committee on Rules & Administration for the record. We have made every effort to provide accurate information to the best of our organization's collective knowledge and experience.

We appreciate the invitation to submit these comments. We hope this will help inform the Committee's on-going work on this urgent topic.

Respectfully Submitted,

Gregory A. Miller
Co-Founder &
Chief Operating Officer

Christine M. Santoro, Esq.
Chief Legal Officer &
Corporate Secretary

Before the
UNITED STATES SENATE COMMITTEE ON RULES & ADMINISTRATION

In the Matter of	)	COMMENTS SUBMISSION
	)	
AI AND THE FUTURE	)	Wednesday, September 27, 2023
	)	
OF OUR ELECTIONS	)	301 Russell Senate Office Building
	)	
	)	

PUBLIC COMMENTS SUBMISSION

UNDERSIGNED ASSOCIATES OF THE OSET INSTITUTE
COMMENTS REGARDING AI AND THE FUTURE OF OUR ELECTIONS

Introduction

Chairwoman Klobuchar, Ranking Member Fischer, and members of the Committee, it is a privilege for us to respectfully submit the following comments in response to your Hearing on AI and the Future of Our Elections. The OSET Institute, Inc. is a 501(c)(3) non-profit non-partisan election technology research organization headquartered in Palo Alto, CA with over seventeen years of experience at the intersection of election technology design and cybersecurity, and in the past eighteen months this has included the application and impact of artificial and augmented intelligence.

The authors of this submission are the Institute's Chief Operating Officer, and principal election technology policy strategist, Gregory Miller[1] — a computer scientist and intellectual property lawyer with over 40 years of experience, and E. John Sebes,[2] the OSET Institute's Chief Technology Officer, with nearly four decades of experience.

[1] https://www.osetinstitute.org/gregory-miller

[2] https://www.osetinstitute.org/john-sebes

Just a year ago, there was little conversation at the time about generative artificial intelligence (AI), let alone its implications on elections. Today, hardly a day goes by without a news headline involving AI. Given the unprecedented pace of innovation, today we believe AI presents a serious threat, not to democracy itself, but as an intense and insidious amplifier of pre-existing threats to the administration of democracy.[3] While we believe there is good potential for beneficial application of AI to the administration of elections, in the wrong hands, AI could be used to disrupt, if not subvert electoral processes.

You heard useful testimony from the witnesses in this Hearing, and in particular, the comments by Secretary Simon were on-point with any comments we would offer about the impact and implication. We need not rehash those or his on-point remarks about how elections officials can best address the situation in the very short term.

Our comments hereunder focus on recommendations for how the Rules Committee can proceed to delve deeper into the issues and how to pursue the necessary fact-finding and related considerations in order to address the challenging policy issues and potentially fashion legislative initiatives at the federal level. For the Rules Committee, of course, this would be to ensure more stability in our electoral proceedings, perhaps as early as the 2024 cycle (though we're not optimistic given current Congressional conduct and timing, so more likely 2025 and beyond). For elections, legislative initiatives must specifically include requisite additional election security funding to provide the resources necessary to combat the AI impact on electoral processes. An excellent example of one nonprofit effort in this regard, "AI&You," was reported in POLITICO contemporaneously with the Hearing.[4] Initiatives like this and ones within government deserve funding to continue and expand where necessary.

1. Summary Recommendations

Hearings like this one are a fine and important first step in building the institutional knowledge necessary to create effective legislation. Below we offer some recommendations on how to pursue the subject-matter and the experts necessary, as the Committee continues to grapple with public policy on the impact of AI in the future of elections.

1. We encourage Congress to call upon the Federal Election Commission (FEC), The Election Assistance Commission (EAC), The Department of Homeland Security (DHS), Cybersecurity Infrastructure Security Agency (CISA) and other relevant agencies, such as the Office of the Director of National Intelligence (ODNI), the FBI, and DOJ to closely monitor the use of AI technologies during the entire election cycle, starting later this fall in order to catalog impacts, which will more clearly inform legislative initiatives, which we assume are more likely to be implemented as a consequence of and after the 2024 general election.

---

[3] We consider the "democracy administration" writ-large to mean all things concerning how a democracy is conducted, which includes the entire political ecosystem, plus the administration of government (in ours, the conduct of all three branches), plus political campaigns, and of course, election administration. Accordingly, election administration is a subset of the larger topic of administering a democracy.

[4] See: https://www.politico.com/news/2023/10/03/ai-campaigns-nonprofit-misinformation-00119579/ AI&You educates marginalized communities about the benefits and risks of artificial intelligence (AI) in their daily lives. See: https://aiandyou.org/ and specifically, https://aiandyou.org/elections/

2. Given the depth, breadth, and rate of change in AI technology, Congress should establish a central source of expertise and advisory on all aspects of AI in general; not just the security impacts to government, but the opportunities, applications, research, and impacts in all respects. In our professional opinion, the National Institute of Standards and Technology (NIST) would be the right agency for this responsibility. NIST can provide education and advice to all federal agencies attempting to figure out the impact and implications of AI. Alternatively, for this resource Congress could look to the National Science Foundation (NSF), which is certainly on point with AI research, although the focus on research, rather than promulgation of education, information and standards (like NIST) may be a barrier to their effective assistance.
3. Shifting from awareness, knowledge, and decision-support to addressing the implications of AI, there is a need that is unlikely to be addressed without federal government support: a narrow and currently impossible kind of AI agent for government and public benefit. The public needs Natural Language Agents (NLA) that accurately provide people with information from a specific knowledge base of real use to the person. We're referring to information that is large complex (e.g., 50 states and 5 territories, with ~10,000 jurisdictions) of elections processes and regulations) and daunting to individuals to navigate, but trivial for large scale machine learning.
    3.1. This is not unprecedented; we encourage the Committee to inquire about specific NLA projects underway such as the Department of Veterans Affairs and other federal agencies.[5]
    3.2. OSET Institute CTO, John Sebes has laid this out in a 4-part Blog series[6] that explains why the commercial sector has no business incentive to tackle this problem.
    3.3. The effort to support elections is modest in scale, as the base model would be trained on the narrow knowledge base of election administration, and just enough to deliver accurate assertions, but not built on the current full-breadth ("large language") models that lead to inaccurate information (and as the science refers to it, "hallucinations").
    3.4. Importantly, these models (regardless of domain) have a very low tolerance for inaccuracy and requires training and testing employed by subject matter experts and not simply AI training experts.
    3.5. And these systems must provide evidence of authoritative source material behind the statements to users.
    3.6. Accordingly, the question becomes, "*How best to determine, create, and make available these public benefit AI resources?*" This leads us to our last observation and question for policy-makers for which we do not yet have a solid answer.
4. Given this need across government service sectors and for the public benefit, how is the technology going to be developed for cost effective feasible use by ordinary organizations?

---

[5] A growing number of government agencies are using NLA-based solutions to improve investigations in critical areas such as law enforcement, defense, and intelligence. For example, the DoD's Deep Exploration and Filtering of Text (DEFT) program uses NLAs to uncover connections implicit in large text documents. See: https://cdt.org/insights/more-than-meets-the-ai-a-look-at-some-trends-in-federal-ai-inventories/

[6] See: https://www.osetinstitute.org/blog

4.1. While commercial AI vendors are expanding capacity and Band-Aids against hallucinations, they are not allocating capacity to address these far simpler needs for very high accuracy with very low tolerance for errors, hallucinations, or lies, because there simply is insufficient commercial incentive to do so. That leaves open the question: "*If not industry, then who*?"

4.2. The OSET Institute offers some initial steps in addressing this in the blog series mentioned above that presents the research agenda and opportunity, and the Office of CTO is pursuing research now on the development of safe small language models (SSLMs). Further, the Office of CTO issued a statement of research principles this past summer.[7]

4.3. Finally, the OSET Institute, in collaboration with a significant and broad coalition of stakeholders is organizing an Elections & AI Security Initiative for 2024. We would be honored to share more about the initiative and our current research underway at the Institute. For reference on the OSET Institute and its work in general, please contact Senator Wyden's office and Chris Soghoian.[8]

2. Recommended Strategies for Continuing Committee Investigation

For the Committee and its staff, we believe it is worth taking a step back to consider some strategies for proceeding into further investigation on the subject of AI and its impact. We believe these considerations are equally useful for any committee considering the impact of AI on their area of responsibility.

1st: Clarify the Subject Matter

The Committee should separate out the current "AI" systems into two parts:

1. Systems that are based in a large-language model ("LLM") used to make natural-language applications ("NLA"); and
2. The remaining wide range of machine-learning ("ML") techniques that are rapidly maturing and entering use in many ways.

2nd: Focus on Augmented Intelligence, Separate From Generative AI

We encourage any Committee investigating AI impact to acquire testimony from those thought-leaders at the forefront of machine learning techniques and applications. This should include both academics who have been refining ML, and those who are early adopters of systems based on that work.

There's a range, but here are two for example: medical diagnostics, and financial fraud detection. These are probably better described as ML-powered "augmented intelligence" systems; the machine intelligence being used to augment human intelligence (*and vice versa*). ML helps people perform their job better when there are large volumes of data to analyze. The key challenges to explore are two:

---

[7] See: https://www.osetinstitute.org/research/2023/0707/ai-elections-principled-research

[8] For a quick snapshot of the Institute's work and activities, please see: https://bit.ly/OSET2023-q1

1. Issues of accuracy, false positives, and false negatives; and
2. The necessity for humans to be in the process of creating findings and planning actions, but where there is human failure to detect an error made by the ML assistant that requires rectification before further action.

3rd: Acquire Testimony from the Most Useful Subject Matter Experts

We encourage any Committee investigating AI impact to focus testimony narrowly on LLM-based systems; specifically, the current LLMs that are polluted by their vast scope including a huge chunk of human text records of fact and fiction, decency and hatred, the whole range of humans good bad and ugly.

Computer security expert Bruce Schneier recently co-wrote a helpful article[9] that identifies 3 categories of "AI experts" who are busy building brands around a theme:

1. "Doomsayers:" We do not believe you need testimony from these type of experts about how badly broken or at risk the current LLMs are. Take “judicial notice” that there are "lies, damn lies, and hallucinations" in current LLMs and move on.
2. “Reformers:" These experts are a better source of insight about how using current polluted LLMs can have an array of practical consequences relevant to public policymaking.
3. “Warriors:" These are experts who claim that current LLMs can be improved to do wonderful things. We recommend not inviting them. The Committee can read a few of these types of experts’ statements and remember: there is a good change that they are wrong, and their optimism is born of greed.

Instead of Doomsayers or Warriors (*and to a lesser extent, Reformers*), seek-out and request testimony from technical professionals who can explain in simple terms how LLMs are built, and why it is as unavoidable as the law of gravity that current LLMs are and will be polluted. The undersigned to these comments would be honored to assist you in that sourcing.

The experts you actually need can explain mitigation methods like human guided training and necessary guardrails, or multi-agent voting on responses, and more. These experts (similar to the Reformers) will explain all the efforts to put current polluted LLMs to use despite the innate flaws.

Regardless, remember this: all of these mitigations are to LLM systems, as anti-virus techniques are to computers. In other words, they are reactive patches to limit the consequences of using fundamentally flawed systems. To this end, consider one of the latest trends in AI is adversarial testing of Natural Language Agents (NLAs), to reveal bypasses to limited safety mechanisms — which is very similar to demonstrating vulnerabilities of limited security mechanisms in conventional software.

4th: Ignore the Sexy; Focus on the Substantive

Separate all of the exciting and provocative capabilities of LLM-based systems from the simpler, but very valuable class of Natural Language (NL) systems with real benefits; that is, NL agents that are trained on a specific body of textual information in one narrow area (or "domain") of knowledge.

---

[9] See: https://www.nytimes.com/2023/09/28/opinion/ai-safety-ethics-effective.html

These systems are built for the sole purpose of communicating with a person to help the person navigate this body knowledge and find the information that is of use to them. And for purposes of understanding the deeper implications of AI in elections, we believe the NLAs are the item to focus on more than headline-stealing announcements such as conversational chatbots and personal agents (*although the latter is a potential serious haven for trouble; see below*).

These so-called "domain specific" NL agents show the promise of this sole-purpose benefit, which can be applied to a very broad range of government use and other public-benefit systems (*and which brings us back to elections*). Yet, let's be clear: today, these systems are based or depend on the pre-existing polluted LLMs. We firmly believe this is a huge mistake — akin to building a house on sand — using tools never intended for safety-critical systems and other systems with an extremely low tolerance for inaccuracy, and nearly zero necessity for human-like characteristics.

We note above that OSET Institute's CTO has written about why these systems do not currently exist, and are not sufficiently profitable for the private sector to address. These "safe small language models" are currently non-existent, but are essential for highly reliable mission-critical government applications.

5th: Avoid Distraction with Separate Issues

Finally, we believe there are a couple of topics to avoid as the Rules Committee drills into the impact of AI on government, the administration of democracy writ large, and the future of elections. These are important issues, for sure, but largely a detour or distraction from the important work the Committee has to better understand the specific issues of AI and its impact, threat, and yes, even opportunity. However, the two topics to set aside are:

1. AI-Empowered-Disinformation. That's right, and while we understand this may be counter-intuitive, please read on: AI is nothing more than an anabolic steroid for disinformation—it does not create a new topic of concern. Disinformation in elections continues to be a serious, systemic, and difficult problem. AI simply makes it faster and more insidious (*i.e., deep-fakes and ultra-deep-fakes*). We encourage the Committee to recognize that point, continue to address disinformation as it is, and not be derailed by a notion that AI creates a new category of the problem; it simply exacerbates the existing one. There is no doubt, we have adversaries using every kind of information warfare techniques to sow malice and chaos. Their tools are advancing, including applications of AI, but the issue is not the tools, but the adversaries, the threats, risks, mitigations. You already have plenty of national security people to tell you about that — AI experts will not shed new light on that.

   However, we recommend examination of the new Democracy by Design initiative[10]—a coalition supporting a content-agnostic election integrity framework for online platforms. This coalition proposes, among other things, prohibiting all use of generative AI or manipulated media to depict election irregularities, misrepresent public figures, or micro-target voters with ads generated using personal data, as well as requiring strong disclosure standards for any political ads that feature AI-generated content.

[10] https://accountabletech.org/wp-content/uploads/Democracy-By-Design.pdf

2. AI and Intellectual Property and the Human Right to Work. Of course, we recognize that this is beyond the scope of responsibility for the Rules Committee, and yet the issue can entangle other proceedings. It is clear that current LLMs have ingested enormous amounts of I.P. without permission. That, unfortunately, cannot be "undone." At the same time, there are real policy issues about how people should be protected or empowered to continue to own the fruit of their labor and creativity—that surely includes content within the administration of democracy in general, and elections in particular. However, at base, that is not about AI itself, rather it is simply the latest example of reckless siphoning of data and knowledge with total disregard for intellectual property rights in order to make money. Stepping back, this is not very different than tracking web activities and traffic in order to sell ads. It is simply far more pervasive with far greater impact.

Recap of Recommendations

1. Obtain testimony specifically from professionals who have failed at trying to build a highly-reliable mission-critical applications on polluted LLMs.
2. Obtain testimony from experts who understand how to build language models, and about the prospects for safe small language models (SSLMs) built with just enough natural language capability to be trained on a domain of knowledge and communicate to people about it, but without the ability to hallucinate.
3. Consider testimony from technology professionals who have been working in realms parallel to AI, and have witnessed two critical mistakes in history (*described below*), and understand the very similar mistake occurring now: trying to build high-accuracy systems on top of fundamentally-flawed and dangerous LLMs.
   3.1. The First Mistake: In the early days of computing, asserting that personal computers (PCs) were great, and they could do nearly anything, so they were used to build safety-critical systems that should only perform limited and very important functions. As a result, today U.S. critical infrastructure is littered with these insecure-by-design PCs with security Band-Aids trying to perform important tasks safely, relying on a security strategy of "patch and pray."
   3.2. The Second Mistake, In the early days of networking and the dawn of the public Internet, asserting that open networks were great, they could connect anything (*to anything*); and the network is the computer, so they were used to connect nearly everything to everything. One problem: everything connected can be attacked on a network that was designed to withstand a nuclear attack, but was never designed to be inherently data secure or private, so, this required more security Band-Aids for insecure-by-design networks.

   Now we have unsafe-by-design artificial intelligence LLMs with tech-giants extolling the virtues of their use in everything, and using the same Band-Aid mentality to patch them vainly in order to build money-making systems that will fail by inaccuracy and hallucination. In the world of elections, similarly, this same mistake was made with social media; let's not make the same mistake again, this time with AI.
4. New testimony could and should provide some guidance on how government can foster the required technology research and development work that's not highly profitable, but essential

for government systems and other public benefit systems to help large numbers of people better navigate complex information bases, far beyond the ability of the limited number of experts and human navigators available.

5. All congressional committees with the appropriate jurisdiction should continue useful — but separate — investigation on the impact of AI more broadly, such as on intellectual property rights, and on the challenges of disinformation. For the Rules Committee the latter, of course, specifically in elections. To that, we advise simply accepting that AI is an "anabolic steroid" to the pre-existing challenge of disinformation, misinformation, and now mal-information.

A Closing Thought

In our Recap of Recommendations above, we observe in item 3 that historically, the computer technology sector made two mistakes—

- The first was overloading the utility of PC technology to serve purposes for which the design of PCs was never intended to support and is inadequate for such; and
- The second was overloading the utility of computer networks—specifically TCP/IP networks (namely the Internet)—to be used in a manner and for services for which its' design was never intended to support.

Of course, we'd like to avoid repeating history; however, we're less than optimistic. At this writing in the 1st week of October, the industry's latest initiative (*and seemingly making the same class of mistake*) is the AI assistant. Last week, Google, Meta, and OpenAI all launched new features for their AI chatbots that allow them to search the web and act as a type of personal assistant.

- OpenAI introduced new ChatGPT features that enables the user to have a verbal conversation, allowing them to receive (*synthetically*) verbal responses to their spoken questions, with the system leveraging search engines.
- Google's rival Bard, has been integrated with Google's services infrastructure, including Calendars, Docs, Gmail, Maps, and YouTube. The objective is for the individual to use the chatbot to ask questions about their own content (e.g., searching their eMail or managing their calendar.)
- Meta also announced that it will front everything with AI chatbots. Users will be able to ask AI chatbots on Messenger and Instagram.

We believe this is a very risky move, considering the known limitations of the current technology. The fact is, the AI tech-giants have not solved the persistent problems of the LLMs, including their tendency to simply make things up (*hallucinate*). Of even more concern to us is that these LLMs are a privacy and security nightmare. We strongly encourage the Rules Committee (*and other Committees of jurisdiction for other domains*) to investigate the risk of putting this flawed technology in front of millions of users and empowering AI models to consume sensitive personal information—such makes everyone more vulnerable to massively scalable attacks (*e.g., hacks, scams phishing, etc.*). Obviously, this goes far beyond Elections to a wide spectrum of applications.

For example, we encourage the Committee to inquire about one issue in particular: a type of attack called indirect prompt injection. It's very easy to perform, and there is no known remedy. In this

type of attack, a website is altered by incorporating hidden text meant to change the AI application's behavior when coming across the content. Attackers could then direct unwitting users to these websites, where subsequently their personal AI agents would consume these subliminal prompts.

With this new generation of personal AI agent plugged into one's eMail, social media, and searching activities, the opportunities for a new generation of hacking and attacks are nearly limitless in capability, complexity, and undetectability.[11] The point is, there is much more to be investigated and considered, and the Committee should set forth to ferret out these issues in order to gain a fuller picture of the policymaking challenges and required legislative architecture.

We hope these remarks are of help to the Committee's on-going work and we would be honored to provide further advise, assistance, or response to any questions or comments.

Respectfully Submitted,

Gregory A. Miller, JD
Election Technology Policy Strategist
OSET Institute, Inc.
Portland, Oregon USA
https://www.osetinstitute.org/

E. John Sebes
Chief Technology Officer
OSET Institute, Inc.
Palo Alto, California USA

[11] "Prompts" are the inputs or queries that a user or a program gives to an LLM AI, in order to elicit a specific response from the model. Prompts can be natural language sentences or questions, or snippets of software code or commands, or any combination of text or code, depending on the domain and the task. Indirect prompt injection attacks, are concealed instructions that make an AI system behave in unintended ways. They are rapidly emerging as a significant cybersecurity concern. Indirect prompt injections involve a third-party providing instructions to the LLM through sources like websites or PDFs. Simply put, these prompt injections (if nefarious or unauthorized, then considered an "attack") take advantage of weaknesses in the chatbot's prompting system. Users enter commands that trigger an unrestricted mode, causing the AI to disregard its built-in safety measures and guidelines. This enables the chatbot to respond without the usual restrictions on its output.

For more background, see: https://www.wired.com/story/generative-ai-prompt-injection-hacking/. There is considerable activity and interest in these prompt attacks also known as Jailbreaking. Jailbreaking in AI, as described above, is the act of cleverly convincing chatbots to bypass restrictions, revealing their capabilities and limitations. AI jailbreaking is a hobby and research field, testing the boundaries of AI and serving as a form of quality assurance and safety testing. In fact, there is a website focused on "Jailbreaking;" see: https://www.jailbreakchat.com/

215 Pennsylvania Avenue, SE • Washington, D.C. 20003 • 202/546-4996 • www.citizen.org

September 26, 2023

U.S. Senate Committee on Rules & Administration
Sen. Amy Klobuchar, Chairwoman
305 Russell Senate Office Building
Washington, D.C. 20510

Dear Chairwoman Klobuchar:

Public Citizen strongly supports legislative and regulatory efforts to rein in potential abuses of deliberately deceptive Artificial Intelligence (A.I.) content in campaign communications that are intended to cause harm to a candidate and deceive voters, commonly referred to as "deepfakes." We are particularly supportive of the legislation introduced by Sen. Klobuchar, one of which would require transparency of the use of A.I.-generated content in ads, and the other which would prohibit the use of deepfakes.

Public Citizen has petitioned the Federal Election Commission for rulemaking on deepfakes under the "fraudulent misrepresentation" law (52 U.S.C. §30124). However, this law only applies to candidates misrepresenting their opponents. Further legislation is required to extend the restriction against deliberately deceptive A.I.-content to super PACs and outside groups as well, which is what the Klobuchar legislation would achieve.

Extraordinary advances in artificial intelligence now provide political operatives with the means to produce campaign ads and other communications with computer-generated fake images, audio or video of candidates that appear real-life, fraudulently misrepresenting that what candidates say or do. Generative artificial intelligence and deepfake technology – a type of artificial intelligence used to create convincing images, audio and video hoaxes – is evolving very rapidly. Deceptive deepfakes are already appearing in elections and it is a near certainty that this trend will intensify absent action from the Federal Election Commission and other policymakers:

- In Chicago, a mayoral candidate in this year's city elections complained that AI technology was used to clone his voice in a fake news outlet on Twitter in a way that made him appear to be condoning police brutality.[5]
- As the 2024 presidential election heats up, some campaigns are already testing A.I. technology to shape their campaign ads. The presidential campaign of Gov. Ron DeSantis, for example, posted deepfake images of former President Donald Trump hugging Dr. Anthony Fauci.[6]

As the technology continues to improve, it will become increasingly difficult and, perhaps, nearly impossible for an average person to distinguish deepfake videos and audio clips from authentic

media. It is an open question how well digital technology experts will be able to distinguish fakes from real media.

A blockbuster deepfake video with this kind of fraudulent misrepresentation could be released shortly before an election, go "viral" on social media, and be widely disseminated, with little ability for voters to determine that its claims are fraudulent.

Public Citizen encourages the Senate Rules Committee to proceed expeditiously with legislation that will curb the potential abuses of deliberately deceptive deepfakes in campaign communications. The 2024 election is already upon us and, without appropriate regulation, deepfakes in campaign communications are likely to become common in campaign attack ads.

Sincerely,

Craig Holman, Ph.D.
Government affairs lobbyist
Public Citizen
215 Pennsylvania Avenue, SE
Washington, D.C. 20003

Lisa Gilbert
Executive Vice President
Public Citizen
1600 20th Street, NE
Washington, D.C. 20009


VOTE
As a Matter of Fact
The Harms Caused by
Election Disinformation
Common Cause
Education Fund

About Common Cause Education Fund

The Common Cause Education Fund is the research and public education affiliate of Common Cause, founded by John Gardner in 1970. We work to create open, honest, and accountable government that serves the public interest; promote equal rights, opportunity, and representation for all; and empower all people to make their voices heard in the political process.

Acknowledgments

This report was co-authored by Sylvia Albert, Yosef Getachew, Jesse Littlewood, Beth Rotman, Paul S. Ryan, Emma Steiner, and Jonathan Walter and published by Common Cause Education Fund.

We also thank Karen Hobert Flynn, Michael Copps, Marilyn Carpinteyro, Susannah Goodman, Stephen Spaulding, and Aaron Scherb for guidance and editing; Melissa Brown Levine for copyediting; Kerstin Vogdes Diehn for design; and Scott Blaine Swenson, David Vance, and Ashlee Keown for strategic communications support. Finally, special thanks to Austin Graham, legal counsel at the Campaign Legal Center, for consultation regarding exemplary state campaign finance disclosure laws.

Special thanks are due to the thousands of volunteers of Common Cause Education Fund's Stopping Cyber Suppression program who spent tens of thousands of hours monitoring social media for disinformation that could disenfranchise voters. Many of the examples in this report came from volunteers in this program.

Knowledge planted in truth grows in truth.

—Aberjhani[1]

CONTENTS

EXECUTIVE SUMMARY

In America, whatever our background, color, or zip code, we value our freedom. Generation after generation has fought for the freedom to have a say in decisions that impact our lives—the freedom to participate fully in our country. But in recent years, a small faction has grown increasingly skilled at spreading lies about our elections, lies that targeted Black communities and other communities of color to suppress their votes, lies that fueled a deadly attack on our Capitol in January 2021 to disrupt the peaceful transfer of power, lies that threaten to suppress votes and undermine public confidence in future elections. This intentional use of false information to affect the participation of voters in elections is known as "election disinformation."

The United States is at a critical juncture. More than 1 in 3 U.S. residents—and nearly 80% of Republicans—wrongly believe that President Joe Biden did not legitimately win the election.[2] And a majority say they "do not have confidence that elections reflect the will of the people."[3] Donald Trump's Big Lie is working, and we have to respond. Just as we came together last year, rising up to vote safely and securely in record numbers during a global pandemic, we must now rise up to stop election disinformation efforts in future elections. This report is a game plan for success.

The United States is at a critical juncture. More than 1 in 3 U.S. residents—and nearly 80% of Republicans—wrongly believe that President Joe Biden did not legitimately win the election.

Election disinformation is not a new phenomenon. Indeed, for nearly two decades Common Cause has been monitoring and working to stop election disinformation as a part of the national Election Protection coalition.[4] As explained in our 2008 report *Deceptive Practices 2.0*,[5] co-authored with the Lawyers' Committee for Civil Rights Under Law and the Century Foundation, false or misleading information about the voting process, often targeting Black communities and intended to suppress votes, has historically been disseminated via flyers and "robocalls." But by 2008, disinformation was beginning to move to email and websites. And as explained in our 2012 report *Deceptive Election Practices and Voter Intimidation*, disinformation tactics continued to evolve: "Over time, they have become more sophisticated, nuanced, and begun to use modern technology to target certain voters more effectively."[6] The volume and sophistication of online disinformation, particularly via social media platforms, continued to rise in 2016 and 2018 preceding a veritable explosion of election disinformation throughout the 2020 election cycle.

As online election disinformation has increased, Common Cause Education Fund's commitment to monitoring and stopping it has likewise increased. During the 2020 election cycle, we led an Anti-Disinformation Working Group of the Election Protection coalition, hired experienced disinformation analysts, and trained dozens of partner organizations and thousands of volunteers in disinformation monitoring. We catalogued more than 3,000 disinformation posts in our election disinformation database, requested the removal of posts from social media platforms when they violated corporate policies, developed messaging to "pre-bunk" the disinformation, and dis-

seminated accurate voting and election information in partnership with the Election Protection coalition.

We continue our election disinformation work in the 2021 "off-year" elections and prepare for elections in 2022 and beyond. As part of our plan to combat election disinformation, Common Cause Education Fund has prepared this report to explain the problem of election disinformation in detail and propose commonsense public and corporate policy reforms to reduce the harmful impacts of election disinformation in future elections.

Section 1 provides an overview of election disinformation, explaining what it is, how it's being spread, and who is spreading it. Understanding the threat of election disinformation is the first step toward eliminating the threat. Common examples of election disinformation include communications providing the wrong election date, bogus election rules, voter intimidation, untrue claims about election integrity/security, and untrue claims post-election about results. Today, the most common means of disseminating disinformation include social media platforms like Facebook and Twitter, junk websites, mainstream media like Fox News, search engines like Google, email, text messages, and robocalls.

For example, in the spring of 2020, former president Donald Trump repeatedly and falsely claimed that mail-in ballots were less secure and part of a plan to rig the election against him and Republicans, generally. Supporters of Trump then repeated these claims, driving a false narrative of voter fraud. Experts analyzed social media and found a massive 3.1 million mentions of disinformation about voting by mail between January 2020 and September 2020.[7] Election disinformation is spread before, during, and after Election Day. The 2020 false voter fraud narrative fed a post-election false narrative that the election was "stolen" from Trump (i.e., Trump's "Big Lie"), giving energy to the so-called Stop the Steal movement and the deadly January 6 insurrection. These false narratives persist today, undermining public confidence in future elections and being used as justification for new voter suppression laws in states around the nation.

Section 2 details current federal and state laws regulating election disinformation—voting rights, campaign finance, communications, consumer protection, media literacy, and privacy laws—and the shortcomings of current laws. These laws are tools we must use to thwart election disinformation efforts. A primary purpose of election disinformation is to suppress and sometimes intimidate voters. Federal law and laws in nearly every state contain provisions explicitly prohibiting voter intimidation, with many of these laws being rightly interpreted as prohibiting election disinformation. A handful of states have enacted laws explicitly prohibiting knowingly disseminating materially false information about the time, place, or manner of elections with the intent to impede voting. Such laws play an important role in fighting election disinformation and should be widely enacted and enforced.

Several other bodies of law are also critically important to combating election disinformation. Strong campaign finance disclosure laws can shine the light of publicity on those seeking to undermine our elections from the shadows. Federal communications law provides digital platforms with legal protections to moderate content online without fear of liability and directly impact election disinformation. Consumer protection laws can protect us from deceptive data collection and data security breaches and have been used to punish some who have contributed to the spread of disinformation. State media literacy laws can help people build the skills necessary to discern

fact from opinion and fiction, news from infotainment, and real information from disinformation. And state privacy laws can protect personal data to prevent bad actors from precision targeting of election disinformation. All of these laws can play a part in effectively regulating and deterring election disinformation.

Section 3 describes the civic integrity policies of some of the largest social media companies, the policies Facebook, Twitter, and YouTube have put in place to address abuses of their platforms for the dissemination of election disinformation. Across all these platforms, content that is misleading about how to participate in elections is actionable and should be removed, including misleading information about the date or time or requirements to participate in an election and statements advocating for violence because of voting, voter registration, the administration, or outcome of an election.

Unfortunately, current civic integrity policies have significant loopholes that have allowed content contributing to voter suppression and election disinformation to remain on social media platforms. Part of the problem is frequent changes to civic integrity policies. For example, during the 2020 election cycle, Facebook changed its election-related misinformation policies 21 times, Twitter changed its policies 16 times, and YouTube changed its policies 12 times.[8] Most of these changes involved adding, subsequently rolling back, and then reinstating new rules concerning key issues like mail-in voting fraud or false victory claims.[9] Another problem is a lack of transparency regarding how well these policies were enforced and their impact on election misinformation. Making matters worse, Facebook and Twitter have now rolled back policies they put in place during 2020 and stopped enforcing existing policies to the degree they did during the 2020 election cycle. Our research shows that there are many pieces of content being left on the platform that would have been taken down months ago. Social media platforms must take additional steps to strengthen their policies on combating content designed to undermine our democracy.

Finally, **Section 4** identifies gaps in current laws and policies that have allowed election disinformation to flourish and recommends reforms to better enable us to fight back against election disinformation.

Federal and State Voting Rights Reforms

The **single most important tool to stop election disinformation is a statute prohibiting knowingly disseminating materially false information regarding the time, place, or manner of elections or the qualifications or restrictions on voter eligibility, with the intent to impede voting.** While the U.S. Department of Justice (DOJ) and some state law enforcement agencies have interpreted existing civil rights laws, specifically those prohibiting voter intimidation or interference, as applying to election disinformation via social media platforms, this application of the law has not yet been thoroughly tested in courts. **Congress and state legislatures should remove any doubt by enacting statutes prohibiting such false election speech**, with both criminal and private civil remedies and a mandate that the government corrects materially false election information.

Federal and State Campaign Finance Reforms

Congress and state legislatures must update campaign finance disclosure laws for the digital age. Strong campaign finance disclosure laws are key to curbing the harmful impacts of election

disinformation. Unfortunately, federal campaign finances laws and the laws of most states are out-of-date, lacking clear mandates and guidance for **"paid for by" disclaimers on digital advertising**, and effective **provisions shining a light on money transferred between groups to evade disclosure**.

Federal and State Privacy Law Reforms

Congress should pass comprehensive data privacy legislation to protect consumers from the abusive collection, use, and sharing of personal data. At a minimum, federal legislation should (1) require companies to minimize the data they collect; (2) prohibit predatory and discriminatory data practices on the basis of protected characteristics with respect to access to credit, housing, education, employment, and public accommodations; (3) provide for fairness in automated decision-making; (4) grant a private right of action to allow consumers to sue companies that violate their privacy rights; and (5) define permissible and impermissible uses for collecting, sharing, and using personal data.

State legislatures should pass comprehensive consumer privacy laws similar to the California Consumer Privacy Act (CCPA) of 2018 to provide consumers with the right to know about the personal information a business collects about them, the right to delete personal information collected from them, the right to opt out of the sale of their personal information, the right to nondiscrimination for exercising their CCPA rights, and the right for consumers to sue businesses for certain data breaches. And states should go further than the CCPA by including privacy legislation requirements that limit what data entities can collect and how that data can be used, as well as civil rights protections that ensure fairness in both automated decision-making and prohibitions on the use of personal data to discriminate on the basis of race, gender, religion, national origin, sexual orientation, gender identity, disability, familial status, biometric information, or lawful source of income, as well as a robust private right of action for consumers whose rights are violated.

State Media Literacy Law Reforms

State legislatures should experiment with best practices around media literacy and hold convenings with organizations like PEN America that are already engaged in the issue and offering media literacy training to the public to put together a set of agreed-upon principles on which to develop legislation.

Federal Media Law Reforms

Congress should enact legislation strengthening local media and protecting public access to high-quality information about government, public safety, public health, economic development, and local culture, such as the Future of Local News Act, which would create a committee to study the state of local journalism and offer recommendations to Congress.

Congress should pass legislation to protect researchers' and watchdog journalists' access to social media data, enabling researchers to study social media platform practices without fear of interference or retaliation from social media companies.

Congress should pass legislation to prohibit online platform discriminatory algorithms and to create greater transparency about how these algorithms operate.

Federal and State Executive and Regulatory Agency Reforms

The White House must play a leading role in combating election disinformation, including by issuing an executive order directing federal agencies with enforcement, rule-making, and investigatory authorities to use these capabilities in combating election disinformation. The White House should also create a federal interagency task force that would identify tools to combat election disinformation and harmful online speech. **Governors in states around the nation should likewise lead efforts to combat election disinformation on the state level.**

The DOJ and state law enforcement agencies should use all existing statutory and regulatory tools (e.g., existing anti-voter intimidation laws) to more aggressively prosecute those who use disinformation to intimidate voters and interfere with their voting rights.

The **Federal Trade Commission should expand the scope of its rule-making and enforcement practices** to more effectively regulate unfair and deceptive commercial data practices and conduct workshops and issue informal guidance on how social media platforms can provide greater transparency in their content moderation practices.

The **Federal Election Commission and state election agencies** should better use all available rule-making and enforcement authority to implement effective campaign finance disclosure requirements for online political advertising.

Social Media Corporation Policy Reforms

While self-regulation will never alone be sufficient, **social media companies must do a better job curbing the spread of disinformation by strengthening their policies** around combating content designed to undermine our democracy. We make specific recommendations in this report for how social media companies can improve their efforts to provide users with authoritative information regarding voting and elections, reduce the spread and amplification of election disinformation, and provide greater transparency concerning their content moderation policies and practices.

Democracy depends on free and fair elections. Together, we must educate ourselves, demand rigorous enforcement of existing laws to stop election disinformation, and pass new laws to protect our right to vote and to stop a small faction from sabotaging our elections.

INTRODUCTION

For nearly two decades, Common Cause has been monitoring and working to stop election disinformation. As the volume and sophistication of online disinformation have risen in recent years, so too has Common Cause Education Fund's commitment to monitoring and stopping the disinformation.

Throughout the 2020 election cycle, Common Cause Education Fund trained thousands of volunteers, who contributed tens of thousands of hours searching for election disinformation in their own social media networks. We also hired experienced staff and contractor disinformation analysts who monitored the more extreme communities and social media platforms. Through this work, we created a database of election disinformation. We led an Anti-Disinformation Working Group of the Election Protection coalition and trained dozens of partner voter protection groups in how to monitor, analyze, and take action on election disinformation, which was particularly critical for language access (including Spanish, Haitian Creole, Arabic, and Asian and Pacific Islander languages). Last, we opened up a public "tip line" for disinformation reports at ReportDisinfo.org.

Throughout the 2020 election cycle, Common Cause Education Fund trained thousands of volunteers, who contributed tens of thousands of hours searching for election disinformation in their own social media networks.

Our in-house analyst reviewed disinformation reports from all these sources daily, documenting and cataloging them in our database, which was shared with our voter protection partners. Our analyst identified which posts were likely to be "actioned" by the social media companies based on their civic integrity policies and reported these posts to the companies, resulting in over 300 actions (labels, removals, and banning of accounts). Unfortunately, the posts on which the companies took action were only a fraction of the 3,000+ problematic election disinformation posts we added to our database, which either were not actioned when reported or were outside the companies' narrow rules for taking action.

Requesting removal from the social media platforms was just one of our program's interventions on election disinformation. We were in constant communication with more than 40 voter protection organizations, alerting them to disinformation threats and providing resources, including messaging and "inoculation" content to "pre-bunk" disinformation. We secured a partnership with PolitiFact to issue fact-checks on dubious social media content and worked with messaging experts and creative designers to make educational social media content for use by the Election Protection coalition. Volunteers and staff of Common Cause Education Fund and partner organizations likewise posted accurate voting and election information on social media through the Election Protection network—including answering questions raised by voters on their own social media posts (e.g., polling place hours, locations, rules, and regulations). We logged over 6,000 of these "voter assistance" posts **(see Figure 1)**.

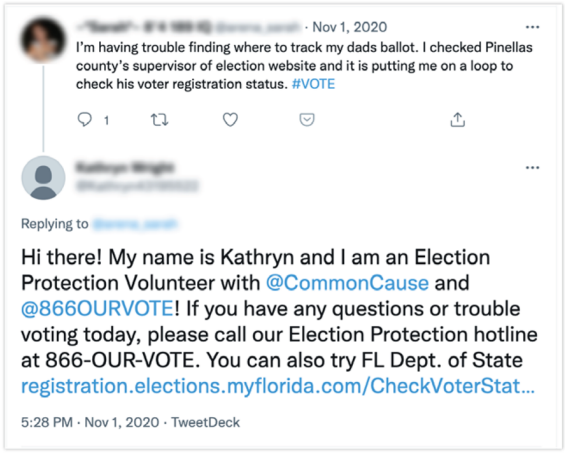


Figure 1: A "voter assistance" post by a Common Cause volunteer.

Our election disinformation monitoring and analysis program has continued during the 2021 "off-year" elections, and we are preparing for 2022. Our focus was (and remains) on **nonpartisan** election disinformation or "cyber suppression"—i.e., mis- and disinformation that could prevent voters from participating in the election and/or undermine their faith in the integrity of the electoral process. We identify and take action on disinformation from both Democrats and Republicans, left-leaning posts and right-leaning posts. However, when volunteers find disinformation about a specific candidate (e.g., disinformation about a candidate's policies, personal history, or activities), we do not catalogue or take action on the candidate-specific disinformation. Doing so would not only stretch our capacity far beyond our limits but also require us to wade into matters that are often more subjective and partisan, in tension with our strict nonpartisanship policy. Candidates and parties are better suited to handle candidate-specific disinformation and typically dedicate resources to doing so.[10]

This report is built on decades of experience monitoring and responding to election disinformation. **Section 1** provides an overview of election disinformation, explaining what it is, how it's being spread, and who is spreading it. Understanding the threat of election disinformation is the first step toward eliminating the threat. **Section 2** details current federal and state laws regulating election disinformation—voting rights, campaign finance, communications, consumer protection, media literacy, and privacy laws. These laws are tools we must use to thwart election disinformation efforts. **Section 3** describes the civic integrity policies of some of the largest social media companies, the policies Facebook, Twitter, and YouTube have put in place to address abuses of their platforms for the dissemination of election disinformation. Finally, **Section 4** identifies gaps in current laws and policies that have allowed election disinformation to flourish and recommends reforms to better enable us to fight back against election disinformation.

SECTION 1: ELECTION DISINFORMATION OVERVIEW

What Is Election Disinformation?

Broadly, election disinformation refers to intentional attempts to use false information to affect the participation of voters in elections. There is a long history of tactics used to disenfranchise voters, and our previous reports[11] detail how flyers, billboards, and other offline tactics are used to tell voters incorrect information that could prevent them from participating in an election. These reports also highlighted some of the emerging online digital tactics used to spread election disinformation, including email, the web, and Facebook, which were just gaining mainstream popularity.

Our earlier reports make clear that as communication methods and channels mature, malign actors adopt them in the service of election disinformation and voter suppression. In the present era, widely adopted social media, where anyone can be a publisher of content, often anonymously or semi-anonymously, has become the most effective communication method of election disinformation. Although purveyors of election disinformation are not limited to social media, where about half of us find our news, they have aggressively adopted the medium.[12]

For nearly two decades, Common Cause Education Fund has been monitoring and working to combat election disinformation through our Election Protection coalition. We witnessed a steady rise of disinformation online in 2016 and 2018, and then a veritable explosion of voting-related disinformation throughout the 2020 election cycle.

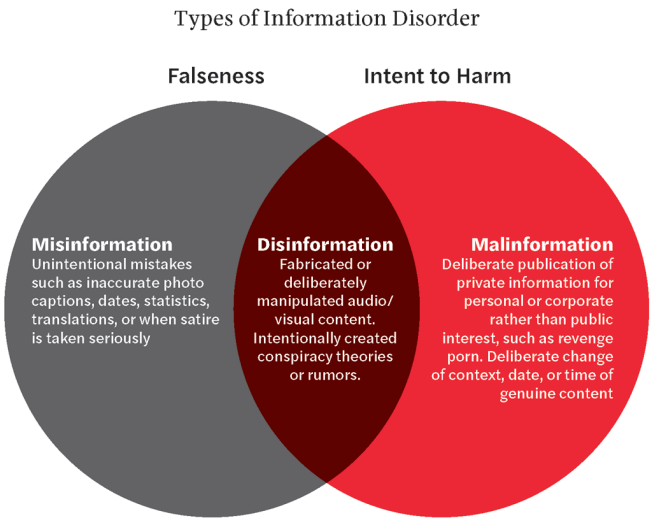


Figure 2: Three types of information disorder. Credit: Claire Wardle and Hossein Derakshan, 2017

"Information disorder" is an emerging term of art used by researchers and media experts[13] that encompasses three related terms **(see Figure 2):**

- **Disinformation** is content that is false (even if it contains some truth) and deliberately created to harm a person, social group, organization, or country.
- **Misinformation** is false information, but it is differentiated from disinformation by lacking an intent to harm any person, group, or organization.
- **Malinformation** is content that is accurate but is intentionally manipulated to cause harm, including voter suppression or voter confusion.

These terms can be accurately applied to individual pieces of *content (a flyer, poster, billboard, text, phone call, or social media post) but also encompass entire narratives (a sequence of pieces of content knitted together that creates a stronger and lasting impact, often referencing previous pieces of content—e.g., Donald Trump's "Big Lie"*[14]). Voting and elections are threatened by individual pieces of content and narratives that fit within each of these three categories of information disorder.

Misinformation

Misinformation is false information, but it is differentiated from disinformation by lacking an intent to harm any person, group, or organization. While it is less intentional, it can be equally harmful. Examples of misinformation include inaccuracies in dates or statistics or incorrectly identified photo captions. Anyone encountering the misinformation could believe it and draw conclusions from it, even if the content provider was not intending to misinform them.

One common misinformation narrative we encountered during the 2020 elections included a widely shared meme that has appeared in multiple election cycles that encourages voters to use "two stamps" when mailing back their absentee ballot under the (false) theory that the U.S. Postal Service (USPS) will ensure delivery or otherwise prioritize your absentee ballot **(see Figure 3). This is misinformation. The USPS stated it would deliver election mail, even if postage is required, without that postage.**[15] While unintentional, this misinformation perpetuated a negative view of the USPS and its ability to manage mail-in ballots and thereby suppress voting. While impression data is not available from social media platforms, some of the content we saw received thousands of shares, and several news organizations, including Reuters, *USA Today*, and PolitiFact responded with fact-checks. Requiring postage for returning voted ballots by mail is a known barrier to voter participation: not everyone has stamps at home, acquiring stamps required potential exposure to COVID during the early stages of the pandemic, and singular stamps are less likely to be available than entire books or packages (which cost more). A voter who believes two stamps are necessary to submit their mail-in ballot but doesn't have access to two stamps may choose not to vote at all. The now-pending Freedom to Vote Act would amend federal law to make clear that no postage is required for completed ballots.[16]

Disinformation

Disinformation content is false and deliberately created to harm a person, social group, organization, or country. Disinformation is deliberately and often covertly spread to influence public opinion

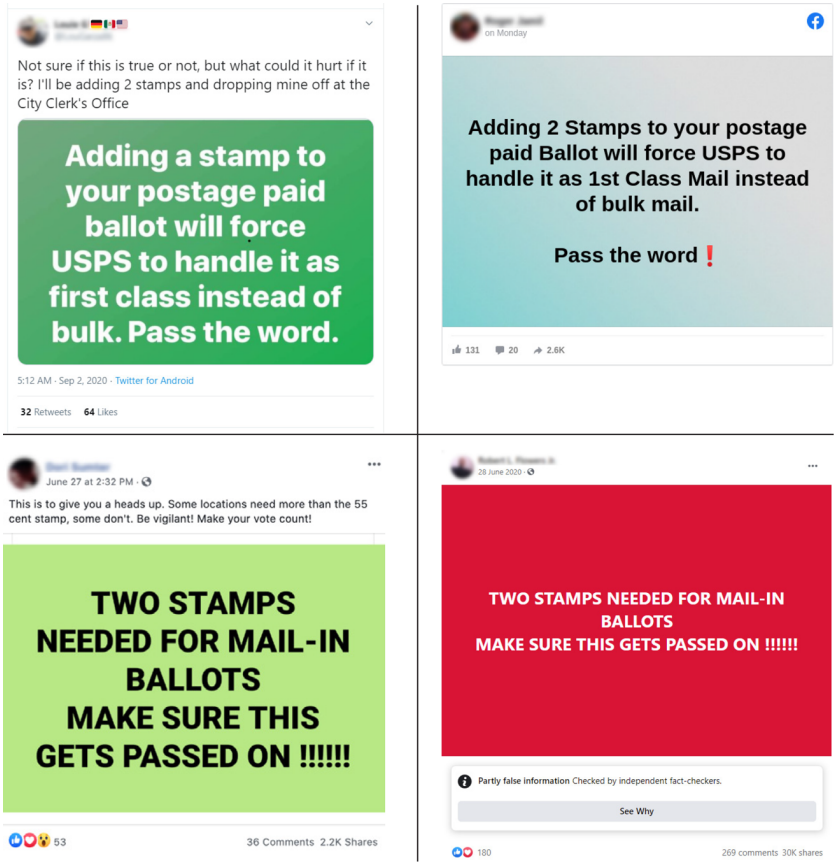


Figure 3: Information disorder posts regarding ballot postage. Credit: Claire Wardle and Hossein Derakshan, 2017

and actions, obscure or alter voting, or provide cause for outrage. Disinformation may contain some true facts, but those facts are either taken out of context or combined with falsehoods to create and support a specific intended message. An example of disinformation using true facts is when old news stories are recirculated to drive people to the wrong conclusions, such as when viral tweets claim that dumped or stolen mail contained ballots or targeted voters specifically. In one instance during the 2020 elections, the pictures that circulated—and garnered thousands of retweets—were actually from a news story two years prior.[17] **See Figure 4 where multiple users posted copycat disinformation posts with this image (successfully reaching a massive audience).**

Common examples of disinformation when it comes to voting and elections include the following:

- Wrong election date, often specific to one group (e.g., "Democrats vote on Wednesday" when the election is on a Tuesday)
- Bogus election rules, often specific to one group (e.g., during the 2016 election numerous social media posts falsely claimed that you could "text your vote" for Hillary Clinton)
- Voter intimidation (e.g., claims that by voting you may put yourself in danger because of the presence of police, Immigration and Customs Enforcement, military, or vigilantes)
- Untrue claims about election integrity/security (e.g., false claims that vote by mail is not secure, false claims that the election process was being rigged or altered)
- Untrue claims post-election about results (e.g., "The Big Lie" false claim that the 2020 election was "stolen" from Trump)

Figure 4: Two disinformation posts that used a photo from an unrelated news story two years prior to make false claims. (These posts were found by a Common Cause volunteer, and our analyst reported them to Twitter, which removed them.)

Malinformation

Malinformation is content that is accurate but is intentionally manipulated to cause harm. This includes misrepresenting the context of a true news story, doxing (releasing personal information like addresses and phone numbers of an individual online to intimidate them), or selectively leaking correspondence. There are multiple ways malinformation negatively impacts voting and elections.

One common use of malinformation is "doxing." "Doxing" is the practice of publishing an individual's personal information online in an effort to intimidate or harass them. After the 2018 midterm elections, when some Florida counties were delayed in reporting the totals of a recount in a very close election, multiple users posted the personal contact information (including home address) of two Florida elections officials while votes were still being cast.[18] The officials were both women of color, and the posts appeared on Facebook pages, including "Confederate Resistance" (which has the Confederate Flag and an image of a gun-holding soldier in its banner).

Malinformation—particularly doxing of elections officials—poses a significant challenge to holding free and fair elections. Election officials are receiving threats[19] and abuse[20] for helping to administer our democracy, fueled by the conspiracy theories that social media platforms allow to thrive through their inaction. Voters, elections officials, poll workers, and volunteer poll monitors have all found themselves as targets of doxing, making our elections more dangerous to participate in, particularly for women and people of color. In one example, Trump and his allies spread conspiracies about election workers in Fulton County, Georgia, claiming that election worker Shaye Moss and her mother Ruby Freeman were involved in a plot to add fraudulent ballots to the count. In Trump's phone call to the Georgia secretary of state urging him to alter the results, he brought up the women, who were the targets of months of threats and harassment.[21] This was all part of what Trump supporters referred to as "Suitcasegate"—their false belief that Fulton County election workers smuggled in fraudulent ballots in suitcases.[22] Other election workers have had to go into hiding, reporting death threats and stalking.[23] A recent survey showed that 1 in 5 election workers have reported receiving threats, and 1 in 3 have felt unsafe at work, all as a consequence of election disinformation.[24]

When elections workers and volunteers are attacked by partisans, it is more likely that only partisans themselves will take the role of administering our elections, which threatens the integrity of elections.

When Is Disinformation Spread?

Election disinformation is spread before, during, and after Election Day.

In the spring of 2020, before voting began, former president Donald Trump and his campaign promoted disinformation falsely, claiming that mail-in ballots were less secure and part of a plan to rig the election against him[25] and Republicans more generally.[26] Junk websites like the Gateway Pundit and Breitbart, as well as Fox News, promote stories of election dysfunction and isolated incidents of voter malfeasance that drive a false narrative of voter fraud, even outside of election periods. Experts analyzed social media between January 2020 and September 2020 and found a massive 3.1 million mentions of disinformation about voting by mail.[27]

With the recent growth in the use of vote-by-mail options and early voting, the active voting "election period" is longer now than in years past. Disinformation spreaders often attack during this longer voting period. During the September 2021 California gubernatorial recall election, which had universal vote by mail (where all registered voters are automatically mailed ballots), during the time that ballots were in mailboxes, Fox News host Tomi Lahren falsely claimed that "the only" thing that will defeat the recall is "voter fraud."[28] Republican candidate Larry Elder made claims of likely voter fraud in the run-up to the election, even creating a website that indicated that the election was rigged while voting was still underway.[29]

In the period after the election concludes and is called by the mainstream media, an increasing number of losers of contests have begun to use claims of a rigged election or unfounded claims of voter fraud to avoid accepting defeat. In the 2019 Kentucky gubernatorial election, after the race was called for his opponent, defeated Gov. Matt Bevin made repeated claims (without evidence) of a rigged election.[30] In the 2020 elections, after the race was called for Joe Biden, Donald Trump's claims of voter fraud and a rigged election were amplified throughout social media and mainstream media—Fox News, on its own, made nearly 800 statements that cast doubt on the results of the election in just two weeks after its own news desk called the election for Biden.[31] These lies gave energy to the so-called Stop the Steal movement that galvanized support for the deadly January 6 insurrection.[32]

Fox News, on its own, made nearly 800 statements that cast doubt on the results of the election in just two weeks after its own news desk called the election for Biden.

Unfortunately, significant damage can occur through the amplification of election disinformation year-round. disinformation efforts not only serve to undermine the legitimacy of the last election but also to lay a foundation of doubt regarding the next election and the integrity of our government, generally. With perpetual campaigns, some politicians continue disseminating election disinformation to keep their donors giving and their names in headlines. Donald Trump, for example, raised more than $100 million peddling lies in the first six months of 2021[33] and continues to tease another run for president in 2024.[34] Some individuals looking to build a following on social media use disinformation to harness the natural outrage we feel about unfairness—especially when it comes to our democracy and our voice in who is elected to lead us.

Disturbingly, from the perspective of social media companies, disinformation is good for business year-round because it drives engagement and use of the platform (which can be monetized by ads and data gathering). Facebook whistleblower Frances Haugen told *60 Minutes* that "Facebook's own research shows that it amplifies hate, misinformation, and political unrest," and that the company prioritizes profit over the public good.[35] Unfortunately, in our current political, media and regulatory frameworks, there is very little to be lost and few systems of accountability that can prevent or hold accountable bad actors or the platforms they use to spread their messages. As a result, we now have a constant, 24/7, year-round public conversation on social media and in the mainstream media anchored with the false narratives of widespread voter fraud and a rigged election. Election disinformation is always in season.

How Is Disinformation Spread?

Disinformation is spread through a variety of communications channels and changes as technology advances. Prior to the widespread adoption of the world wide web and social media, most election disinformation was spread through flyers, billboards, and phone calls.[36] The following are the most-used communications channels for the spread of disinformation today.

Websites and Media Outlets

Junk websites are frequent purveyors of disinformation. *PolitiFact*'s "Junk News Almanac" in November 2017 listed over 300 websites that frequently share mis- and disinformation.[37] But even legitimate websites and media outlets can send mis- or disinformation to voters. Even the most trusted sources of election information sometimes contain misinformation. For example, in the primary elections in New Hampshire in September 2020, at least two county websites stated incorrect information about the acceptance of absentee ballots by election workers at polling places.[38]

Search Engines

A disinformation narrative can take root or spread when users see search engine results for search queries. Google's search engine will return content from junk websites, although Google's recommendation algorithm will often surface more trustworthy sources first. However, even before users click on content that Google's results surface during their search, there are additional opportunities for election disinformation. In 2020, Common Cause researched how search engines responded to a set of voting-related queries and found multiple occasions where search results contained incorrect voting information (e.g., "online voting") or phrases that indicated a disinformation narrative surfaced by scammers or those attempting to suppress the vote.

For example, at the bottom of every page of results, Google shows a "related searches" panel **(see Figure 5).** This panel shows keywords or phrases that other users who searched for the same term also requested. For example, if many users search for "voting" followed by "register to vote," "register to vote" might show up as a related search for "voting." Google's "related searches" feature can steer users toward misinformation.

This "related searches" feature can have a significant impact on the user experience.[39] One market research firm found that 18% of searches involve the user changing the search query before they click on any results and speculated that Google promotes related searches to target such users.[40]

As one troubling example, we found that Google searches for "vote," "how to vote," and "voting" all directed users to related searches about online voting, such as "vote online," "how to vote online," and "online voting website." Encouraging people to vote online—an option that generally

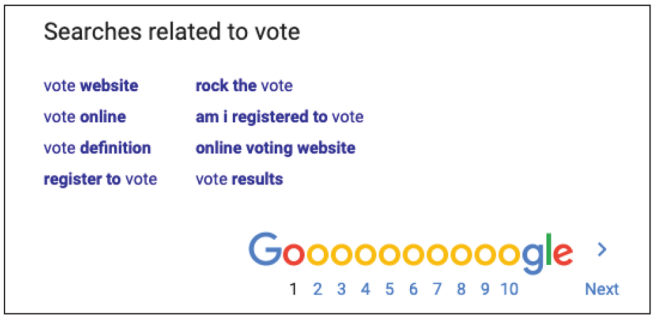


Figure 5: Searches related to "voting" shown in Arizona.

doesn't exist in the United States for the voting public[41]—is a known scam and form of disinformation designed to keep people away from the real polls.[42] Disinformation asserting that voters can cast their ballots via text messages or the web plagues modern elections.[43]

Another area of concern is Google's "autocomplete" feature, which matches the characters that a user has typed into the search bar with previous searches that start with those characters.[44] Autocomplete is different from related searches in that autocomplete attempts to predict how a user will finish a query. According to Google, the search engine turns off the autocomplete feature when the query cannot be reliably matched with related content, when the predictions contain sexual or other policy-violating content, or when a user has previously reported a prediction as inappropriate.[45] Google has a special rule against autocomplete predictions that could affect election integrity.

> We don't allow predictions that could be interpreted as a position for or against any candidate or political party, nor which could be interpreted as claims about the participation in or integrity of the electoral process.[46]

On September 10, 2020, Google clarified that this ban extends to search predictions that suggest donating to a particular candidate or that discuss election processes or requirements, whether they are accurate or not.

However, we found numerous autocomplete suggestions that appear to violate Google's rules **(see Figure 6)**. For example, the third suggestion when we typed "ballot" into a clean instance of Google was "ballot harvesting," a loaded term that has been used by Trump and others to raise suspicions about the practice of ballot collection (i.e., when a person other than the voter collects a completed absentee ballot to drop off).[47]

Existing research finds that a majority of internet users trust Google to provide them with accurate information and that the way Google presents information has the potential to sway public opinion.[48] A 2015 study found that changing the order of search results had a statistically significant impact on undecided voters' candidate preferences.[49] The authors of the study noted that the "impact of such manipulations would be especially large in countries dominated by a single search engine company," like the United States, where Google's market share approaches 90%.[50]

Common Cause engaged in dialogue with Google about these examples, and the company pledged to take action where it identified that our research showed examples that were against its terms of service.

One additional way search results can harm voters is through scam advertisements. Research by the Tech Transparency Project found "search terms like 'register to vote,' 'vote by mail,' and 'where is my polling place' generated ads linking to websites that charge bogus fees for voter registration, harvest user data, or plant unwanted software on people's browsers."[51] After Google pledged to correct this issue, Common Cause collaborated with the Tech Transparency Project and found examples that the problem persisted.[52] This highlights the need for watchdogs and ongoing monitoring of different vectors where disinformation can spread.

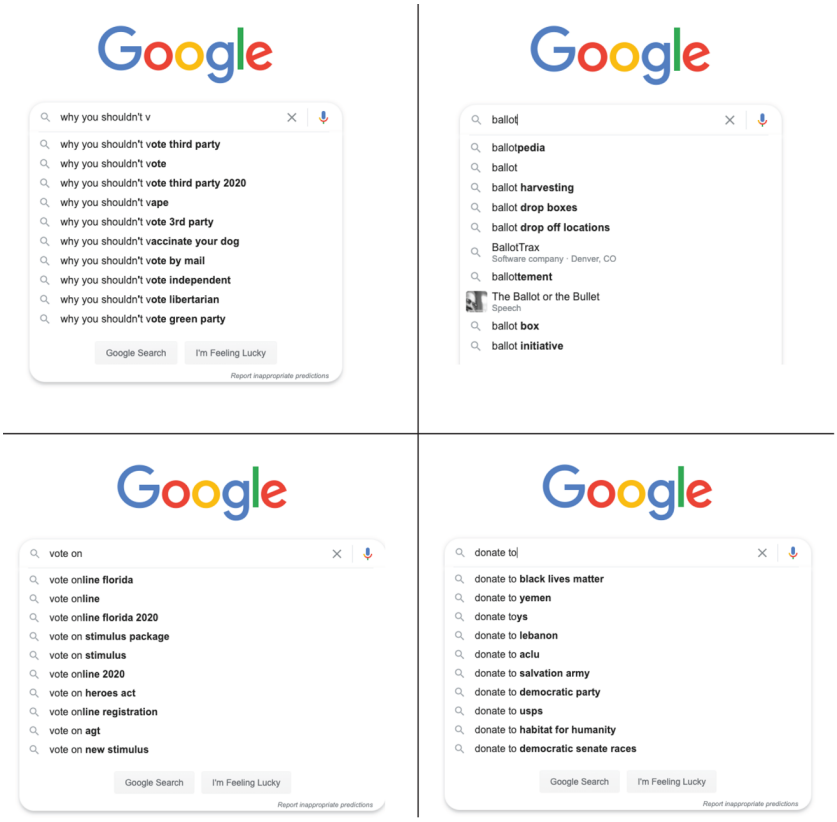


Figure 6: Google autocomplete suggestions for various election-related terms.

Email

In October 2020, the *Washington Post reported that registered Democratic Party voters in four states received threatening emails from unknown actors using the Proud Boys domain that took advantage of information from voter files to harass voters.*[53] The emails reportedly targeted Democratic voters in swing states and told them to change their votes to Trump, or "we will come after you." The emails were reported in Florida, Arizona, Pennsylvania, and Alaska. The Proud Boys denied involvement, pointing to the unsecured status of the domain as evidence that other provocateurs may have hijacked it. The FBI later reported that these emails were the work of Iranian intelligence.[54]

Robocalls

Robocalls (i.e., automated telephone calls that deliver recorded messages) are still used to spread disinformation, in part because they can target individual voters—or segments of voters. On August 27, 2020, Michigan's secretary of state tweeted a notice[55] that Detroit voters were receiving robocalls purporting to be from Jack Burkman and Jacob Wohl of "The 1599 Project" and posted a link to a YouTube audio recording of the call.[56] The robocall tells recipients that voting by mail will enter their information into a public database that "will be used by police departments to track down old warrants," be used to "collect outstanding debts," and enlist people into a mandatory vaccine program from the Centers for Disease Control and Prevention. The call then warns people not to give their information to "the Man."

Soon after these Michigan robocalls were made public,[57] reports surfaced of the same robocalls being made to voters in Philadelphia and Pittsburgh.[58] The two men cited in the call were known for their history of attempting to entrap public officials with bizarre schemes. The attorney general and secretary of state in Michigan announced an inquiry into the source of the calls.[59] On August 24, 2021, the Federal Communications Commission proposed a $5 million fine to the perpetrators.[60]

Text Messages

Similar to robocalls, text messages can be sent directly to individual voters via phone numbers and automated by computers. Text messages are also important disinformation vectors from individuals who, of their own volition, want to share disinformation with their contacts. Text messages are private communication, and most cellphone carriers, because of privacy concerns, cannot or will not actively monitor or interfere with users' text messages. This can make it more difficult to combat disinformation.

Some social media platforms, like WhatsApp, operate similarly to text messages—they are confidential and encrypted, one-to-one (or group) messages where only the receiver can view them. Similar to text messages, it is possible to send unsolicited messages on WhatsApp and similar platforms. However, most of the disinformation spread on WhatsApp in the 2020 election appears to have come from within a "group" of contacts, not from an outside interloper.[61]

Social Media

More than 70% of U.S. residents use social media,[62] and half of the adults in the United States "often" or "sometimes" get their news from social media.[63] With this increasing adoption of social media by voters, social media has become a critical vector for election disinformation. Social media is a broad category that creates multiple vectors for disinformation to spread, and the ways different social media platforms work create different challenges and opportunities to combat disinformation. Disinformation proliferates on social media. Global human rights group Avaaz found that just a small group of disinformation spreaders are responsible for a large portion of election and voting disinformation online, spawning millions

More than 70% of U.S. residents use social media, and half of the adults in the United States "often" or "sometimes" get their news from social media.

of interactions around false and misleading stories.[64] The Jacobs Technion-Cornell Institute at Cornell Tech found 7.6 million tweets and 25.6 million retweets from 2.6 million users that included key terms relating to voter fraud spanning from October 23 to December 16, 2020.[65] The spread of election disinformation via social media platforms is a huge and growing problem.

Some social media platforms, like Twitter, are "open." That is, most users can see most of the content. While some users on Twitter choose to keep their content private, and there are private "direct messages," most Twitter content is available to any user, can be searched and found, and has the potential to find its way into the "feed" of any user. YouTube and (generally) Instagram also fit into this category **(see Figure 7)**.

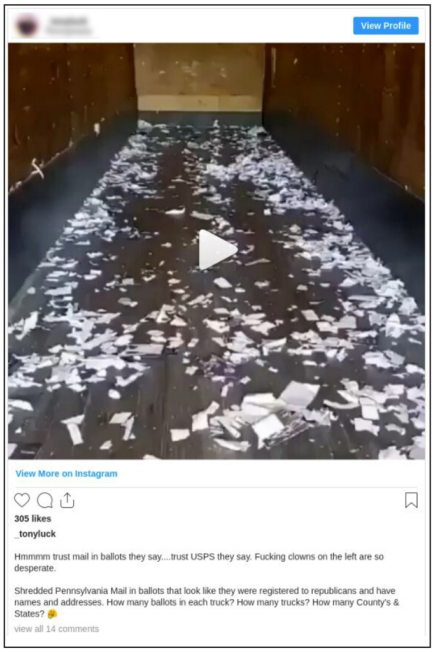


Figure 7: Instagram. The picture is from a video that went viral after people mistook residual shredding for mail-in ballots.

Some social media platforms, like WhatsApp, are "closed" and content is sent (and seen) by specific groups of users, not everyone on the platform. While WhatsApp allows group chats, you cannot search for content across the platform in the way you can with Twitter currently. NextDoor is another "closed" platform where most of the content posted there can only be seen by users in the specific neighborhood they reside in or in neighboring areas **(see Figure 8)**. Other posts on NextDoor can be public posts, but the majority are location specific.

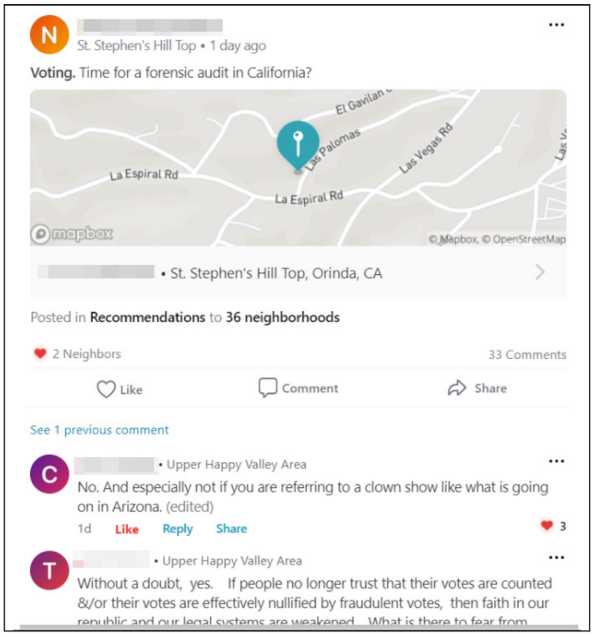


Figure 8: NextDoor. A NextDoor user calls for a "forensic audit."

Many social media platforms are a hybrid. Facebook, the platform with the largest user base in the United States, is a hybrid with both open and closed content. Open content on Facebook consists of posts by users on their own profile or page (as long as they are set to be publicly viewable). Closed content on Facebook includes not only direct messages but also groups. Groups have been growing in importance on Facebook. Between February 2017 and April 2019, active users of Facebook groups grew from 100 million people to 400 million people.[66] Groups have many different privacy settings,[67] but often their content is "private"—that is, only the users of those groups can see that content (though it appears through the main "news stream" of content when they log into Facebook). The upshot is that while some Facebook content is public (and searchable if, and only if, you have access to their CrowdTangle tool), that search will return only a portion of the content on the platform.

Telegram is another hybrid, combining encrypted one-on-one instant messaging, public posts (one-way broadcasts), and both public and private groups.[68] Telegram is most similar to WhatsApp in appearance and messaging functions but has channels where users can broadcast one-way posts, as well as capabilities for massive group chats in these channels. Posts on public groups can be forwarded to other channels and users. Because of these features and the encryption it

provides, Telegram has been a vital tool to organize against authoritarian rulers.[69] Recently, Telegram has been used by many right-wing activists and white supremacists in the United States **(see Figure 9)**.

A core tenet of social media is that users can create content that is seen, immediately, by other users. This means that any social media platform or user-generated content platform is a potential vector for election disinformation. Throughout our Stopping Cyber Suppression program, our monitoring efforts found mis- and disinformation on mainstream platforms like Facebook and Twitter, and on platforms with smaller user bases like NextDoor. We even found an example of election disinformation on the online tag function of the companion app to the Peloton exercise company **(see Figure 10)**. Peloton later banned the use of the "Stop the Steal" tags.[70]

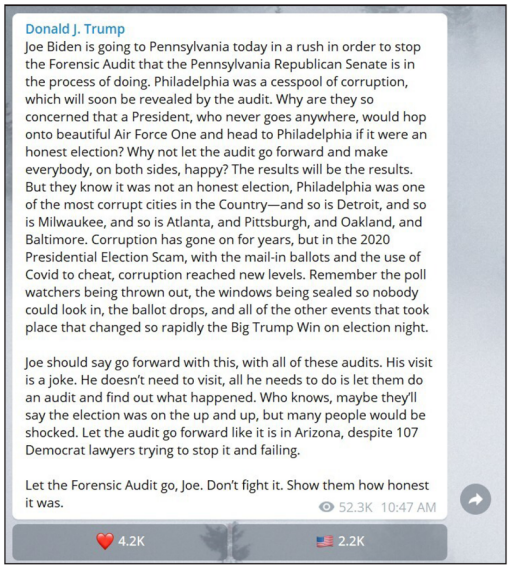


Figure 9: Telegram. This is an example of how Trump uses other social media networks to continue to spread his disinformation about 2020.

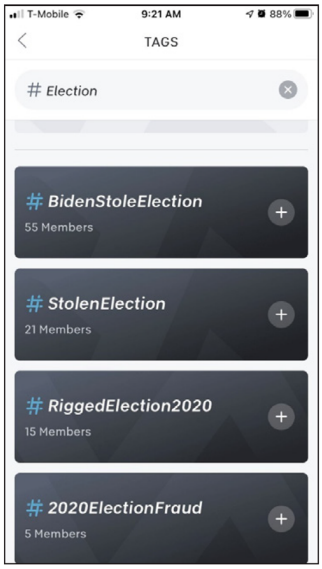


Figure 10: Peloton

Only the social media companies themselves have full access to the content and therefore are the only ones who would conclusively know how much election disinformation appeared on their platforms. A report from Stanford University found disinformation in the 2020 elections across multiple platforms.[71]

Facebook, as the social media platform with the largest user base in the United States (and worldwide), is the most important platform when it comes to preventing the spread of disinformation. According to Facebook's own reporting, between March 1 and November 2, 2020, Facebook applied some kind of label or warning to 180 million posts that shared election mis- or disinformation and removed 265,000 pieces of content for breaking the company's rules against voter interference **(see Figure 11).**[72]

Figure 11: Facebook. Here, a filing from a vote suppression group is used as proof of "voter fraud."

YouTube is owned by Google and is one of the most popular online platforms in the United States, used by seven-in-ten Americans including 26% of U.S. adults who get news there.[73] YouTube hosted videos that promoted election disinformation that had significant views: one study showed that YouTube was a key vector for disinformation used on other platforms, where Twitter users would tweet out disinformation-filled YouTube videos.[74] An independent analysis of YouTube videos revealed that during the week of November 3, 2020, videos supporting the false claim of widespread election fraud were viewed more than 138 million times.[75] YouTube's ability to grow a large audience quickly has helped spread election disinformation narratives. A group of pro-Trump channels connected to the far-right newspaper *Epoch Times* that launched on November 10, 2020, grew to 200,000 subscribers and 11 million views in less than two months with videos that contained election disinformation.[76]

From September to December 9, 2020, YouTube claimed to have removed "8000 channels and thousands of harmful and misleading elections-related videos for violating our existing policies."[77] While YouTube also pledged to disallow any "content alleging widespread fraud or errors changed the outcome of a historical U.S. Presidential election,"[78] research found that many election disinformation videos remained on the platform **(see Figure 12).**[79]

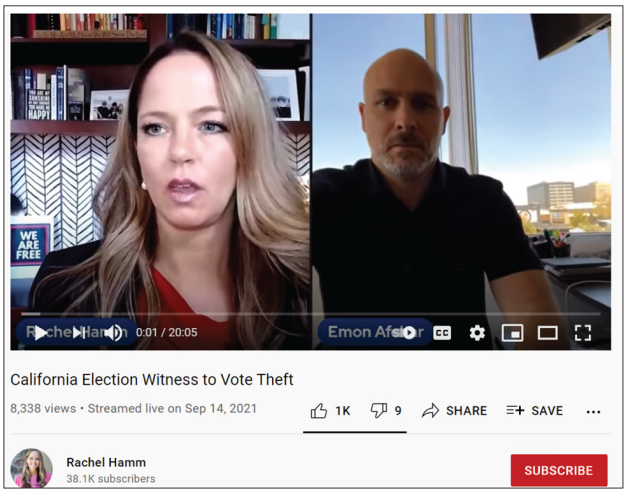


Figure 12: YouTube. In this video, a woman running for secretary of state claims votes were stolen in the California recall based on the experiences of the man interviewed.

Twitter has fewer users than Facebook and YouTube but maintains an important place in the rapid sharing and spreading of election disinformation. Most tweets are public and can be easily searched. And Twitter's application programming interface, which opens Twitter's data and functionality to external third parties, is accessible to social media monitoring tools and researchers, making it much easier for independent researchers (and Common Cause's Social Media Monitoring volunteers) to find and report disinformation. Twitter released a report that claimed it labeled 300,000 tweets containing "disputed and potentially misleading" information about the election between October 27 and November 11, 2020.[80] Twitter has not released any additional reports **(see Figure 13)**.

TikTok, a popular video app, also took measures against election disinformation. Despite reporting in February 2021 the removal of over 300,000 videos for election disinformation, it continues to surface videos to users that contain false claims **(see Figure 14).**[81] TikTok boasts up to one billion monthly users, and videos posted by users and surfaced to audiences via the algorithmic For You Page can receive thousands of views and shares before they are removed, even with a robust enforcement policy that acts to take them down within hours.[82]

Figure 13: Twitter. Here, a popular conservative influencer claims that votes were deleted live on CNN.

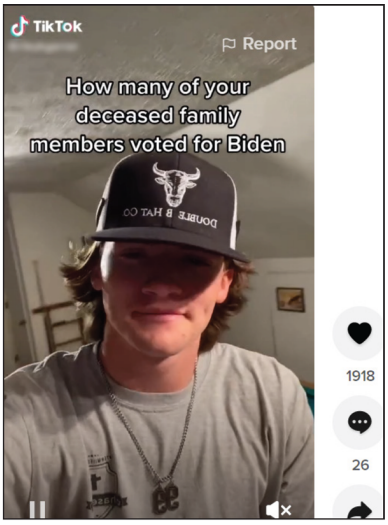


Figure 14: TikTok. A TikTok user posts a commonly used trope about "dead voters."

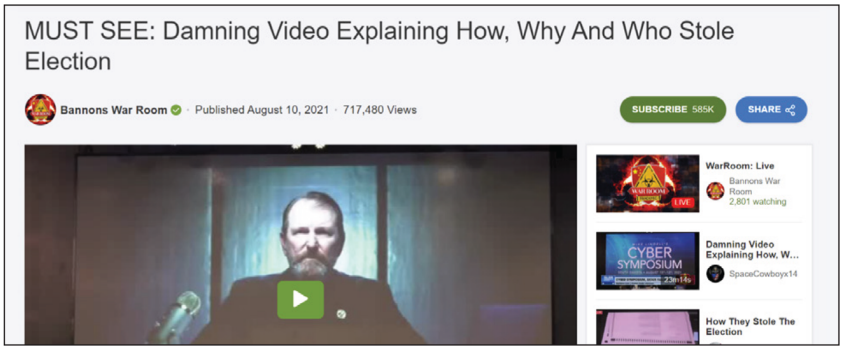


Figure 15: Rumble. Steve Bannon's show earned almost a million views with this video on "how, why, and who stole election."

Rumble is a video platform where users upload videos that can be monetized and licensed. Rumble is now the home of many viral right-wing disinformation videos, such as Steve Bannon's Pandemic War Room show, which has almost 600,000 subscribers and frequently posts clips making false claims about the 2020 presidential election **(see Figure 15)**. Analyses of Rumble show that it surfaces QAnon and conspiracy content to users at rates higher than accurate and

factual information, exposing its millions of monthly users to disinformation on a variety of subjects. According to one analysis, searching the word "election" on the platform led to two times the amount of disinformation as correct information.[83]

Who Is Spreading Election Disinformation and Why?

Few who intentionally spread election disinformation would publicize this fact because the behavior is sometimes illegal and always despicable. The ability of individuals to anonymously spread election disinformation is part of the problem—and strengthening transparency laws as recommended later in this report is part of the solution. Nevertheless, here is what we know about those spreading election disinformation in recent years.

Both foreign and domestic actors have used—and likely will continue to use—election disinformation. During the 2016 elections, the Russian Internet Research Agency created numerous posts on multiple social media platforms. According to the U.S. Senate Select Committee on Intelligence, this foreign interference was "at the direction of the Kremlin" and created social media content in support of then-candidate Trump and against Hillary Clinton.[84] In particular, the content was "principally aimed at African-Americans in key metropolitan areas."[85] Russian disinformation efforts included the use of the Facebook page Blacktivist, which purported to be a Black empowerment page and garnered 11.2 million engagements with Facebook users.[86] Both advertisement and organic (non-ad) content was published through this program. This Russian social media content was designed to drive divisions between voters and cause general political instability in the United States, a tactic that differed from more direct efforts to disenfranchise voters used by some other purveyors of election disinformation.[87]

Russian disinformation efforts included the use of the Facebook page Blacktivist, which purported to be a Black empowerment page and garnered 11.2 million engagements with Facebook users.

Whereas Russia's 2016 election interference exploited fissures between U.S. social groups, foreign interference by Russia and others in our 2020 election primarily entailed amplifying existing election disinformation narratives created by other bad actors, including then-president Donald Trump. As noted previously, intimidating emails sent to voters in 2020 purported to be from the white supremacist Proud Boys organization, but the Department of Homeland Security investigated and accused Iran of producing them.[88] Russian media in 2020 capitalized on the false narratives Trump and others spread about a "rigged" election and vote by mail in particular.[89] A report from the director of National Intelligence found that in the 2020 elections, Iran and Russia "spread false or inflated claims about alleged compromises of voting systems to undermine public confidence in election processes and results."[90] These were the same claims that domestic malign actors—including then-president Donald Trump and his party—were actively spreading.

Notwithstanding some foreign involvement in the spread of election disinformation in the United States, **the vast majority of election disinformation that plagues our politics appears to originate with and is amplified by domestic sources.**

A number of social scientists are working to understand the psychology behind individuals spreading disinformation. In our observations, gleaned from over 15,000 volunteer hours spent monitoring social media for mis- and disinformation during the 2020 election cycle, we have found that election **misinformation** is often spread by those sincerely attempting to be helpful in a climate of uncertainty and distrust (particularly when it came to the USPS and its ability to manage vote by mail in the 2020 elections) and **disinformation** is spread by individuals with partisan goals, including intraparty contests, like the Democratic Presidential Primary **(see Figure 16).**[91]

Figure 16: Disinformation image circulated on Twitter with incorrect election date for Super Tuesday primaries, branded as coming from the Biden campaign.

In an age of hyperpartisanship, spreading election disinformation can both serve to attack your political opponents and show that you are aligned with other members of your political tribe. Election disinformation—in particular, the narrative of a rigged election and pervasive voter fraud committed by Democrats—existed long before the rise of Donald Trump but now has become party orthodoxy. You can signal that you are a Trump-supporting "MAGA Republican" (an acronym for Trump's campaign slogan "Make America Great Again") by spreading stories that reinforce a narrative (however false) about a political system rigged against other MAGA Republicans. This creates a negative feedback loop of distrust in government and elections: a September 2021 poll showed that 78% of Republicans believe that Joe Biden did not win the presidency.[92] Numerous states and counties are proceeding with sham ballot reviews—even in areas where Trump won decisively.[93] Among 15 Republican candidates currently running for secretary of state in five battleground states, 10 have "either declared that the 2020 election was stolen or called for their state's results to be invalidated or further investigated."[94] **Election disinformation is spread by activists and candidates in the same way that political messaging and issue priorities used to be.**

While disinformation is spread by a large number of social media platform users, highly influential accounts and pages matter most, as the social media algorithms are more likely to promote content created by a user with a large following. These algorithms have empowered a small number of disinformation "superspreaders" to instigate the bulk of disinformation about COVID-19.[95] There are a few accounts with strong influence on social media that made the biggest contributions to spreading disinformation, and they are almost exclusively conservative. For example, Douglass Mackey, who the *New York Times* describes as a "far-right Twitter troll" and "right-wing

provocateur"[96] with nearly 60,000 Twitter followers, is currently being prosecuted by the DOJ for spreading election disinformation in the weeks leading up to the 2016 presidential general election and seems to have been driven by partisan and anti-Black racist motives.[97] Mackey's stated goal for his Twitter disinformation campaign was to "drive up turnout with non-college whites, and limit black turnout," with memes intended to suppress the votes of Hillary Clinton supporters.[98]

According to the Stanford Election Integrity Partnership's report on mis- and disinformation, "Influential accounts on the political right rarely engaged in factchecking behavior, and were responsible for the most widely spread incidents of false or misleading information in our dataset."[99] That included 15 "verified" Twitter accounts including Eric Trump, Donald Trump, Donald Trump Jr., and social media influencers like James O'Keefe, Tim Pool, Elijah Riot, and Sidney Powell.[100] Similarly, an analysis by the advocacy group Avaaz concluded that Facebook missed an opportunity to dramatically limit election disinformation by acting early on a select few accounts and content. The report noted that "the top 100 false or misleading stories related to the 2020 elections" were viewed 162 million times in three months. Moreover, Avaaz researchers found that 100 of the top Facebook pages that have spread disinformation were viewed more than 10 billion times between March and October.[101]

Wealthy conservatives with partisan motives spend big money, both directly and through "dark money" groups, to spread election disinformation. Jane Mayer, a preeminent investigative journalist for the *New Yorker* covering money in politics and author of the 2017 bestseller book Dark Money, recently turned her attention to the funding of election disinformation.[102] Mayer cites the conservative Lynde and Harry Bradley Foundation, with its $850 million endowment, as a major funder of recent election disinformation efforts through numerous nonprofits, including the Heritage Foundation, American Legislative Exchange Council, Honest Elections Project, Election Integrity Project California, and FreedomWorks.[103] Mayer also cites multimillionaire founder of Overstock.com, Patrick Byrne, as a purveyor of election disinformation in the form of his film The Deep Rig, which "asserts that the 2020 Presidential election was stolen by supporters of Joe Biden, including by Antifa members who chatted about their sinister plot on a conference call."[104]

Wealthy conservatives with partisan motives spend big money, both directly and through "dark money" groups, to spread election disinformation.

Some Republican politicians are also superspreaders of election disinformation,[105] seemingly motivated by at least two factors, raising money and rationalizing new voter suppression laws—both of which will help them win future elections. As noted earlier, Donald Trump raised more than $100 million peddling the "Big Lie" in the first six months of 2021,[106] and other Republicans have jumped on Trump's election disinformation gravy train. The *New York Times* analyzed campaign finance data from the first quarter of 2021 and observed that "leaders of the effort to overturn Mr. Biden's electoral victory have capitalized on the outrage of their supporters to collect huge sums of campaign cash," singling out Senators Josh Hawley and Ted Cruz, and Representatives Marjorie Taylor Greene and Kevin McCarthy.[107] The authors concluded, "Far from being punished for encouraging the [January 6] protest that turned lethal, they have thrived in a system that

often rewards the loudest and most extreme voices, using the fury around the riot to build their political brands."[108]

Relatedly, since the beginning of the year, calls for "forensic audits" of the 2020 election have gained steam as a means for Trump supporters to allegedly collect evidence of election fraud and for those in right-wing spaces to profit off of these endeavors. The most infamous of these is the recently concluded sham ballot review in Maricopa County, Arizona, which cost up to $7 million and ended up affirming a Biden victory—as expected.[109] The election results in Maricopa County had been accurately counted, certified, and audited by the county, using processes that exist all around the United States to ensure the accuracy and integrity of our elections, before Arizona Senate Republican leaders launched their own Trump-inspired partisan review. In the following section, we profile the Arizona sham ballot review as a case study in election disinformation. And the "audit" push shows no signs of stopping. Ten other states are in various states of either conducting or instigating sham ballot reviews.[110] The end result of these sham ballot reviews isn't renewed confidence in elections but a calcified and further-reinforced belief on the part of Trump supporters that there is a "there" there, and to keep their attempts to undermine the election process going.

Disinformation Case Study: Arizona Sham Ballot Review

Late in September, national media outlets delivered expected news. A Republican-commissioned review of nearly 2.1 million ballots cast in Arizona's November 2020 election, carried out over many months by a wholly unqualified firm known as Cyber Ninjas, was finally over and reaffirmed what we already knew: Joe Biden won Maricopa County.[111] Arizona's sham ballot review illustrates the many facets and problems of election disinformation—a perfect case study.

Veteran voting rights advocate and lawyer Ralph Neas oversaw a study of the Arizona process for the nonpartisan Century Foundation and explained that though the process was a "farce," it may nonetheless have "extraordinary consequences."[112] Neas explained: "The Maricopa County audit exposes exactly what the Big Lie is all about. If they come up with an analysis that discredits the 2020 election results in Arizona, it will be replicated in other states, furthering more chaos. That will enable new legislation. Millions of Americans could be disenfranchised, helping Donald Trump to be elected again in 2024. That's the bottom line. Maricopa County is the prism through which to view everything. It's not so much about 2020—it's about 2022 and 2024."[113]

There has been, and will continue to be, a strategic effort to spread disinformation by bad actors about the 2020 election using the Arizona Maricopa County sham ballot review process. This disinformation is being promoted and amplified through major social media platforms, including Twitter and Facebook. Bad actors point to Facebook and Twitter disinformation content as evidence of the validity of their conspiracy theories and share this content on Telegram and other spaces where they are organizing their efforts, creating an echo chamber of disinformation.

We've found examples of disinformation about the Arizona sham ballot review on Twitter and Facebook, but also other platforms like Telegram. QAnon influencers and audit supporters repost the @ArizonaAudit Twitter account content and other misinformation to various other platforms, ranging from Instagram to Telegram.

A large portion of online organizing and chatter surrounding the review is happening on Telegram, where users can discuss the stolen election myth without fear of violating platform policies (Telegram has no prohibition against election disinformation).[114] A critical component of these discussions is the sharing of screenshots or original posts from Twitter. In the first example, the @ArizonaAudit tweet is shared on a popular QAnon influencer's Instagram account **(see Figure 17)**. In the second, an accusation about Arizona secretary of state Katie Hobbs is disseminated from Twitter into the rumor mill of Telegram **(see Figure 18)**.

Figure 17: Instagram post of @ArizonaAudit tweet.

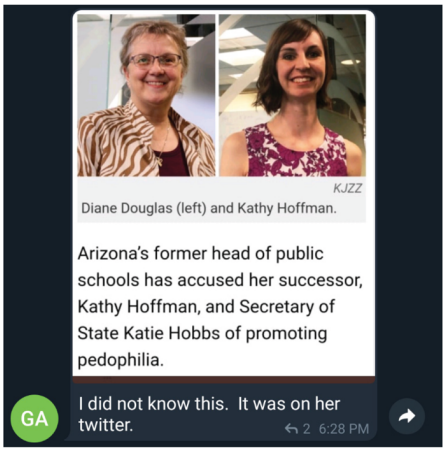


Figure 18: Telegram post republishing disinformation from Twitter.

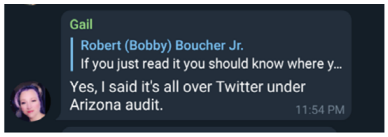


Figure 19: Telegram post referring to disinformation on Twitter.

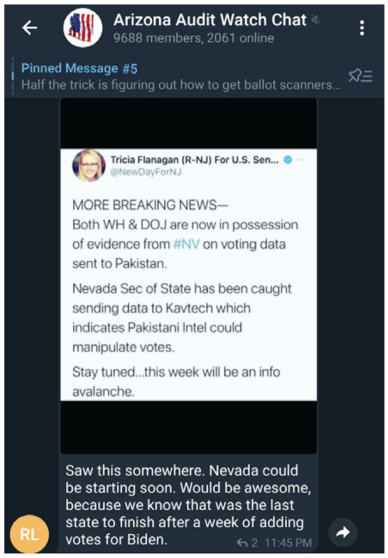


Figure 20: Telegram post sharing disinformation tweeted by congressional candidate.

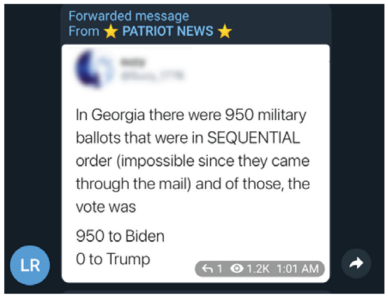


Figure 21: Telegram post forwarding debunked election disinformation.tweeted by congressional candidate.

Users on Twitter and Facebook are able to go viral making debunked claims[115] about the Arizona audit, such as the viral claim that there were 250,000 "illegal votes" found or that databases were deleted.[116]

This disinformation doesn't stay on Twitter or Facebook—it migrates to a new home on Telegram, where users add commentary and further radicalize. As seen in **Figure 19**, two users are debating in a Telegram chat for the Arizona audit (that boasts 14.3k members) where to find sources. One tells the other, "it's all over Twitter under Arizona audit." In **Figure 20**, a screenshot of viral disinformation from a verified congressional candidate account is shared in the same chat. In the third example, **Figure 21**, a forward of yet another debunked claim[117] is shared. Finally, CodeMonkeyZ (Ron Watkins, major QAnon influencer), shared Rep. Paul Gosar's (R-AZ) claims about fraudulent Arizona votes to his 245,000 followers on Telegram **(see Figure 22)**.

In some cases, they organize online actions, such as Twitter hashtags. For example, the hashtag #FraudVitiatesEverything is based on a saying coined by CodeMonkeyZ (Ron Watkins, major QAnon influencer), who says it as a claim that the election will be overturned due to fraud **(see Figure 23)**.

Figure 22: Telegram post republishing election disinformation tweeted by Rep. Gosar.

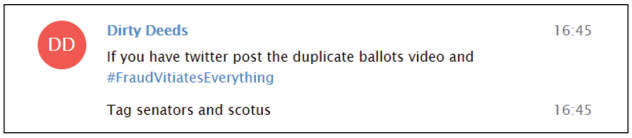


Figure 23: Disinformation hashtag campaign.

Disinformation spreads and jumps between social media platforms regularly. And those spreading disinformation point to mainstream platforms like Facebook and Twitter to show the efficacy of their efforts. A viral tweet or Facebook post is a trophy for a disinformation spreader. Because these platforms have more mainstream users and because they have an active content moderation regime that purports to remove some kinds of disinformation, it is far more dangerous to have disinformation spread on them as they appear to have the imprimatur of truth. In particular, Facebook algorithmically amplifies content to users, even if they didn't specifically ask to receive it, because of their recommendation engine. Facebook's own internal research, according to documents provided by whistleblower Frances Haugen, shows that the divisive, polarizing, angry content—like election disinformation—spreads better and faster than other content. Content is shown to users through algorithmic amplification—invitations to groups, suggested pages to follow and content promoted to users in their feed. Based on the content of the pages users follow, they will have content "pushed" to their newsfeed. According to Haugen, Facebook knows that this can lead users into an experience filled with extremist content—not from their own choosing but from simply following the recommendations of the platform.[118]

Disinformation spreads and jumps between social media platforms regularly.

If election disinformation were limited to a self-selected group of conspiracy theorists, it would continue to be a problem—but a much smaller and more manageable one than the mainstreaming of disinformation currently happening on Facebook and Twitter.

SECTION 2: STATE AND FEDERAL LAWS REGULATING ELECTION DISINFORMATION

Several different bodies of law provide tools for fighting election disinformation. A primary purpose of election disinformation is to suppress and sometimes intimidate voters. Consequently, election laws prohibiting voter intimidation and false election speech play an important role in fighting election disinformation. Several other bodies of law are also critically important to the fight. Strong campaign finance disclosure laws can shine the light of publicity on those seeking to undermine our elections from the shadows and help ensure existing laws are enforced. Communications laws, consumer protection laws, media literacy laws, and privacy laws can all play a part in effectively regulating and deterring election disinformation.

Federal and state laws of all these types are detailed in the following sections, highlighting some of the presently available legal tools for stopping election disinformation. To be certain, current laws across the United States are not entirely up to the task of preventing the increasingly sophisticated election disinformation tactics that will be deployed in 2022 elections and beyond. And there's no single "silver bullet" reform that would fix everything. But there are some important, effective laws on the books today that should be expanded to other jurisdictions and vigorously enforced. Such "best practices" are included in the final section of this report, along with other recommended reforms.

There's no single "silver bullet" reform that would fix everything. But there are some important, effective laws on the books today that should be expanded to other jurisdictions and vigorously enforced.

Voter Intimidation and False Election Speech Laws

Federal law and laws in nearly every state contain provisions explicitly prohibiting voter intimidation, with many of these laws being rightly interpreted as prohibiting election disinformation.

Some states have enacted laws explicitly prohibiting various types of false election-related speech—e.g., false statements about voting procedures/qualifications, candidates, incumbency, endorsements, veteran status, or ballot measure effects. In this report, we focus only on the first of these types: laws prohibiting false statements about voting procedures and qualifications such as where and when to vote. Our reasons are twofold and related to one another.

First, the veracity of statements about voting procedures and qualifications (e.g., the date of the election, the hours polls are open) is easily ascertainable, and determining such veracity can be done in an entirely nonpartisan, objective fashion. By contrast, determining the veracity of statements about a candidate (e.g., a candidate's stance on an issue) is often more subjective, as reflected by the rating systems some prominent fact-checkers use. For example, the Poynter Institute's PolitiFact uses a "truth-o-meter" with six grades: true, mostly true, half-true, mostly false, false, pants on fire.[119]

Second, and relatedly, courts have for years been divided on the constitutionality of laws prohibiting false speech characterizing candidates and ballot measures, with at least two federal appellate courts in recent years striking down such laws as unconstitutionally vague and overbroad.[120] Courts are much more likely to uphold as constitutionally permissible narrower laws prohibiting false statements about the procedures and qualifications of voting. As Professor Richard L. Hasen argued in a 2013 law review article, "The strongest case for constitutionality is a narrow law targeted at false election speech aimed at disenfranchising voters."[121]

The following section summarizes voter intimidation and false speech laws at the federal level and in numerous states. And the recommendations section at the end of this report identifies the best features of these laws, urging their adoption throughout the United States.

Federal Voter Intimidation and False Election Speech Laws

Federal government efforts to protect the right to vote and prevent voter intimidation date back to the period immediately following the Civil War and the creation of the DOJ in 1870[122] and the passage of the Ku Klux Klan Acts in 1870–71.[123] Several federal voting rights laws relate directly to election disinformation and voter intimidation. And though the DOJ has found some forms of voter intimidation to be "difficult to prosecute" because the intimidation is "both subtle and without witnesses,"[124] such is not the case for voter intimidation via election disinformation, which is often blatant and in full public view.

The National Voter Registration Act of 1993 makes it a crime to knowingly and willfully intimidate or threaten any person for voting, registering to vote, or aiding others to register and vote.[125] Another federal criminal statute similarly provides that "[w]hoever intimidates, threatens, coerces, or attempts to intimidate, threaten, or coerce, any other person for the purpose of interfering with the right of such other person to vote" in a federal election has committed a crime subject to fines or imprisonment.[126] The DOJ explains that this statute "criminalizes conduct intended to force prospective voters to vote against their preferences, or refrain from voting, through activity reasonably calculated to instill some form of fear."[127]

Conspiracy to "injure, oppress, threaten, or intimidate any person…in the free exercise or enjoyment of any right or privilege secured to him by the Constitution or laws of the United States"—including the right to vote—is a felony under federal law.[128] This criminal code provision covers voter suppression schemes, **including "providing false information to the public—or a particular segment of the public**—regarding the qualifications to vote, the consequences of voting in connection with citizenship status, the dates or qualifications for absentee voting, the date of an election, the hours for voting, or the correct voting precinct."[129]

In January 2021, the DOJ charged Twitter user Douglass Mackey (a.k.a. "Ricky Vaughn") with violation of this statute for conspiring "to injure, oppress, threaten and intimidate persons in the free exercise and enjoyment of a right and privilege secured to them by the Constitution and laws of the United States, to wit, the right to vote[.]"[130] The DOJ alleges that in the weeks leading up to the November 2016 presidential election, Mackey conspired with others to spread memes on Twitter falsely claiming that Hillary Clinton supporters could vote via text message to a specific phone number included in the memes **(see Figure 24).**[131] At least 4,900 individuals attempted to vote by texting "Hillary" to the number included in the memes.[132] *The New York Times* reported

Figure 24: Voter suppression Twitter meme.

that this "appeared to be the first criminal case in the country involving voter suppression through the spread of disinformation on Twitter" and that the "case will test the novel use of federal civil rights laws as a tool to hold people accountable for misinformation campaigns intended to interfere with elections[.]"[133] The case remains pending as of this writing.

In addition to the federal criminal code provisions detailed in the preceding paragraphs, the **Voting Rights Act of 1965 and other civil rights laws** also prohibit disinformation activities that amount to voter intimidation or suppression. The Voting Rights Act provides that no person "shall intimidate, threaten, or coerce, or attempt to intimidate, threaten, or coerce any person for voting or attempting to vote."[134] For example, this statute was successfully relied on by the DOJ to win a consent decree in a 1990 lawsuit against the North Carolina Republican Party, which had mailed disinformation postcards to 125,000 Black voters throughout the state, incorrectly stating that recipients could not vote if they had moved within 30 days of the election and threatening criminal prosecution.[135]

And as voting rights lawyer Michael Weingartner explains in a forthcoming law review article, recently, some plaintiff victims of election disinformation have turned to a provision of the Ku Klux Klan Acts that provides for an award of monetary damages to victims of conspiracies to prevent giving their "support or advocacy" to federal political candidates.[136] This statute, Weingartner argues, holds promise to "redress modern voter intimidation, deter bad actors, and provide an incentive to plaintiffs to bring suit."

State Voter Intimidation and False Election Speech Laws

The federal laws detailed earlier prohibiting voter intimidation and suppression—including some disinformation tactics—generally apply to any election with candidates for federal office on the ballot. Nearly every state, likewise, has laws prohibiting voter intimidation and suppression, applicable to elections even when no federal office candidates are on the ballot. A few states have laws explicitly regulating false election-related speech, and a few others have interpreted more general anti-intimidation laws to prohibit false election speech.

APPENDIX I summarizes the voter intimidation and false speech laws of several states. Among the best state laws worthy of emulating around the nation, **Colorado** law provides that no person shall knowingly or recklessly "make, publish, broadcast, or circulate or cause to be made, published, broadcasted, or circulated…any false statement designed to affect the vote on any issue submitted to the electors at any election or relating to any candidate for election to public office."[137] The Colorado attorney general's guidance makes clear that disinformation tactics—including "misleading phone calls, texts, or emails to a voter"—can constitute illegal voter intimidation."[138]

Similarly, **Hawaii** law provides that any person who "knowingly broadcasts, televises, circulates, publishes, distributes, or otherwise communicates…false information about the time, date, place, or means of voting with the purpose of impeding, preventing, or otherwise interfering with the free exercise of the elective franchise" has committed illegal election fraud.[139]

And **Virginia** explicitly outlaws communicating to a "registered voter, by any means, false information, knowing the same to be false, intended to impede the voter in the exercise of his right to vote," including information "about the date, time, and place of the election, or the voter's precinct, polling place, or voter registration status, or the location of a voter satellite office or the office of the general registrar."[140] Importantly, Virginia law includes a private right of action for registered voters to whom such false information is communicated, enabling them to seek an "injunction, restraining order, or other order, against the person communicating such false information."[141]

Campaign Finance Laws

In 1933, Supreme Court Justice Lewis D. Brandeis famously wrote, "Publicity is justly commended as a remedy for social and industrial diseases. Sunlight is said to be the best of disinfectants; electric light the most efficient policeman."[142]

The Supreme Court cited this Brandeis quote in its 1976 seminal campaign finance law decision *Buckley v. Valeo* in which the Court upheld as constitutionally permissible federal campaign finance disclosure requirements.[143] While many of those spending money to spread election disinformation prefer to hide in the shadows, knowing that disclosure of their identity would dishonor them and make clear their partisan motivations, strong campaign finance laws can force them into the light of day. The Buckley Court explained that "disclosure provides the electorate with information…in order to aid the voters in evaluating those who seek federal office" and informing voters of the "sources of a candidate's financial support [to] alert the voter to the interests to which a candidate is most likely to be responsive[.]"[144] Disclosure laws also "deter actual corruption and avoid the appearance of corruption by exposing large contributions and expenditures to the light of publicity."[145] Such exposure, the Court reasoned, "may discourage those who would use money for improper purposes either before or after the election."[146]

For nearly a half-century, federal and state courts around the nation have stood by the Buckley Court's reasoning. Just last month, a federal appellate court upholding a challenged disclosure law wrote, "A well-informed electorate is as vital to the survival of a democracy as air is to the survival of human life."[147]

Federal Campaign Finance Disclosure Laws

Federal law imposes thorough disclosure requirements on candidates, political parties, and other political committees. They must disclose the name and other identifying information of any donor who contributes more than $200, as well as any recipient of a payment exceeding $200 from the candidate or committee.[148] They must also include a "paid for by" disclaimer on any public communication they pay to distribute.[149] Consequently, if a candidate, party, or other political committee is paying to distribute disinformation, the public can know about it.

However, federal law disclosure requirements are weak and ineffective with respect to individuals and nonpolitical committee organizations such as so-called 501(c)(4) social welfare organizations, labor unions, and trade associations like the U.S. Chamber of Commerce. Only a narrow range of political spending by such organizations triggers disclosure and disclaimer requirements. Unless an ad expressly advocates[150] the election or defeat of a candidate (e.g., vote for candidate Smith), mentions a candidate and is aired on TV or radio in close proximity to an election,[151] or solicits a contribution to a candidate or political committee,[152] such ads are not subject to federal campaign finance law disclosure and "paid for by" disclaimer requirements.

In other words, **federal law leaves plenty of opportunities for individuals and nonpolitical committee organizations to disseminate disinformation without triggering disclosure or disclaimer requirements**. As long as they stay off TV and radio, and avoid express phrases like "elect Jones," their disinformation campaigns go unregulated by campaign finance law. This is a campaign finance law problem. In the final section of this report, we recommend some solutions.

Another provision of federal campaign finance law related to election disinformation is a statute prohibiting a candidate or employee of a candidate from fraudulently misrepresenting that they are acting for or on behalf of any other candidate or political party in a manner that is damaging to such other candidate or party.[153] The same law likewise prohibits any person from fraudulently misrepresenting that they are "speaking, writing, or otherwise acting for or on behalf of any candidate or political party…for the purpose of soliciting contributions or donations."[154]

However, the Federal Election Commission (FEC) "has a long history of finding no misrepresentation where communications contain disclaimers accurately identifying the true sponsor," unless the body of the communication contains an explicit misrepresentation that "countermands an otherwise accurate disclaimer."[155] Even a technically deficient disclaimer may suffice, so long as the disclaimer accurately identifies the sponsor.[156, 157]

In short, so long as an implicitly misleading political communication contains the required fine print or quickly spoken "paid for by" language at the end, the FEC will likely conclude the communication **does not violate** the "fraudulent misrepresentation" law. This is another area of federal campaign finance law that needs to be strengthened to reduce the spread of election disinformation. We recommend some fixes in the final section of this report, including one reform with bipartisan support among FEC commissioners.

State Campaign Finance Disclosure Laws

Spending in state and local candidate and ballot measure elections is regulated entirely by state (and sometimes local) campaign finance laws. Like federal law, most states' campaign finance laws are quite effective concerning spending by candidates and political committees but deficient when it comes to spending by individuals and nonpolitical committee entities.

APPENDIX II summarizes several states' campaign finance disclosure laws relevant to election disinformation. **Alaska has enacted one of the nation's most effective laws for tracing the source of funds** spent on election advertising by groups, including those that do not qualify as political committees,[158] requiring such groups to disclose the identity of any contributor who has given the group more than $250 in the aggregate during the calendar year "for the purpose of influencing the outcome of an election," as well as all election-related contributions and expenditures made by such groups, including contributions to other such groups.[159] The purpose of this statute is to reveal contributors whose funds are transferred through multiple organizations before being spent on election advertising—contributors who would evade disclosure under most jurisdictions' laws.

California has likewise led the way in recent years, strengthening campaign finance disclosure laws applicable to common sources and types of election disinformation. In 2014, the state strengthened disclosure laws applicable to "multipurpose organizations" spending money to influence California elections (e.g., 501(c)(4) social welfare organizations, often referred to as "dark money" organizations because, in most jurisdictions, they are not required to publicly disclose their funders).[160] And in 2018, California took another step by enacting the "Social Media DISCLOSURE Act,"[161] strengthening disclosure requirements by requiring "paid for by" disclaimers on a broad array of political advertising disseminated via social media platforms. The state's campaign finance regulatory agency, the California Fair Political Practices Commission, has done a good job implementing these laws and continually monitoring evolving campaign finance practices in an effort to keep state campaign finance laws and policies up to date.

California has likewise led the way in recent years, strengthening campaign finance disclosure laws applicable to common sources and types of election disinformation.

Maryland,[162] **Minnesota,**[163] and **Rhode Island**[164] have also enacted legislation requiring certain tax-exempt organizations that are often the source of undisclosed "dark money" political spending in other jurisdictions to disclose their donors and political spending.

Finally, the state of **Washington has some of the strongest disclosure laws in the nation applicable to election disinformation and other digital political advertising.**[165] Washington Public Disclosure Commission regulations provide for modified "paid for by" disclaimers on certain digital ads[166] and require online platforms that sell paid political advertising to provide the public with access to detailed digital ad information.[167]

Federal Communications Laws

Section 230 of the federal Communications Decency Act has long provided digital platforms with legal protections to moderate content online without fear of liability.[168] Section 230 immunizes websites, including internet platforms such as Facebook and Twitter, from liability as a publisher of third-party content.[169] The statute's "Good Samaritan" provision has two primary components that are often described as a "shield and sword." First, Section 230 shields websites from lawsuits regarding content posted by third parties on their platform.[170] For example, Facebook would be protected from lawsuits under the statute for third-party user posts hosted on their platform. Second, the statute provides platforms with a sword to remove content they determine is obscene, violent, or otherwise objectionable without fear of liability.[171]

The broad protections the statute provides empower platforms to take down disinformation while also providing them cover should they choose to leave the misleading content in question up. However, because platforms have broad discretion to moderate content without fear of liability, they are more likely to leave offending content up instead of taking it down.[172] We saw this time and time again during the 2020 election season. A report on Facebook's content moderation failures from advocacy group Avaaz found that Facebook's failure to take down misinformation resulted in over 10.1 billion estimated views of content from top-performing pages that repeatedly shared misinformation over the eight months before the U.S. elections.[173] Common Cause's research found many examples of social media posts generating high engagement on provably false claims that are similar to posts that were labeled or removed months prior.[174]

A report on Facebook's content moderation failures from advocacy group Avaaz found that Facebook's failure to take down misinformation resulted in over 10.1 billion estimated views of content from top-performing pages that repeatedly shared misinformation over the eight months before the U.S. elections.

In the 117th Congress, both Democrats and Republicans have proposed to modify or outright repeal Section 230,[175] but as of this writing, none of these proposals have been passed into law. Introduced legislation generally falls into a few different categories: (1) bills that limit the scope of Section 230, (2) bills that impose new obligations on companies that want to use Section 230 as a defense, (3) bills that want to make changes to the "Good Samaritan" provision of Section 230, and (4) bills that repeal Section 230 outright.[176] A number of bills introduced by Republicans, like Representative Louie Gohmert's (R-TX) Abandoning Online Censorship Act[177] and Senator Bill Hagerty's (R-TN) 21st Century Foundation for the Right to Express and Engage in (FREE) Speech Act,[178] would repeal Section 230 outright. These proposals are often predicated on the false idea that social media platforms are "censoring" conservatives, and this is reflected in public statements from the sponsors of the legislation.[179] Other proposals are bipartisan and would make less drastic changes to Section 230. For example, the Platform Accountability and Consumer Transparency Act, introduced by Senator Brian Schatz (D-HI) and Senator John Thune (R-SD), would condition Section 230 immunity on the publication of an acceptable use policy that would detail the types of content the provider allows, explain how the provider enforces its content policies, and describe

how users can report policy-violating or illegal content.[180] Most recently, Representatives Frank Pallone (D-NJ), Mike Doyle (D-PA), Jan Schakowsky (D-ILL), and Anna Eshoo (D-CA) announced that they would be introducing legislation that would amend Section 230 to remove immunity for platforms that knowingly or recklessly use an algorithm or other technology to recommend content that materially contributes to physical or severe emotional injury.[181]

As demonstrated by the number of proposals introduced in the 117th Congress, no one has a silver bullet solution to reforming Section 230 given the challenges and unintended consequences amending the statute could create. If Congress amends the statute, it could significantly limit free expression online, diminish the internet as a tool for grassroots mobilization, and open the door to liability for smaller websites and online companies, further cementing the dominance of large social media platforms.[182] Therefore, any Section 230 reform deserves careful and nuanced consideration.

Federal Consumer Protection Laws

The Federal Trade Commission (FTC) is charged with protecting consumers and promoting competition.[183] Its primary consumer protection authority comes from Section 5 of the FTC Act, which prohibits "unfair or deceptive acts or practices in or affecting commerce."[184] An "unfair" act or practice is defined as an act or practice that causes or is likely to cause substantial injury to consumers, cannot reasonably be avoided by consumers, and is not outweighed by countervailing benefits to consumers or competition.[185] A representation, omission, or practice is "deceptive" if it "is likely to mislead consumers acting reasonably under the circumstances and is material to consumers—likely to affect the consumer's conduct or decision with regard to a product or service."[186]

Historically, the FTC's Bureau of Consumer Protection has used its Section 5 authority to bring enforcement actions for a wide range of privacy and data security violations, such as deceptive data collection and failure by a company to adequately assess and address data security risks.[187] Remedies include requiring violators to implement comprehensive privacy and security programs, deletion of illegally obtained consumer information, and providing robust transparency and choice mechanisms for consumers.[188]

Bad actors have exploited unfair and deceptive data collection practices to help spread disinformation, and the FTC has used its authority to enforce against these actions. In 2016, for example, political data analytics and consulting company Cambridge Analytica used an app to collect the personal data of millions of Facebook users without their consent.[189] The company was able to find out where people worked, what they looked like, where they lived, what kind of car they drove, who they've voted for in past elections, what kind of music they liked, how much money they made, whether or not they were married, whether or not they owned a gun, and more all without their consent.[190] This data was then used to develop psychological profiles to

Bad actors have exploited unfair and deceptive data collection practices to help spread disinformation, and the FTC has used its authority to enforce against these actions.

help the Donald Trump and Ted Cruz presidential campaigns target voters with false or misleading political ads and allowed other bad actors to spread disinformation.[191]

The FTC launched an investigation into both Cambridge Analytica and Facebook for engaging in deceptive acts and practices in violation of Section 5 of the FTC Act. The investigation into Facebook culminated in a $5 billion penalty against the company for violating a 2012 consent decree, which prohibited Facebook from making misrepresentations about the privacy or security of its users' personal information.[192] Facebook had undermined user privacy preferences by deceiving users when the company shared the data of users' Facebook friends with third-party developers and by misrepresenting the ability of users to control the use of facial recognition technology with their accounts, among other violations.[193] The FTC also issued an opinion and order finding that Cambridge Analytica engaged in deceptive practices to harvest the personal information of tens of millions of Facebook users for voter profiting and targeting.[194] The FTC's order prohibits Cambridge Analytica from making misrepresentations about the extent it protects the privacy and confidentiality of personal information.[195]

The investigation into Facebook culminated in a $5 billion penalty against the company for violating a 2012 consent decree, which prohibited Facebook from making misrepresentations about the privacy or security of its users' personal information.

While the FTC has broad authority to enforce against unfair and deceptive practices, there are limits to the effectiveness of its current enforcement capabilities. First, the FTC is limited in its ability to seek civil penalties for first-time violations of Section 5, and in many cases, the agency levies penalty fines against companies for violations of consent decrees.[196] Second, the FTC is severely limited in its resources—its budget is roughly $350 million,[197] and it only has around 40 staffers working on privacy issues.[198] This pales in comparison to the budgets of other privacy enforcement agencies around the world.[199] If the FTC is to meaningfully protect consumers from a myriad of privacy violations, many of which lead to the spread of disinformation, it will need adequate funding. Finally, the FTC's current enforcement actions have proven inadequate in changing the business models and practices of the largest online companies. For example, Facebook's stock went up after the agency imposed a record-breaking $5 billion fine on the company.[200] As then-FTC commissioner Rohit Chopra noted in his dissenting statement, the $5 billion settlement imposes no meaningful changes to the company's structure nor does it include any changes to the company's surveillance and advertising practices that exposed millions of users to propaganda, manipulation, and discrimination.[201] Future FTC oversight and enforcement must be able to address corporate business models that lead to the spread of disinformation and other harmful content.

State Media Literacy Laws

People of all ages need media literacy skills now more than ever to tackle the myriad of problems caused by disinformation. In 2019, a Stanford University study found 52% of students assessed

believed a grainy video claiming to show ballot stuffing in the 2016 Democratic primaries (the video was actually shot in Russia) constituted "strong evidence" of voter fraud in the United States and concluded today's high school students "lack the skills to judge the reliability of information online."[202] Teaching media literacy in K–12 schools is critical to providing young people with the skills they need to navigate the internet, critically evaluate the content received and consumed online, and protect them from misinformation.[203]

Students are not the only cohort that needs media literacy skills. Older individuals have been found to engage with fake news at a disproportionately higher rate than younger people, and a study of Twitter during the final month of the 2016 presidential election showed users over 50 were overrepresented among users responsible for spreading 80% of fake content.[204] Given that older individuals are more likely to register and vote, it is equally as important that this group is able to learn media literacy skills.[205]

While not a silver bullet by any stretch of the imagination, a greater emphasis on the skills necessary to discern trustworthy from untrustworthy opinion, fact from opinion, news from infotainment, and real information from misinformation will go a long way toward protecting our democracy.[206]

As of August 2021, roughly 15 states have some variation of media or information literacy laws on the books. APPENDIX III summarizes several of these states' laws. States have taken a wide variety of approaches, including requiring media literacy classes in schools, providing resources for teachers, and developing state media literacy committees.[207] Earlier this year, for example, **Illinois** passed a media literacy law requiring every public high school in the state to include in its curriculum a unit of instruction on media literacy, **making it the first state to mandate media literacy classes**.[208] And in 2019, **Colorado** created a media literacy advisory committee within the Colorado Department of Education, which later that year submitted a report to the General Assembly recommending revision of Colorado academic standards, provision of materials and resources to teachers, and legislation to support effective implementation of media literacy programs in schools throughout the state.[209]

Earlier this year Illinois passed a media literacy law requiring every public high school in the state to include in its curriculum a unit of instruction on media literacy, making it the first state to mandate media literacy classes.

Media literacy laws have gained traction in the past few years because they do not run into the same First Amendment concerns that other laws designed to target misinformation may face.[210] Additionally, media literacy laws are often able to find bipartisan support, as most of the laws discussed earlier were enacted with support from both Republicans and Democrats in their respective state legislatures. State governments should study best practices around media literacy in collaboration with expert organizations like PEN America and experiment with legislation based on best practices.

State Privacy Laws

Privacy laws (or lack thereof) play an important role in how misinformation is allowed to spread on the internet and how we can combat it. Access to personal data gives bad actors the ability to target individuals with a precision that has never been seen before. Without detailed data about a user's political beliefs, age, location, and gender, it is far more difficult for bad actors to target them with disinformation.[211] To quote Alex Campbell in his piece for *Just Security,* "Fake news becomes a lot less scary if it can't choose its readers."[212]

While efforts to pass a comprehensive federal privacy law have stalled, a few states have passed legislation. These laws vary in scope, but each of them requires a company operating in the state to inform users if they are selling the users' data and gives users the right to access, delete, correct, or move their data.[213]

California became the first state with comprehensive consumer privacy laws on the books when its legislature passed the California Consumer Privacy Act of 2018 (CCPA) and expanded it in 2020 with the Consumer Privacy Rights Act. It provides consumers with the right to know about the personal information a business collects about them, the right to delete personal information collected from them, the right to opt out of the sale of their personal information, and the right to nondiscrimination for exercising their CCPA rights.[214] The CCPA applies to companies that generate more than $25 million a year in revenue; buy, receive, or sell the personal information of 50,000 or more California residents; or derive 50% or more of their annual revenue from selling California residents' personal information.[215] Companies that do not comply can be fined by the California Office of the Attorney General. **The CCPA is considered by advocates to be the strongest of the state privacy laws on the books,** in part because it contains a limited private right of action against certain types of breaches, which allows consumers to directly sue the company committing the breach.[216]

California became the first state with comprehensive consumer privacy laws on the books.

However, while CCPA has some strong elements, it is important to recognize where it (and the other state privacy laws) falls short. Strong, comprehensive privacy legislation should have data minimization requirements limiting what data entities can collect and how that data can be used, as well as civil rights protections that ensure fairness in both automated decision-making and prohibitions on the use of personal data to discriminate on the basis of race, gender, religion, national origin, sexual orientation, gender identity, disability, familial status, biometric information, or lawful source of income.

Colorado is the most recent state to enact privacy legislation, having passed the Colorado Privacy Act (CPA) in July of 2021.[217] The law shares similarities with the privacy laws of Virginia and California, as it also allows consumers to opt out of data collection while requiring companies to disclose what data they collect, what they do with that data, and how long they keep it.[218] Like Virginia's law, Colorado's law applies to entities that "control or process" the information of 100,000 or

more residents or entities that make 50% or more of their gross revenue from the sale of personal data if they hold the information of about 25,000 or more consumers.[219] Also, like the Virginia law, it only applies to "Colorado residents acting only in an individual or household context."[220] One place where the CPA differs slightly from Virginia's law (and California's) is in enforcement. Under the CPA, both the Colorado attorney general and district attorneys have enforcement authority and can bring actions against businesses that violate the law.[221]

The **Virginia** Consumer Data Protection Act (VCDPA) was passed by the Virginia General Assembly in 2021 with bipartisan support. The VCDPA gives consumers the same core rights California's CCPA does but applies to entities that "control or process" the information of 100,000 or more Virginia residents in a calendar year or entities that make 50% or more of their gross revenue from the sale of personal data if they hold information for about 25,000 or more Virginia residents.[222] It provides Virginia residents with the right to confirm if a controller has their data, the right to correct inaccuracies in the data the controller has, the right to have a controller delete personal data provided by or obtained about them, and the right to opt out of having their data used for targeted advertising.[223]

However, unlike California's law, the VCDPA does not contain a private right of action and is written more narrowly, as it only covers individuals acting on their own or in a "household context," not those acting in a "commercial or employment context."[224] The lack of a private right of action is problematic because it means the only party that can enforce the law is the Virginia Attorney General's Office, and the General Assembly only gave the Attorney General's Office $400,000 in additional funding to do so.[225]

SECTION 3: SELECT SOCIAL MEDIA CIVIC INTEGRITY POLICIES

Social media platforms from Facebook to Twitter and YouTube to TikTok have civic integrity policies in place designed to combat disinformation related to elections and other civic processes.[226] These policies often work in tandem with the platforms' other policies, which address things like fraud, violent content, hate speech, and other content the platform may find objectionable.[227] A piece of content may violate multiple policies at once, like a post advocating violence against a specific group.

Platform civic integrity policies primarily focus on prohibiting content that is misleading about how to participate in the civic process. This includes misleading statements or information about the official announced date or time of an election,[228] misleading information about requirements to participate in an election,[229] and content containing statements advocating for violence because of voting, voter registration, or the administration or outcome of an election.[230]

These policies are not exhaustive though and have significant loopholes that allow for certain disinformation-oriented content to stay up on the platforms. This includes narratives contributing to voter suppression, disinformation from world leaders/public figures, and political ads.

This is in part because platforms are frequently changing and updating their policies. For example, the Mozilla Foundation (Mozilla), a nonprofit whose advocacy work includes using data visualization and original reporting to track the ways the internet is helping and hurting users around the world, found that during the 2020 election cycle (October 2019 through January 2021), Facebook changed its election-related misinformation policies 21 times, Twitter changed its policies 16 times, and YouTube changed its policies 12 times.[231] Most of these changes involved adding, subsequently rolling back, and then reinstating new rules concerning key issues like mail-in voting fraud or false victory claims.[232]

During the 2020 election cycle (October 2019 through January 2021), Facebook changed its election-related misinformation policies 21 times, Twitter changed its policies 16 times, and YouTube changed its policies 12 times.

While Mozilla was able to track the policies of Facebook, Twitter, and YouTube during the 2020 election, it concluded that "there remains a persistent lack of data about how well these policies were enforced and their impact on election misinformation."[233] This reflects a gap in understanding between the public and the platforms about which policies were most effective and which were not.[234] It is also concerning because of how prevalent these platforms have become in our nation's politics and the way the Big Lie about how the 2020 election was "rigged and stolen" from Trump has metastasized via social media platforms throughout our national civic dialogue.

During the 2020 election specifically, mainstream social media platforms like Facebook, Twitter, and YouTube expanded their policies and enforcement against election-related disinformation, removing or labeling many dangerous false claims of widespread voter fraud.[235] These changes generally correlated with major events like the beginning of the impeachment inquiry into then-president Donald Trump, the death of George Floyd, Trump's first public claim that mail-in ballots would lead to election fraud, and his first public refusal to commit to accepting election results, Election Day, and when the Electoral College confirmed President Biden's victory.[236]

Next, we summarize only the policies that Facebook, Twitter, and YouTube implemented during the 2020 elections and soon after. We also discuss how inconsistent enforcement and policy loopholes led to the spread of disinformation during and after the election, how the actions taken (or not taken) by the platforms contributed to the insurrection at the Capitol complex on January 6, and how the platforms reacted in the aftermath.

Unfortunately, Facebook and Twitter have stopped enforcing existing policies to the degree they did during the 2020 election.[237] Our research shows that there are many pieces of content being left on the platform that would have been taken down months ago.[238]

Facebook

It has been well documented that Facebook is inconsistent in its enforcement of existing policies. In September of 2020, the *Wall Street Journal* flagged over 200 pieces of content for Facebook that appeared to violate the platform's rules against the promotion of violence and dangerous information, only to have Facebook respond by taking down around 30 pieces of flagged content and later conceding that more than half of the pieces of content should have been taken down for violating their policies.[239]

In addition to inconsistent enforcement, Facebook also had two major loopholes that contribute significantly to the spread of disinformation on the platform: the newsworthiness exemption and its policy of not fact-checking political ads.

The newsworthiness exemption applies to any content that Facebook believes "should be seen and heard"[240] and meets a balancing test that weighs the public benefit of having the content up versus the harm keeping the content in question up could cause.[241] This is extremely subjective, and this subjectivity is reflected in Facebook's use of the newsworthiness exemption over time. Through 2020 and the first half of 2021, content from certain users, including politicians, was presumed to be newsworthy and left up.[242] However, following criticism from its oversight board, Facebook eliminated the presumption that posts by politicians are automatically considered newsworthy.[243] While this is a step forward, Facebook is still able to apply its newsworthiness exemption to any piece of content it chooses without giving much justification as to why the content is left up. This gives politicians and other bad actors with large public followings the ability to spread disinformation that Facebook may consider "newsworthy" with confidence that it is unlikely to be taken down.

Facebook's decision to exempt political ads has proven to be equally controversial, if not more, than their newsworthiness exemption. This loophole is straightforward: Facebook will not fact-check political advertisements on the platform.[244] During the 2020 election, then-candidate Donald Trump took advantage of this loophole several times and placed ads on Facebook intending to mislead voters about then-candidate Joe Biden and his son Hunter.[245] If Facebook is to get serious about cracking down on disinformation, this loophole is one of the first they need to address.

This laissez-faire approach to content moderation allowed bad actors to spread content that contributed to the January 6 insurrection.[246] Right-wing groups were able to use Facebook's Groups feature to plan their assault on the Capitol building, while prominent pages pushed harmful content delegitimizing the election and took advantage of relaxed enforcement of live videos to urge violence.[247]

This laissez-faire approach to content moderation allowed bad actors to spread content that contributed to the January 6 insurrection.

Following the insurrection on January 6 and significant criticism by both members of Congress and civil society, Facebook made a few different changes. This included suspending President Trump's account indefinitely (which was later reduced to two years after the indefinite suspension was appealed to Facebook's oversight board), increasing its monitoring for calls to violence and protest, and updating its election label to read, "Joe Biden has been elected President with results that were certified by all 50 states. The US has laws, procedures, and established institutions to ensure the peaceful transfer of power after an election."[248] *Additionally, the company put in place heightened penalties for public figures "during times of civil unrest and ongoing violence."*[249] However, it is important to note that while Facebook gives an example of actionable content (someone sharing a link to a statement from a terrorist group in the aftermath of an attack), they do not define what constitutes a period of "civil unrest and ongoing violence."[250]

Significant questions exist as to how seriously Facebook takes the threat of disinformation. Even the changes Facebook made following January 6 are riddled with loopholes. As Common Cause has documented, former president Trump is still able to run political ads on Facebook and solicit donations from supporters, even though he is suspended from the platform.[251] And as with most Facebook policies, the new rules put in place are arbitrary and subject to human discretion.

Twitter

Although Facebook tends to dominate the conversation about content moderation practices and the spread of disinformation on social media, Twitter is guilty of many of the same things: inconsistent enforcement of existing policies, loopholes in policies that allow for the spread of disinformation, and relatively weak policy responses to the January 6 insurrection. While Twitter may want to be viewed as better on content moderation than its peers, it has been equally as slow to deal with the misinformation that is found all over the platform.

Media Matters highlighted Twitter's inconsistent enforcement in a post discussing the platform's treatment of a doctored video of House Speaker Nancy Pelosi (D-CA) appearing to slur her words.[252]

While one version of the video on the platform has received a "manipulated media" tag, other versions of the video remained on the platform untouched.[253]

Just like Facebook's newsworthiness exemption, Twitter has a major loophole that contributes significantly to the spread of disinformation called the "public interest exception." This exception applies to tweets from elected and government officials that Twitter believes "directly contribute" to the understanding or discussion of a matter of public concern.[254] Tweets that are found to be in the public interest but break other rules may have a label put on them but will not be taken down.[255] Even though the platform insists that this does not mean public officials can post whatever they want (even tweets in violation of their rules), in reality, public officials are generally allowed to get away with posting whatever they want.[256] The consequences of this loophole were on full display when then-outgoing president Donald Trump live tweeted throughout the insurrection, adding gas to the fire by inciting his supporters while they stormed the Capitol.[257]

Twitter permanently suspended Donald Trump's account and CEO Jack Dorsey acknowledged the platform's role in the insurrection, but issues remain.[258] Today, public officials like Rep. Marjorie Taylor Greene (R-GA) are able to take advantage of this exemption, and if Twitter is serious about getting rid of disinformation on the platform, they also need to look into closing this loophole.[259]

YouTube

Compared to Facebook and Twitter, YouTube's policies have not been scrutinized to the same degree, but like the other social media platforms mentioned here, YouTube is also inconsistent in its enforcement of existing policies.[260] **However, instead of having one or two major loopholes in which disinformation is able to spread, YouTube's policies are overall far more permissive than that of Facebook and Twitter.**[261]

YouTube's inconsistency in policy enforcement is well documented. In 2019, the platform announced that it would be making changes to its hate speech policy and taking down thousands of videos that were in violation of the new policy, but Gizmodo found that many of the videos remained up.[262] To make matters worse, YouTube's own algorithm will frequently recommend content that violates its own policies.[263] One example of this is a user watching a music video from Art Garfunkel, one half of the popular 1960s pop duo Simon & Garfunkel, recommending a video titled *Trump Debate Moderator EXPOSED as having Deep Democrat Ties, Media Bias Reaches BREAKING Point.*[264] A study done by Mozilla found that nearly 200 videos YouTube's algorithm recommended to volunteers had a collective 160 million views before the platform took them down for violating YouTube's policies.[265]

Like its peers, YouTube took some actions following the January 6 insurrection. First, YouTube suspended Donald Trump's account until the risk of violence associated with the account had decreased.[266] Second, they introduced new rules, giving "strikes" to channels whose videos violate the platform's policies and permanently removing channels that receive three strikes within the same 90-day period.[267] Since YouTube has not explained what they mean by "risk of violence," it is unclear when and if they will let Donald Trump back on the platform or how they will apply this standard to other accounts in the future.

SECTION 4: RECOMMENDATIONS

Federal laws and the laws of many states contain important provisions to reduce the harmful impact of election disinformation. Social media company civic integrity policies are likewise critically important. These current laws and policies leave much room for improvement. There is no single policy solution to the problem of election disinformation. We need strong voting rights laws, strong campaign finance laws, strong communications and privacy laws, strong media literacy laws, and strong corporate civic integrity policies. In this section, we recommend reforms in all these policy areas, highlighting both pending legislation that should be passed and existing state laws that should be replicated in other jurisdictions.

There is no single policy solution to the problem of election disinformation. We need strong voting rights laws, strong campaign finance laws, strong communications and privacy laws, strong media literacy laws, and strong corporate civic integrity policies.

Statutory Reforms

Voter Intimidation and False Election Speech Reforms

The Voting Rights Act of 1965 and other federal laws prohibit voter intimidation and have been interpreted by the DOJ and courts as prohibiting certain forms of election disinformation.[268] **However, existing federal statutes and the statutes in most states do not explicitly prohibit election disinformation and should be amended to do so.** And as the *New York Times* wrote earlier this year in response to the DOJ prosecution of Douglass Mackey for disseminating election disinformation via Twitter, the "case will test the novel use of federal civil rights laws as a tool to hold people accountable for misinformation campaigns intended to interfere with elections[.]"[269] Rather than relying on courts to appropriately apply long-standing statutes to new modes of election disinformation, Congress and state legislatures should update statutes to explicitly prohibit election disinformation.

Congress should enact the Deceptive Practices and Voter Intimidation Prevention Act of 2021,[270] which would modernize federal law to address the worst election disinformation practices being used today. This bill has been incorporated into both the Freedom to Vote Act[271] and the For the People Act.[272] It twice-passed the House in 2019 as part of the For the People Act and the Stopping Harmful Interference in Elections for a Lasting Democracy (SHIELD) Act but died in the Senate. Among other things, this legislation would amend the anti-intimidation statute at 52 U.S.C. § 10101(b) to add a new subsection **explicitly outlawing false statements regarding federal elections**. Under this amendment, it would be illegal to knowingly disseminate materially false information within 60 days before a federal election regarding the time, place, or manner of holding any federal election or the qualifications or restrictions on voter eligibility—with the intent to impede or prevent another person from exercising the right to vote in an election.[273] Importantly, the bill not only contains criminal enforcement provisions but also would create a private right of action, enabling those harmed by disinformation to file a civil lawsuit against the

perpetrator.[274] Finally, the legislation would require the attorney general to communicate with the public to correct any materially false election information if state and local election officials have failed to do so.[275]

States should likewise enact legislation explicitly prohibiting dissemination of false election speech, modeled on the federal Deceptive Practices and Voter Intimidation Prevention Act of 2021[276] or similar legislation already enacted in Virginia prohibiting knowingly communicating false information "intended to impede the voter in the exercise of his right to vote," including information "about the date, time, and place of the election, or the voter's precinct, polling place, or voter registration status, or the location of a voter satellite office or the office of the general registrar."[277]

Congress should also adopt a proposed amendment to the National Defense Authorization Act, a military appropriations and policy bill, which would require national intelligence agencies to include in their regularly scheduled post-election report the identification of any Russian government official or agent who used "social or traditional media to spread significant amounts of false information to individuals in the United States" as a form of election interference.[278] And Congress should go beyond this Russia-specific amendment to require identification of any foreign government spreading disinformation to interfere in U.S. elections.

Campaign Finance Reforms

Strong, up-to-date campaign finance disclosure laws are key to curbing the harmful impacts of election disinformation. Whereas social media and other internet platforms have become favored means of disseminating election disinformation, federal campaign finance disclosure laws and the laws of most states were written decades ago, before these avenues for disinformation existed. **Congress and state legislatures should update disclosure laws so that the public has easy access to accurate information regarding who is spending money on political advertising online and through more traditional modes of communication.**

Congress should pass the DISCLOSE Act of 2021[279] and the Honest Ads Act,[280] both of which have been incorporated into the Freedom to Vote Act[281] and the For the People Act.[282] The For the People Act passed the House in March 2021. The Freedom to Vote Act is Senate legislation that includes most of the core pillars of the For the People Act, including the DISCLOSE Act and the Honest Ads Act. The DISCLOSE Act would **shine a light on presently dark money in federal elections** by expanding the definition of what constitutes a reportable "campaign-related disbursement" to include certain ads that support or oppose candidates and transfers between organizations—a tactic used to evade disclosure under current law.[283] The DISCLOSE Act would also strengthen "paid for by" disclaimers for robocalls, a popular mode of communication for election disinformation.[284] The Honest Ads Act would require political ads sold online to be covered by the same rules as political ads sold on TV, radio, and satellite.[285] It would also expand disclosure rules to include any online ads that mention a candidate and require social media platforms and websites with 50 million or more unique monthly visitors that sell political advertising to maintain a database of all online political ads—both necessary reforms to improving transparency around online ads.[286]

Congress should also act on the FEC's recommendation to amend and strengthen the "fraudulent misrepresentation of campaign authority" statute by expanding the law to prohibit any

person—not just candidates and their agents—from fraudulently claiming to be acting on behalf of a candidate or political committee.[287]

And **Congress should restructure and strengthen the FEC to ensure more effective implementation and enforcement of all federal campaign finance laws,** as would result from the passage of the Restoring Integrity to America's Elections Act,[288] which was last introduced in 2019 but has been incorporated into both the For the People Act[289] and the Freedom to Vote Act.[290]

Similarly, states should follow the leads of Alaska, California, New York, Washington, and the handful of other states profiled earlier in this report that have enacted campaign finance disclosure laws to bring previously dark money spending on digital and other election advertising into the light. Alaska's[291] and California's[292] laws that enable voters to trace election spending to the source are similar to the reforms contained in the federal DISCLOSE Act, all providing models to emulate throughout the states. California,[293] New York,[294] and Washington[295] have enacted cutting-edge requirements for digital political ad "paid for by" disclaimers and public access to digital ad databases that can serve as models for the nation.

State Media Literacy Laws

Common Cause does not recommend any specific media literacy laws, but **we do encourage policymakers to experiment with best practices around media literacy** and advocacy groups interested in pushing for media literacy laws to work with stakeholders to determine the approach that best fits their state. This involves holding convenings and bringing to the table organizations like PEN America, which are already engaged in the issue and offering media literacy training to the public, to put together a set of principles and best practices on which to develop legislation.[296] Media Literacy Now has put together a model bill that established an advisory council within the respective state's Department of Education,[297] which is one example of an approach that could be taken.

State Privacy Laws

While California, Colorado, Virginia, and Washington have all successfully passed comprehensive privacy legislation, each bill was lacking in some respects. None of them have specific civil rights protections, and only California has a (very limited) private right of action. **Advocates interested in passing comprehensive privacy legislation should look to the Digital Fairness Act, which was introduced in the New York State Assembly this year, as a model.**[298] The Digital Fairness Act includes heightened protections for biometric information, strong civil rights protections by explicitly making it unlawful to process personal information or target advertising in ways that discriminate in employment, finance, health care, credit, insurance, housing, education opportunities, or public accommodations based on an individual's or class of individuals' actual or perceived age, race, creed, color, national origin, sexual orientation, gender identity or expression, sex, disability, predisposing genetic characteristics, or domestic violence victim status, as well as and a robust private right of action for consumers whose rights are violated.[299] **Any legislation considered needs to have strong civil rights protections, a strong private right of action, and strong data minimization provisions.** Further, if a state Attorney General's Office is going to play a meaningful role in enforcement, it must receive enough funding to be able to effectively bring lawsuits and protect consumers in their state.

Federal Legislative Reforms to Mitigate Platform Business Practices

Algorithmic Accountability

Social media algorithms have contributed to the spread of disinformation given that platforms have optimized them for user engagement, which has led users down a rabbit hole of hate speech, conspiracy theories, and harmful content.[300] Algorithms can also promote the amplification of disinformation as conspiracy theorists used the "stop the steal" moniker across platforms to organize and mobilize offline violence.[301] To hold platforms accountable for the algorithms they deploy, **we urge Congress to pass the Algorithmic Justice and Online Platform Transparency Act.**[302] The legislation would **prohibit discriminatory algorithms and create greater transparency about how these algorithms operate.** The bill also addresses election disinformation specifically by prohibiting online platforms from processing personal information "in a manner that intentionally deprives, defrauds, or attempts to deprive or defraud any individual of their free and fair exercise of the right to vote in a Federal, state, or local election."[303]

Comprehensive Privacy Legislation

Comprehensive federal privacy legislation is key to placing limits on the collection and sharing of personal data, which has contributed to the spread of disinformation. As discussed, social media platforms collect vast amounts of data on their users, and bad actors exploit these data practices by targeting harmful content. For example, the Trump campaign used Facebook to target millions of Black voters with deceptive information to deter them from voting.[304]

At a minimum, federal legislation should (1) require companies to minimize the data they collect; (2) prohibit predatory and discriminatory data practices on the basis of protected characteristics with respect to access to credit, housing, education, employment, and public accommodations; (3) provide for fairness in automated decision-making; (4) grant a private right of action to allow consumers to sue companies that violate their privacy rights; and (5) define permissible and impermissible uses for collecting, sharing, and using personal data.[305]

Strengthening Local Media

Local media plays a critical role in supporting civic engagement and provides communities with vital information on issues such as public safety, economic development, and health care. Unfortunately, the decade-long decline in local media has robbed communities of critical news and information, creating an "infodemic" that has helped disinformation flourish.[306] The economic decline in local news can be attributed in part to large social media platforms that are now dominating the advertising market, making the ad-driven business model for journalism unsustainable.[307]

Congress can pass **legislation that funds local media and community and public media of all kinds.** Funds should be targeted at preserving newsrooms and reporting jobs at local commercial and nonprofit news outlets, as well as investments to address the civic information needs of communities most affected by the long-term decline of local news. In addition to short-term spending, we need long-term solutions about how journalism can meet the civic information needs of communities in the 21st century. To that end, Congress should pass the Future of Local News Act, which would **create a committee to study the state of local journalism and offer recommendations to Congress.**[308] The bill would provide a first step in determining what transformative investments are needed to create a sustainable local journalism landscape.

Empowering Researchers and Watchdog Journalists

Transparency is more important now than ever, as political campaigns and disinformation agents use manipulative tactics to target voters on social media platforms. Ensuring researchers can study the major platforms without fear of interference is crucial to understanding how misinformation spreads and for developing policies that address the many harms to society the platforms have caused.

Congress should **pass legislation that ensures researchers and watchdog journalists have sufficient access to social media data and protect them from retaliation by the platforms.** New York University researcher Laura Edelson, whose account was banned by Facebook while she was researching ad transparency and the spread of misinformation, recommended in her recent congressional testimony that there be universal digital ad transparency, a researcher safe harbor law, and greater access to public content with "meaningful reach or content from public figures with meaningful audiences."[309] Members of Congress have also put forward their own legislative solutions. Representative Lori Trahan (D-MA) has introduced legislation that would require covered platforms to grant academic researchers and the FTC access to an ad library with select information about each ad.[310]

Executive and Regulatory Agency Reforms

In addition to legislative reforms to fight election disinformation, there are regulatory tools and other actions federal and state executive branch agencies can take to combat disinformation, including stronger enforcement of existing laws and promulgation of new regulations to rein in social media business practices that bad actors exploit to spread and amplify harmful content that interferes with our democracy. As a general matter, statutory reform is preferable to executive and regulatory agency reform because some types of executive and regulatory agency reform can easily be reversed when a new president's or governor's administration comes into power. Nevertheless, executive branch reforms may be easier to attain than the passage of legislation and can play an important role in reducing election disinformation.

Presidential and Gubernatorial Leadership

The White House under the Biden administration must play a leading role in combating election disinformation. A recently published report from the U.S. surgeon general shows that the Biden administration has already recognized that the spread of COVID-19 misinformation poses serious risks to the nation's public health.[311] The surgeon general's report identified several recommendations the government, social media platforms, and other stakeholders can take to stop the spread of false content related to the pandemic.[312] Similarly, the Biden administration should take a whole-of-government approach to combating election disinformation. To start, the administration can issue an executive order directing federal agencies with enforcement, rule-making, and investigatory authorities to use these capabilities in combating election disinformation.

Next, the administration should create a federal interagency task force that would identify tools to combat election disinformation and harmful online speech.[313] The task force, composed of senior officials from executive agencies such as the DOJ and Department of Commerce and independent agencies such as the FTC, among others, would develop initiatives to mitigate the impact of disinformation, particularly on African American communities and other communities

of color that are disproportionately targeted by disinformation campaigns.[314] Given the scope and complexity of how disinformation spreads, a whole-of-government approach is necessary, and the White House must lead on this initiative.

Similarly, governors in states around the nation can and should play a leading role in stopping election disinformation, establishing task forces like the one described earlier, and using all available resources of state agencies under their control.

U.S. DOJ and State Law Enforcement Agencies

Existing statutes give federal and state law enforcement officials many tools to fight election disinformation. The federal DOJ, for example, has long interpreted voting rights laws as prohibiting election disinformation that interferes with the fundamental right to vote.[315] However, the DOJ prosecution begun earlier this year against Twitter user Douglass Mackey[316] for illegally disseminating election disinformation in 2016 appears to be the first criminal prosecution in the United States involving voter suppression through the spread of disinformation on Twitter.[317] **The DOJ should be more aggressive in its criminal prosecution and civil litigation against those who use disinformation to intimidate voters and interfere with their voting rights.**

State law enforcement officials should likewise use state laws prohibiting voter intimidation, election interference, and false statements regarding elections—including those that do not explicitly name election disinformation as a form of illegal interference—to stem the tide of election disinformation and hold perpetrators accountable.

FTC Reforms

There are several different actions the FTC could take to improve its role as the country's privacy enforcement agency. Under the current administration, the FTC could expand the scope of its rule-making and enforcement practices. Senate Democrats[318] and civil society groups[319] have both asked the FTC to **initiate rule-making to regulate unfair and deceptive commercial data practices**. This rule-making would consider strong protections for members of marginalized communities, data minimization practices, prohibitions on certain practices, opt-in consent rules on the use of personal data, and global opt-out standards.[320] In addition to rule-makings, the FTC can conduct workshops and issue informal guidance on how platforms can provide greater transparency in their content moderation practices.

State law enforcement officials should likewise use state laws prohibiting voter intimidation, election interference, and false statements regarding elections.

FEC and State Election Agency Reforms

The FEC has an important role to play in combating disinformation in federal elections. The FEC is our nation's frontline enforcer of campaign finance disclosure laws in federal elections. Yet, despite the proliferation of online political advertising over the past decade-plus, the FEC has failed to update its "paid for by" disclaimer rules for digital ads. In October 2011, the Commission

published an advance notice of proposed rule-making on "internet communication disclaimers"[321] and has, over the decade, invited public comment on the issue several times. Common Cause filed comments in 2018 on behalf of more than 25,000 members and supporters urging the Commission to adopt regulations applying to digital ads the full disclaimer requirements now applicable to radio, television, and print ads.[322] But today, more than a decade after the rule-making began, the Commission still has not adopted final regulations. With bad actors continuing to target Black and other communities of color with election disinformation using digital political ads, **it is long past time for the FEC to promulgate clear and enforceable disclaimer rules for online political advertising.**

State campaign finance agencies similarly have an important role to play in implementing and enforcing effective disclosure laws to shine a light on those trying to undermine our elections with disinformation. All states' campaign finance enforcement agencies should follow the leads of the Washington Public Disclosure Commission, the California Fair Political Practices Commission, and others that have worked hard to effectively apply campaign finance laws to the digital landscape.

Social Media Corporation Policy Reforms

While self-regulation on its own has proven ineffective in curbing the spread of disinformation, **social media platforms must take additional steps to strengthen their policies on combating content designed to undermine our democracy.** The recommendations that follow focus on how platforms can improve their efforts to provide users with authoritative information concerning voting and elections, reduce the spread and amplification of election disinformation, and provide greater transparency regarding their content moderation policies and practices. While not an exhaustive list, these recommendations represent the basic steps all social media platforms should take.

Provide Users With Authoritative Information About Voting and Elections

Platforms should help users identify official voting information such as registering or updating their registration, tracking ballots in the mail, and identifying in-person polling sites. *Platforms should direct their users to authoritative sources of information regarding voting and elections. Authoritative sources come from state and local election officials.*

Consistent Enforcement of Civic Integrity Policies During Both Election and Nonelection Cycles

Platforms have failed to consistently enforce the civic integrity policies they have in place to combat the spread of election disinformation.[323] Further, enforcement tends to become more relaxed during nonelection cycles. **Platforms must commit to upholding their own civic integrity policies** and consistently enforce them throughout election, as well as nonelection, cycles. Consistent enforcement includes rapid removal, labeling, and de-prioritizing of content that violates civic integrity policies.

Platforms should also **close loopholes in their civic integrity policies** bad actors exploit to spread disinformation. For example, platforms should apply third-party fact-checkers to political adver-

tisements and remove exemptions for public figures that allow them to spread disinformation with impunity.

Reducing the Spread and Amplification of Disinformation

Platforms must **reduce the spread and amplification of disinformation caused by the algorithms they deploy**. As discussed, platforms optimize their algorithms to maximize user engagement. Content that generates the most engagement and gets amplified tends to focus on lies, conspiracy theories, and incitements of violence. Platforms can take steps to limit amplification by fashioning artificial intelligence systems and algorithms so that engagement does not prioritize disinformation. Further, platforms should conduct third-party human and civil rights audits of their algorithms to ensure voter suppression content is not getting amplified.

Provide Researchers and Watchdog Journalists Greater Access to Social Media Data

Platforms must **provide researchers and watchdog journalists with sufficient and reliable access to social media data.** As discussed, researchers and watchdog journalists play a critical role in shedding light on how platforms enforce and interpret their content moderation policies in practice. For example, researchers exposed numerous instances where Facebook failed to properly disclose election advertisements[324] despite their policies and the ability of campaigns to use manipulative targeting practices to reach voters on the platform.[325] Further, watchdog journalists have uncovered how Facebook deliberately designed its algorithms to optimize for engagement,[326] incentivizing the spread of disinformation and selectively choosing to apply its content moderation policies.[327] Unfortunately, platforms have resisted providing researchers and watchdog journalists greater access to data, likely because of the risk of embarrassment from failure to adhere to their own policies and public statements.[328] Giving researchers and watchdog journalists greater access to data will not only provide a better picture of how disinformation gets spread, targeted, and amplified but also ensure the integrity of our elections.

Invest Greater Resources in Combating Disinformation Targeting Non-English-Speaking Communities

Platforms must invest greater resources in combating election disinformation in non-English-speaking communities. Research has shown that non-English-language disinformation has continued to spread.[329] Further, disparities exist in the level of enforcement between English and non-English disinformation, leaving non-English-speaking communities more vulnerable to disinformation. **Platforms can remedy these disparities in enforcement by investing greater resources to combating non-English disinformation**, including hiring more content moderators to monitor and combat disinformation in languages other than English.

CONCLUSION

For decades, Common Cause Education Fund has worked on public education and systemic reforms to build a better democracy. The harmful impact of election disinformation makes it clear that our core programmatic work is needed now more than ever. We must and will educate and mobilize our communities to curb the harmful, rapid growth of election disinformation. Doing so will help deliver on America's promise of a functioning 21st-century democracy that's open, accessible, responsive, and accountable to the people. We need your support and your activism to fix the problem of election disinformation. Together, we can build a democracy that works for everyone.

APPENDIX I—STATE VOTER INTIMIDATION AND FALSE ELECTION SPEECH LAWS

Under **Arizona** law, it is a criminal misdemeanor to knowingly "make use of force, violence or restraint, or to inflict or threaten infliction…of any injury, damage, harm or loss, or in any manner to practice intimidation upon or against any person, in order to induce or compel such person to vote or refrain from voting for a particular person or measure" or to "impede, prevent or otherwise interfere with the free exercise of the elective franchise of any voter."[330] Other Arizona laws apply specifically to employer intimidation of their employees[331] and voter interference within a 75-foot buffer zone outside of polling places.[332]

California (among other states[333]) invites candidates and committees to sign a voluntary "Code of Fair Campaign Practices" that includes a promise not to "use or permit any dishonest or unethical practice that tends to corrupt or undermine our American system of free elections, or that hampers or prevents the full and free expression of the will of the voters including acts intended to hinder or prevent any eligible person from registering to vote, enrolling to vote, or voting."[334]

In **Colorado**, it is illegal to "impede, prevent, or otherwise interfere with the free exercise of the elective franchise of any elector."[335] Colorado law also explicitly provides that **no person shall knowingly or recklessly "make, publish, broadcast, or circulate or cause to be made, published, broadcasted, or circulated…any false statement designed to affect the vote** on any issue submitted to the electors at any election or relating to any candidate for election to public office."[336] Colorado attorney general guidance makes clear that disinformation tactics—including "misleading phone calls, texts, or emails to a voter"—can constitute illegal voter intimidation. Examples of illegal voter intimidation include "texting voters deliberately false information about voting locations" and "calling voters to tell them that they must have an identification card or be vaccinated in order to vote."[337]

Under **Florida's** "Voter Protection Act," it's a felony to "directly or indirectly use or threaten to use force, violence, or intimidation or any tactic of coercion or intimidation to induce or compel an individual" to register or vote or refrain from doing so.[338] Importantly, it's a **felony under Florida law to "knowingly use false information"** to "challenge an individual's right to vote" or "induce or attempt to induce an individual to refrain from voting or registering to vote."[339]

Georgia law makes it a felony to use or threaten "violence in a manner that would prevent a reasonable elector from voting or actually prevents any elector from voting."[340] It is likewise a felony in Georgia to use or threaten "violence, or act[] in any other manner to intimidate" another person to vote or refrain from voting or registering to vote.[341]

Hawaii law provides that any person who "knowingly broadcasts, televises, circulates, publishes, distributes, or otherwise communicates, including by electronic means or advertisement, false information about the time, date, place, or means of voting with the purpose of impeding, preventing, or otherwise interfering with the free exercise of the elective franchise" has committed illegal election fraud.[342] It is also a crime in Hawaii to "in any way practice[] intimidation upon or against any person in order to induce or compel the person to vote or refrain from voting" or to

impede, prevent, or otherwise interfere with voting.[343] And earlier this year, Hawaii extended its "no campaigning" zone to protect from harassment voters waiting in lines extending far outside voting centers.[344]

Maine law makes it a crime to "interfere[] with a voter attempting to cast a vote"[345] or to "knowingly cause[] a delay in the registration…of another or…in the delivery of an absentee ballot or absentee ballot application with the intent to prevent a person from voting or to render that person's vote ineffective."[346]

Maryland law makes it a crime to knowingly and willfully "influence or attempt to influence a voter's voting decision through the use of force, threat, menace, [or] intimidation" or to "influence or attempt to influence a voter's decision whether to go to the polls to cast a vote" through use of force, fraud, or threat.[347] It is also a crime to "engage in conduct that results or has the intent to result in the denial or abridgement of the right of any citizen of the United States to vote on account of race, color, or disability."[348]

Michigan law makes it a felony to "attempt, by means of bribery, menace, or **other corrupt means** or device, either directly or indirectly, to influence an elector in giving his or her vote, or to deter the elector from, or interrupt the elector in giving his or her vote at any election held in this state."[349] It is also **illegal in Michigan to knowingly disseminate an "assertion, representation, or statement of fact concerning a candidate…that is false, deceptive, scurrilous, or malicious**, without the true name of the author being subscribed" to the statement.[350]

Minnesota law prohibits directly or indirectly using or threatening "force, coercion, violence, restraint, damage, harm, loss, including loss of employment or economic reprisal, undue influence, or temporal or spiritual injury against an individual to compel the individual to vote for or against a candidate or ballot question."[351] Further, "fraud may not be used to obstruct or prevent the free exercise of the right to vote."[352]

Under **Nevada** law, it is illegal to impede or prevent by "fraudulent contrivance, the free exercise of the franchise by any voter."[353] It is likewise a felony to use or threaten to use any force, intimidation, coercion or undue influence, or to "inflict any mental injury, damage, harm or loss upon" a person on connection with registering or voting in an election.[354]

New Mexico law makes it a felony to use or threaten "force, violence, infliction of damage, harm or loss or any form of economic retaliation, upon any voter…for the purpose of impeding or preventing the free exercise of the elective franchise."[355] Guidance from New Mexico's secretary of state explains that "disseminating false or misleading election information" is a form of voter intimidation, stating, "It is unlawful to disseminate misleading information about elections, including flyers or other communication that purposely misstate the time and date of an election, where it will be held, and how voting will happen."[356]

North Carolina law makes it illegal for any person to "intimidate or oppose any legally qualified voter on account of any vote such voter may cast or consider or intend to cast, or not to cast, or which that voter may have failed to cast."[357]

Pennsylvania law prohibits using "coercion, threats of bodily injury or intimidation" to intentionally prevent or attempt to prevent someone from registering to vote.[358] The Pennsylvania Department of State elaborates on guidance that it is likewise illegal to use such intimidation to compel or prevent someone from voting and that "[d]isseminating false or misleading election information, including information on voting eligibility, polling place procedures, polling place hours, or voting methods" is a form of illegal voter intimidation.[359]

Virginia law makes it a crime to "hinder, intimidate, or interfere with any qualified voter so as to prevent the voter from casting a secret ballot"[360] or to interfere or attempt to interfere with a person registering to vote.[361] Virginia explicitly outlaws communicating to a "registered voter, by any means, false information, knowing the same to be false, intended to impede the voter in the exercise of his right to vote," including information "about the date, time, and place of the election, or the voter's precinct, polling place, or voter registration status, or the location of a voter satellite office or the office of the general registrar."[362] Virginia law includes a private right of action for registered voters to whom such false information is communicated, enabling them to seek an "injunction, restraining order, or other order, against the person communicating such false information."[363]

Wisconsin law prohibits the use or threat of force, violence, duress, or any fraudulent device or contrivance to "impede or prevent the free exercise of the franchise at an election."[364] Wisconsin law also provides that no person "may knowingly make or publish, or cause to be made or published, a false representation pertaining to a candidate or referendum which is intended or tends to affect voting at an election."[365]

APPENDIX II—STATE CAMPAIGN FINANCE DISCLOSURE LAWS

Alaska has enacted one of the nation's most effective laws for tracing the source of funds spent on election advertising by groups, including those that do not qualify as political committees,[366] requiring such groups to disclose the identity of any contributor who has given the group more than $250 in the aggregate during the calendar year "for the purpose of influencing the outcome of an election," as well as all election-related contributions and expenditures made by such groups, including contributions to other such groups.[367] The purpose of this statute is to reveal contributors whose funds are transferred through multiple organizations before being spent on election advertising—contributors who would evade disclosure under most jurisdictions' laws.

California has led the way in recent years in strengthening campaign finance disclosure laws applicable to common sources and types of election disinformation. In 2014, the state strengthened disclosure laws applicable to nonpolitical committee groups spending money to influence California elections (e.g., 501(c)(4) social welfare organizations, 501(c)(5) labor organizations, 501(c)(6) trade associations).[368] Under the 2014 reform, so-called multipurpose organizations that spend more than $50,000 during a 12-month period or more than $100,000 in a period of four consecutive years to influence California elections must register with the state and disclose the donors whose funds were used for the California political spending. And in 2018, California took another step by enacting a "Social Media DISCLOSURE Act,"[369] strengthening disclosure requirements by requiring "paid for by" disclaimers on a broad array of political advertising disseminated via social media platforms.

Maryland has enacted legislation requiring certain tax-exempt organizations that are the source of undisclosed "dark money" political spending in many jurisdictions—501(c)(4), 501(c)(6), and 527 organizations—that make aggregate election-related disbursements of $10,000 or more in an election cycle, to file a report disclosing the identity of each person that made cumulative donations of $10,000 or more to the organization during the period covered by the report.[370]

Minnesota law requires likewise targets' would-be "dark money" with specific donor disclosure requirements for "associations" that contribute more than $5,000 in a calendar year to independent expenditure or ballot measure committees. Such associations must disclose the "name, address, and amount attributable to each person that paid the association dues or fees, or made donations to the association that, in total, aggregate more than $5,000 of the contribution from the association to the independent expenditure or ballot question political committee or fund."[371]

New York has enacted legislation imposing disclosure requirements on paid internet and digital political ads and requiring the state board of elections to maintain a publicly available database of such ads.[372] New York explicitly permits modified "paid for by" disclaimers for digital political advertising, so long as the ad "contains a link to another webpage where the "paid for by" statement is prominently displayed."[373]

Rhode Island law stems the flow of "dark money" with a disclosure requirement not only for independent political expenditures but also for certain transfers of funds between organizations

(defined in the statute as a "covered transfer") exceeding $1,000 in a calendar year for such political spending.[374]

Finally, the state of Washington has some of the strongest disclosure laws in the nation applicable to election disinformation and other digital political advertising. Washington's overall campaign finance disclosure regime is appropriately broad in its application and mandates on-ad identification of top donors to the sponsor for certain political advertising.[375] Washington Public Disclosure Commission regulations provide for modified "paid for by" disclaimers on certain digital ads[376] and require online platforms that sell paid political advertising (digital "commercial advertisers") to provide the public with access to detailed digital ad information.[377]

APPENDIX III—STATE MEDIA LITERACY LAWS

California's media literacy law, passed in 2018 with bipartisan support in the assembly, requires the Department of Education to list instructional materials and resources on how to evaluate trustworthy media sources.[378]

In 2019, the **Colorado** General Assembly passed legislation creating a media literacy advisory committee within the Colorado Department of Education. The committee submitted a report to the Colorado General Assembly in December of the same year and recommended revising Colorado academic standards, providing materials and resources to teachers, and enacting further legislation directing Colorado to take action to support effective implementation of media literacy programs in schools throughout the state.[379]

Florida's law, which was passed in 2008 and strengthened in 2013, requires media literacy to be integrated into the standards for all subjects in K–12 public schools.[380]

Illinois passed a media literacy law in 2021 requiring every public high school in the state to include in its curriculum a unit of instruction on media literacy, **making it the first state to mandate media literacy classes.**[381]

A **Minnesota** state law passed in 2006 requires the state's education commissioner to embed technology and information literacy standards into the state's academic standards and graduation requirements.[382]

New Mexico has had a law in place since 2009 allowing for media literacy to be offered as an elective in schools.[383]

ENDNOTES

1 Aberjhani and Luther E. Vann, *Elemental: The Power of Illuminated Love* (Columbia, SC: Soar Publishing, 2008).

2 Chris Cillizza, "The Big Lie Is (Unfortunately) Winning," *Washington Post*, September 15, 2021, https://www.cnn.com/2021/09/15/politics/big-lie-republican-belief-trump/index.html.

3 Cillizza, "The Big Lie Is (Unfortunately) Winning," *Washington Post*, September 15, 2021, https://www.cnn.com/2021/09/15/politics/big-lie-republican-belief-trump/index.html.

4 "The national, nonpartisan Election Protection coalition works year-round to ensure that all voters have an equal opportunity to vote and have that vote count. Made up of more than 100 local, state and national partners, Election Protection uses a wide range of tools and activities to protect, advance and defend the right to vote." Election Protection Coalition, "About," https://866ourvote.org/about/.

5 Common Cause, the Lawyers' Committee for Civil Rights Under Law, and the Century Foundation, *Deceptive Practices 2.0: Legal and Policy Responses* (Washington, DC: Self-Published, 2008), https://www.commoncause.org/wp-content/uploads/2018/03/0064.pdf.

6 Common Cause and Lawyers Committee for Civil Rights Under Law, *Deceptive Election Practices and Voter Intimidation: The Need for Voter Protection* (Washington, DC: Self-Published, 2012), https://lawyerscommittee.org/wp-content/uploads/2015/07/DeceptivePracticesReportJuly2012FINALpdf.pdf.

7 Davey Alba, "How Voting by Mail Tops Election Misinformation," *New York Times*, September 30, 2020, https://www.nytimes.com/2020/09/30/technology/how-voting-by-mail-tops-election-misinformation.html.

8 Jon Lloyd et al., "Misinformation in the 2020 US Elections: A Timeline of Platform Changes," Mozilla Foundation, March 8, 2021, https://foundation.mozilla.org/en/blog/misinformation-in-the-2020-us-elections-a-timeline-of-platform-changes/.

9 Jon Lloyd et al., "Misinformation in the 2020 US Elections: A Timeline of Platform Changes," Mozilla Foundation, March 8, 2021, https://foundation.mozilla.org/en/blog/misinformation-in-the-2020-us-elections-a-timeline-of-platform-changes/.

10 *See*, e.g., Matt Viser, "Inside the 'Malarkey Factory,' Biden's Online War Room," *Washington Post*, October 19, 2020, https://www.washingtonpost.com/politics/biden-trump-campaign-disinformation/2020/10/18/99774228-0fdd-11eb-8074-0e943a91bf08_story.html.

11 Common Cause, the Lawyers' Committee for Civil Rights Under Law, and the Century Foundation, *Deceptive Practices 2.0: Legal and Policy Responses*, (Washington, DC: Self-Published, 2008), https://www.commoncause.org/wp-content/uploads/2018/03/0064.pdf; Common Cause and Lawyers Committee for Civil Rights Under Law, *Deceptive Election Practices and Voter Intimidation: The Need for Voter Protection* (Washington, DC: Self-Published, 2012), https://lawyerscommittee.org/wp-content/uploads/2015/07/DeceptivePracticesReportJuly2012FINALpdf.pdf; Liz Kennedy, Stephen Spaulding, Tova Wang, Jenny Flanagan and Anthony Kammer, *Bullies at the Ballot Box: Protecting the Freedom to Vote Against Wrongful Challenges and Intimidation* (Washington, DC: Common Cause and Demos, 2012), https://www.commoncause.org/wp-content/uploads/2018/03/BulliesAtTheBallotBox-Final.pdf.

12 Mason Walker and Katerina Eva Matsa, "News Consumption Across Social Media in 2021," Pew Research Center, September 20, 2021, https://www.pewresearch.org/journalism/2021/09/20/news-consumption-across-social-media-in-2021/.

13 Claire Wardle, "Understanding Information Disorder," First Draft, September 22, 2020, https://firstdraftnews.org/long-form-article/understanding-information-disorder/.

14 *See*, e.g., Stephen Collinson, "Trump's Big Lie Is Changing the Face of American Politics," CNN, September 16, 2021, https://www.cnn.com/2021/09/16/politics/trump-big-lie-gop-election/index.html.

15 Tom Kertscher, "Postal Service Says Its Policy Is to Deliver Even Mail Ballots Lacking Postage," *PolitiFact*, July 23, 2020, https://www.politifact.com/factchecks/2020/jul/23/facebook-posts/postal-service-says-its-policy-deliver-even-mail-b/.

16 S. 2747, Freedom to Vote Act, 117th Cong. (2021), Sec. 1304, "Carriage of Election Mail," https://www.congress.gov/bill/117th-congress/senate-bill/2747/text.

17 Storm Gifford, "Jersey Postal Worker Mails It in Before Quitting, Dumps Undelivered Post on Street," *N.Y. Daily News*, October 4, 2018, https://www.nydailynews.com/news/national/ny-news-postal-carrier-quits-job-on-spot-20181004-story.html.

18 Craig Timberg and Beth Reinhard, "As Recount Politics Heat Up, Two Florida Election Officials Are the Targets of Online Harassment," *Washington Post*, November 13, 2018, https://www.washingtonpost.com/technology/2018/11/13/recount-politics-heat-up-two-florida-election-officials-are-targets-online-harassment/.

19 Linda So and Jason Szep, "U.S. Election Workers Get Little Help from Law Enforcement as Terror Threats Mount," *Reuters*, September 8, 2021, https://www.reuters.com/investigates/special-report/usa-election-threats-law-enforcement/.

20 Linda So, "Trump-Inspired Death Threats Are Terrorizing Election Workers," *Reuters*, June 11, 2021, https://www.reuters.com/investigates/special-report/usa-trump-georgia-threats/.

21 Johnny Kauffman, "'You Better Run': After Trump's False Attacks, Election Workers Faced Threats," *NPR*, February 5, 2021, https://www.npr.org/2021/02/05/963828783/you-better-run-after-trumps-false-attacks-election-workers-faced-threats.

22 Matthew Brown, "Fact Check: Georgia 'Suitcase' Video Is Missing Context," *USA Today*, December 14, 2020, https://www.usatoday.com/story/news/factcheck/2020/12/14/fact-check-georgia-suitcase-video-missing-context/3892640001/.

23 Linda So, "Trump-Inspired Death Threats Are Terrorizing Election Workers," *Reuters*, June 11, 2021, https://www.reuters.com/investigates/special-report/usa-trump-georgia-threats/.

24 "Election Officials Under Attack," Brennan Center for Justice and the Bipartisan Policy Center, June 16, 2021, https://www.brennancenter.org/our-work/policy-solutions/election-officials-under-attack.

25 Salvador Rizzo, "Trump's Fusillade of Falsehoods on Mail Voting," *Washington Post*, September 11, 2020, https://www.washingtonpost.com/politics/2020/09/11/trumps-fusillade-falsehoods-mail-voting/.

26 Reid J. Epstein and Stephanie Saul, "Trump Says Mail Voting Means Republicans Would Lose Every Election. Is That True? No.," *Chicago Tribune*, April 10, 2020, https://www.chicagotribune.com/nation-world/ct-nw-nyt-mail-voting-ballots-20200410-qfnxhakicve3ndpxz64lcsqzr4-story.html.

27 Davey Alba, "How Voting by Mail Tops Election Misinformation," *New York Times*, September 30, 2020, https://www.nytimes.com/2020/09/30/technology/how-voting-by-mail-tops-election-misinformation.html.

28 Justin Baragona, "Fox Host Tomi Lahren Claims 'Only Thing' That Will Save Gavin Newsom Is 'Voter Fraud,'" *Daily Beast*, September 8, 2021, https://www.thedailybeast.com/fox-host-tomi-lahren-claims-only-thing-that-will-save-gavin-newsom-is-voter-fraud.

29 Lara Korte, "Larry Elder Prepares for California Recall Loss with Lawyers, Voter Fraud Website," *Sacramento Bee*, September 10, 2021. https://www.sacbee.com/news/politics-government/capitol-alert/article254078633.html.

30 Ben Tobin and Billy Kobin, "'Absurd' and 'Ridiculous': What Officials, Experts Say about Bevin's Voter Fraud Claims," *Courier Journal*, November 7, 2019, https://www.courier-journal.com/story/news/politics/elections/kentucky/2019/11/07/kentucky-governor-election-fact-check-matt-bevins-voter-fraud-claims/2516391001/.

31 Lis Power, "In 2 Weeks after It Called the Election, Fox News Cast Doubt on the Results Nearly 800 Times," Media Matters for America, January 14, 2021, https://www.mediamatters.org/fox-news/2-weeks-after-it-called-election-fox-news-cast-doubt-results-nearly-800-times.

32 Atlantic Council's DFRLab, "#StopTheSteal: Timeline of Social Media and Extremist Activities Leading to 1/6 Insurrection," *Just Security*, February 10, 2021, https://www.justsecurity.org/74622/stopthesteal-timeline-of-social-media-and-extremist-activities-leading-to-1-6-insurrection/.

33 Isaac Stanley-Becker and Anu Narayanswamy, "Trump Has More than $100 Million in Political Cash after First Six Months of 2021," *Washington Post*, August 1, 2021, https://www.washingtonpost.com/politics/2021/07/31/trump-committees-fundraising-2021-fec/.

34 Michael Scherer and Josh Dawsey, "Trump, Talked Out of Announcing a 2024 Bid for Now, Settles on a Wink-and-Nod Unofficial Candidacy," *Washington Post*, October 4, 2021, https://www.washingtonpost.com/politics/trump-2024-campaign-candidacy/2021/10/03/73af3b12-21f8-11ec-b3d6-8cdebe60d3e2_story.html.

35 Scott Pelley, "Whistleblower: Facebook Is Misleading the Public on Progress Against Hate Speech, Violence, Misinformation," *60 Minutes*, October 4, 2021, https://www.cbsnews.com/news/facebook-whistleblower-frances-haugen-misinformation-public-60-minutes-2021-10-03/.

36 Common Cause and Lawyers Committee for Civil Rights Under Law, *Deceptive Election Practices and Voter Intimidation: The Need for Voter Protection* (Washington, DC: Self-Published, 2012), https://lawyerscommittee.org/wp-content/uploads/2015/07/DeceptivePracticesReportJuly2012FINALpdf.pdf.

37 "PolitiFact's Guide to Fake News Websites and What They Peddle," PolitiFact, April 20, 2017, https://www.politifact.com/article/2017/apr/20/politifacts-guide-fake-news-websites-and-what-they/

38 *See* Casey McDermott, Twitter post, September 8, 2020, 11:50am, https://twitter.com/caseymcdermott/status/1303359983728418816 ("Exeter and Bedford, both with pretty big voting populations, wrongly say on their websites 'election officials are not authorized to accept absentee ballots at the polls.' State officials say they're addressing the error. Again, you can still vote absentee at the polls today.").

39 Danny Sullivan, "How We Keep Search Relevant and Useful," *The Keyword* (blog), July 5, 2019,

40 https://blog.google/products/search/how-we-keep-google-search-relevant-and-useful/.

41 Rand Fishkin, "The State of Searcher Behavior Revealed Through 23 Remarkable Statistics," *Moz* (blog), March 14, 2017, https://moz.com/blog/state-of-searcher-behavior-revealed.

42 Online voting presents enormous risks to voter privacy, ballot secrecy, integrity of election results, and, consequently, national security. The option to vote online is not available to the general public. Thirty-three states do offer some form of online voting to overseas and military voters, and a few are now allowing it for people with disabilities. *See*, e.g., Kaleigh Rogers, "New Laws Let Americans With Disabilities Vote Online. They've Also Resurrected the Debate About Voting Access vs. Election Security," *FiveThirtyEight*, July 7, 2021, https://fivethirtyeight.com/features/new-laws-let-americans-with-disabilities-vote-online-theyve-also-resurrected-the-debate-about-voting-access-vs-election-security/; *see also* Eric Geller, "Some States Have Embraced Online Voting. It's a Huge Risk," *Politico*, June 8, 2020, https://www.politico.com/news/2020/06/08/online-voting-304013.

43 Leanna Garfield, "Don't Fall for These Online Voting Scams Circulating the Internet," *Business Insider*, November 8, 2016, https://www.businessinsider.com/online-text-voting-scams-hillary-trump-election-2016-11.

44 Alexa Corse and Dustin Volz, "No, You Can't Vote Via Text or Tweet," *Wall Street Journal*, August 11, 2018, https://www.wsj.com/articles/no-you-cant-vote-via-text-or-tweet-1533985201.

45 Danny Sullivan, "How We Keep Search Relevant and Useful," *The Keyword* (blog), July 15, 2019, https://blog.google/products/search/how-we-keep-google-search-relevant-and-useful/.

46 "How Google autocomplete predictions work," Google Help Center, https://support.google.com/websearch/answer/7368877?hl=en.

47 "How Google Autocomplete Predictions Work," Google Help Center, https://support.google.com/websearch/answer/7368877?hl=en.

48 Amber Phillips, "What Is Ballot 'Harvesting,' and Why Is Trump so against It?," *Washington Post*, May 26, 2020,

49 https://www.washingtonpost.com/politics/2020/05/26/what-is-ballot-harvesting-why-is-trump-so-against-it/.

50 Lily Ray, "2020 Google Search Survey: How Much Do Users Trust Their Search Results?," *Moz* (blog), March 2, 2020, https://moz.com/blog/2020-google-search-survey.

51 Robert Epstein and Ronald E. Robertson, "The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections," Proceedings of the National Academy of Sciences of the United States of America (PNAS), August 4, 2015, https://www.pnas.org/content/pnas/112/33/E4512.full.pdf.

52 "Search Engine Market Share United States Of America," Statcounter GlobalStats, https://gs.statcounter.com/search-engine-market-share/all/united-states-of-america.

53 "Google Pushing Scam Ads on Americans Searching for How to Vote," Tech Transparency Project, June 29, 2020, https://www.techtransparencyproject.org/articles/google-pushing-scam-ads-americans-searching-how-vote.

54 "Google Fails to Stop Exploitative Ads Targeting American Voters," Tech Transparency Project, October 5, 2020, https://www.techtransparencyproject.org/articles/google-fails-stop-exploitative-ads-targeting-american-voters.

55 Ellen Nakashima, Amy Gardner, Isaac Stanley-Becker, and Craig Timberg, "U.S. Government Concludes Iran Was behind Threatening Emails Sent to Democrats," *Washington Post*, October 22, 2020, https://www.washingtonpost.com/technology/2020/10/20/proud-boys-emails-florida/.

56 Ellen Nakashima, Amy Gardner, Isaac Stanley-Becker, and Craig Timberg, "U.S. Government Concludes Iran Was behind Threatening Emails Sent to Democrats," *Washington Post*, October 22, 2020, https://www.washingtonpost.com/technology/2020/10/20/proud-boys-emails-florida/.

57 Sec. Jocelyn Benson, Twitter Post, August 27, 2020, 12:12 p.m., https://twitter.com/JocelynBenson/status/1299017044554326019?s=20.

58 *Example of False Information Being Used to Suppress Voting in Detroit*, YouTube video, 0:37, posted by Michigan Department of State/Secretary of State, August 27, 2020, https://youtu.be/JVodAUh9kL0.

59 Will Sommer, "Jacob Wohl Accused of Starting a Voter Suppression Scheme," *Daily Beast*, August 27, 2020, https://www.thedailybeast.com/jacob-wohl-accused-of-starting-a-voter-suppression-scheme?ref=scroll.

60 Meryl Kornfield, "Robocall Targets Battleground States with Falsehoods about Mail-in Voting," *Washington Post*, August 27, 2020, https://www.washingtonpost.com/politics/2020/08/27/robocalls-michigan-penn-voting-jacob-wohl/.

61 Stephanie Saul, "Deceptive Robocalls Try to Frighten Detroit residents about Voting by Mail," *New York Times*, August 27, 2020, https://www.nytimes.com/2020/08/27/us/elections/deceptive-robocalls-try-to-frighten-detroit-residents-about-voting-by-mail.html.

62 Federal Communications Commission, "FCC Proposes $5 Million Robocalling Fine Against Jacob Wohl and John Burkman," news release, August 24, 2021, https://docs.fcc.gov/public/attachments/DOC-375180A1.pdf.

63 William Davies, "What's Wrong with WhatsApp," *The Guardian*, July 2, 2020, https://www.theguardian.com/technology/2020/jul/02/whatsapp-groups-conspiracy-theories-disinformation-democracy.

64 Brooke Auxier and Monica Anderson, "Social Media Use in 2021," Pew Research Center, April 7, 2021, https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/.

65 Elisa Shearer and Amy Mitchell, "News Use Across Social Media Platforms in 2020," Pew Research Center, January 12, 2021, https://www.pewresearch.org/journalism/2021/01/12/news-use-across-social-media-platforms-in-2020/.

66 Sheera Frenkel, "How Misinformation 'Superspreaders' Seed False Election Theories," *New York Times*, November 23, 2020, https://www.nytimes.com/2020/11/23/technology/election-misinformation-facebook-twitter.html.

67 "VoterFraud2020," Jacobs Technion-Cornell Institute at Cornell Tech, https://voterfraud2020.io/.

68 Salvador Rodriguez, "Mark Zuckerberg Shifted Facebook's Focus to Groups after the 2016 Election, and It's Changed How People Use the Site," *CNBC*, February 16, 2020. https://www.cnbc.com/2020/02/16/zuckerbergs-focus-on-facebook-groups-increases-facebook-engagement.html.

69 Jordan Davis, "Making Groups Privacy Settings Easier to Understand," Facebook, August 14, 2019, https://about.fb.com/

news/2019/08/groups-privacy-settings/.

70 Dave Johnson, "What Is Telegram? A Quick Guide to the Fast and Secure Messaging Platform," *Business Insider*, March 24, 2021, https://www.businessinsider.com/what-is-telegram.

71 Michael Schwirtz, "Telegram, Pro-Democracy Tool, Struggles Over New Fans From Far Right," *New York Times*, January 26, 2021, https://www.nytimes.com/2021/01/26/world/europe/telegram-app-far-right.html.

72 Ashley Carman, "Peloton Is Blocking the #StopTheSteal Hashtag from Being Created or Used," *The Verge*, January 11, 2021, https://www.theverge.com/2021/1/11/22225197/peloton-stop-the-steal-tag-banned-donald-trump.

73 Center for an Informed Public, Digital Forensic Research Lab, Graphika, & Stanford Internet Observatory, "The Long Fuse: Misinformation and the 2020 Election," Stanford Digital Repository: Election Integrity Partnership v1.3.0, 2021, https://purl.stanford.edu/tr171zs0069.

74 "A Look at Facebook and US 2020 Elections," Facebook, December 2020, https://about.fb.com/wp-content/uploads/2020/12/US-2020-Elections-Report.pdf.

75 Galen Stocking et al., "YouTube News Consumers about as Likely to Use the Site for Opinions as for Facts," Pew Research Center, September 28, 2020, https://www.pewresearch.org/journalism/2020/09/28/youtube-news-consumers-about-as-likely-to-use-the-site-for-opinions-as-for-facts/.

76 Casey Newton, "How YouTube Failed the 2020 Election Test," *Platformer*, March 3, 2021, https://www.platformer.news/p/how-youtube-failed-the-2020-election.

77 Sheera Frenkel, "Election Misinformation Often Evaded YouTube's Efforts to Stop It," *New York Times*, November 18, 2020, https://www.nytimes.com/2020/11/18/technology/election-misinformation-often-evaded-youtubes-efforts-to-stop-it.html.

78 Craig Silverman, "This Pro-Trump YouTube Network Sprang Up Just After He Lost," *BuzzFeed News*, January 8, 2021, https://www.buzzfeednews.com/article/craigsilverman/epoch-times-trump-you-tube.

79 "Supporting the 2020 U.S. Election," YouTube Official Blog (blog), December 9, 2020, https://blog.youtube/news-and-events/supporting-the-2020-us-election/.

80 "Supporting the 2020 U.S. Election," YouTube Official Blog (blog), December 9, 2020, https://blog.youtube/news-and-events/supporting-the-2020-us-election/.

81 "YouTube Still Awash in False Voter Fraud Claims," Tech Transparency Project, December 22, 2020, https://www.techtransparencyproject.org/articles/youtube-still-awash-false-voter-fraud-claims.

82 Shannon Bond, "Twitter Says Steps to Curb Election Misinformation Worked," *NPR*, November 12, 2020, https://www.npr.org/sections/live-updates-2020-election-results/2020/11/12/934267731/twitter-says-steps-to-curb-election-misinformation-worked.

83 Makena Kelly, "TikTok Removed More than 300,000 Videos for Election Misinformation," *The Verge*, February 24, 2021, https://www.theverge.com/2021/2/24/22298024/tiktok-election-misinformation-disinformation-transparency-report.

84 Jessica Bursztynsky, "TikTok Says 1 Billion People Use the App Each Month," *CNBC*, September 27, 2021, https://www.cnbc.com/2021/09/27/tiktok-reaches-1-billion-monthly-users.html; Kari Paul, "TikTok: False Posts about US Election Reach Hundreds of Thousands," *The Guardian*, November 5, 2020, https://www.theguardian.com/technology/2020/nov/05/tiktok-us-election-misinformation.

85 Ellie House, "Rumble Sends Viewers Tumbling Toward Misinformation," *Wired*, May 11, 2021, https://www.wired.com/story/rumble-sends-viewers-tumbling-toward-misinformation/.

86 United States Senate Select Committee on Intelligence, "Report on Russian Active Measures Campaigns and Interference in the 2016 U.S. Election," Vol. 2, 116th Cong. (2019), https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf.

87 United States Senate Select Committee on Intelligence, "Report on Russian Active Measures Campaigns and Interference in the 2016 U.S. Election," Vol. 2, 116th Cong. (2019), https://www.intelligence.senate.gov/sites/default/files/documents/Report_Volume2.pdf.

88 Tim Mak, "Senate Report: Russians Used Social Media Mostly to Target Race In 2016," *NPR*, October 8, 2019, https://www.npr.org/2019/10/08/768319934/senate-report-russians-used-used-social-media-mostly-to-target-race-in-2016.

89 Scott Shane, "These Are the Ads Russia Bought on Facebook in 2016," *New York Times*, November 1, 2017, https://www.nytimes.com/2017/11/01/us/politics/russia-2016-election-facebook.html.

90 Ellen Nakashima et al., "U.S. Government Concludes Iran Was Behind Threatening Emails Sent to Democrats," *Washington Post*, October 22, 2020, https://www.washingtonpost.com/technology/2020/10/20/proud-boys-emails-florida/.

91 Jen Kirby, "Yes, Russia Is Interfering in the 2020 Election," Vox, September 21, 2020, https://www.vox.com/2020/9/21/21401149/russia-2020-election-meddling-trump-biden.

92 "Intelligence Community Assessment: Foreign Threats to the 2020 US Federal Elections (Declassified)," National Intelligence Council, March 10, 2021, https://www.dni.gov/files/ODNI/documents/assessments/ICA-declass-16MAR21.pdf.

93 Carrie Levine, "Online Misinformation during the Primaries: A Preview of What's to Come?," Center for Public Integrity, March 11, 2020, https://publicintegrity.org/politics/elections/online-misinformation-during-the-primaries-a-preview-of-whats-

to-come/.

94 Jennifer Agiesta and Ariel Edwards-Levy, "CNN Poll: Most Americans Feel Democracy Is under Attack in the US," *CNN*, September 15, 2021, https://www.cnn.com/2021/09/15/politics/cnn-poll-most-americans-democracy-under-attack/index.html.

95 Allie Morris, "Hours after Trump Calls for Audit of 2020 Texas Election, State Says It's Auditing 4 Urban Counties," *Dallas Morning News*, September 23, 2021, https://www.dallasnews.com/news/politics/2021/09/23/trump-to-abbott-add-2020-election-audit-bill-to-texas-legislatures-special-session/.

96 Tim Reid and Nathan Layne, Jason Lange, "Special Report: Backers of Trump's False Fraud Claims Seek to Control next Elections," *Reuters*, September 22, 2021, https://www.reuters.com/world/us/backers-trumps-false-fraud-claims-seek-control-next-us-elections-2021-09-22/.

97 Shannon Bond, "Just 12 People Are Behind Most Vaccine Hoaxes On Social Media, Research Shows," *NPR*, May 14, 2021, https://www.npr.org/2021/05/13/996570855/disinformation-dozen-test-facebooks-twitters-ability-to-curb-vaccine-hoaxes.

98 Nicole Hong, "Twitter Troll Tricked 4,900 Democrats in Vote-by-Phone Scheme, U.S. Says," *New York Times*, January 27, 2021, https://www.nytimes.com/2021/01/27/nyregion/douglass-mackey-arrested-far-right-twitter.html.

99 *U.S. v. Douglas Mackey*, Complaint 1 (E.D.N.Y. January 22, 2021), https://www.justice.gov/opa/press-release/file/1360816/download.

100 *U.S. v. Douglas Mackey*, Complaint 21 (E.D.N.Y. January 22, 2021), https://www.justice.gov/opa/press-release/file/1360816/download.

101 Center for an Informed Public, Digital Forensic Research Lab, Graphika, & Stanford Internet Observatory, "The Long Fuse: Misinformation and the 2020 Election," Stanford Digital Repository: Election Integrity Partnership v1.3.0 (2021), https://purl.stanford.edu/tr171zs0069.

102 Center for an Informed Public, Digital Forensic Research Lab, Graphika, & Stanford Internet Observatory, "The Long Fuse: Misinformation and the 2020 Election," Stanford Digital Repository: Election Integrity Partnership v1.3.0 (2021), https://purl.stanford.edu/tr171zs0069.

103 "Facebook: From Election to Insurrection," Avaaz, March 18, 2021, https://secure.avaaz.org/campaign/en/facebook_election_insurrection/.

104 Jane Mayer, "The Big Money Behind the Big Lie," *New Yorker*, August 9, 2021, https://www.newyorker.com/magazine/2021/08/09/the-big-money-behind-the-big-lie.

105 Jane Mayer, "The Big Money Behind the Big Lie," *New Yorker*, August 9, 2021, https://www.newyorker.com/magazine/2021/08/09/the-big-money-behind-the-big-lie.

106 Jane Mayer, "The Big Money Behind the Big Lie," *New Yorker*, August 9, 2021, https://www.newyorker.com/magazine/2021/08/09/the-big-money-behind-the-big-lie.

107 U.S. Rep. Zoe Lofgren (D-CA) has published a 1,939-page catalog of disinformation and misinformation social media posts by Republican members of the House of Representatives who voted to overturn the 2020 presidential election in service of former president Trump's "Big Lie." *See* Rep. Zoe Lofgren, "Social Media Review: Members of the U.S. House of Representatives Who Voted to Overturn the 2020 Presidential Election," https://housedocs.house.gov/lofgren/SocialMediaReview8.pdf.

108 Isaac Stanley-Becker and Anu Narayanswamy, "Trump Has More than $100 Million in Political Cash after First Six Months of 2021," *Washington Post*, August 1, 2021, https://www.washingtonpost.com/politics/2021/07/31/trump-committees-fundraising-2021-fec/.

109 Luke Broadwater, Catie Edmondson and Rachel Shorey, "Fund-Raising Surged for Republicans Who Sought to Overturn the Election," *New York Times*, April 17, 2021, https://www.nytimes.com/2021/04/17/us/politics/republicans-fund-raising-capitol-riot.html.

110 Luke Broadwater, Catie Edmondson and Rachel Shorey, "Fund-Raising Surged for Republicans Who Sought to Overturn the Election," *New York Times*, April 17, 2021, https://www.nytimes.com/2021/04/17/us/politics/republicans-fund-raising-capitol-riot.html.

111 Nicholas Reimann, "Arizona Audit Cost Trump Supporters Nearly $6 Million—Only to Assert Biden Won by Even More," *Forbes*, September 24, 2021, https://www.forbes.com/sites/nicholasreimann/2021/09/24/arizona-audit-cost-trump-supporters-nearly-6-million-only-to-assert-biden-won-by-even-more/?sh=72519b1e2410.

112 "What Happened in Arizona Did Not Stay in Arizona," States United Action, Fair Fight Action, United to Protect Democracy, https://notanaudit.com/.

113 Rosalind S. Helderman, "Arizona Ballot Review Commissioned by Republicans Reaffirms Biden's Victory," *Washington Post*, September 24, 2021, https://www.washingtonpost.com/politics/arizona-ballot-review-draft-report/2021/09/24/7c19ac08-1562-11ec-b976-f4a43b740aeb_story.html.

114 Jane Mayer, "The Big Money Behind the Big Lie," *The New Yorker*, August 9, 2021, https://www.newyorker.com/magazine/2021/08/09/the-big-money-behind-the-big-lie.

115 Jane Mayer, "The Big Money Behind the Big Lie," *The New Yorker*, August 9, 2021, https://www.newyorker.com/magazine/2021/08/09/the-big-money-behind-the-big-lie.

116 Telegram, "Terms of Service," accessed October 19, 2021, https://telegram.org/tos.

117 Tom Kertscher, "No Truth to the Claim That Arizona Audit Found Trump up by 250,000 votes," PolitiFact, May 7, 2021, https://www.politifact.com/factchecks/2021/may/07/facebook-posts/no-truth-claim-arizona-audit-found-trump-250000-vo/.

118 Reuters Fact Check, "Fact Check—Maricopa County Database Was not Deleted," *Reuters*, May 21, 2021, https://www.reuters.com/article/factcheck-maricopa-database/fact-check-maricopa-county-database-was-not-deleted-idUSL2N2N8266.

119 Daniel Funke, "Fact Check: Georgia Military, Overseas Ballots not Evidence of Election Fraud," *USA Today*, May 29, 2021, https://www.usatoday.com/story/news/factcheck/2021/05/29/fact-check-georgia-military-overseas-ballots-not-evidence-fraud/7450430002/.

120 Keith Zubrow, "Facebook Whistleblower Says Company Incentivizes 'Angry, Polarizing, Divisive Content,'" *CBS News*, October 4, 2021, https://www.cbsnews.com/news/facebook-whistleblower-frances-haugen-60-minutes-polarizing-divisive-content/.

121 PolitiFact, "How We Determine Truth-O-Meter Ratings," last updated October 27, 2020, https://www.politifact.com/article/2018/feb/12/principles-truth-o-meter-politifacts-methodology-i/.

122 *See 281 Care Committee v. Arneson*, 766 F.3d 774 (8th Cir. 2014) (holding unconstitutional Minn. Stat. § 211B.06, which prohibited dissemination of false paid political advertising about the personal or political character or acts of a candidate, or about the effect of a ballot question); *see also Susan B. Anthony List v. Driehaus*, 814 F.3d 466 (2016) (holding unconstitutional Ohio Rev. Code Ann. §3517.21(B)(9)-(10), which made it illegal to "[m]ake a false statement concerning the voting record of a candidate or public official," or to "[p]ost, publish, circulate, distribute, or otherwise disseminate a false statement concerning a candidate").

123 Richard L. Hasen, "A Constitutional Right to Lie in Campaigns and Elections?," 74 Mont. L. Rev. 53, 71 (2013), *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2151618.

124 Act to Establish the Department of Justice, ch. 150, 16 Stat. 162 (1870).

125 *See* U.S. Senate, "The Enforcement Acts of 1870 and 1871," https://www.senate.gov/artandhistory/history/common/generic/EnforcementActs.htm.

126 U.S. Dept. of Justice, "Federal Prosecution of Election Offenses," 50, 8th ed. December 2017, https://www.justice.gov/criminal/file/1029066/download.

127 52 U.S.C. § 20511(1).

128 18 U.S. Code § 594.

129 U.S. Dept. of Justice, "Federal Prosecution of Election Offenses," 52, 8th ed. December 2017, https://www.justice.gov/criminal/file/1029066/download.

130 18 U.S.C. § 241.

131 U.S. Dept. of Justice, "Federal Prosecution of Election Offenses," 56, 8th ed. December 2017, https://www.justice.gov/criminal/file/1029066/download.

132 *U.S. v. Douglas Mackey*, Complaint 1 (E.D.N.Y. January 22, 2021), https://www.justice.gov/opa/press-release/file/1360816/download.

133 Tasneem Nashrulla and Ryan Mac, "The Racist Guy Behind One of the Most Influential Pro-Trump Twitter Accounts Was Arrested for Election Interference," *BuzzFeed News*, January 27, 2021, https://www.buzzfeednews.com/article/tasneemnashrulla/ricky-vaughn-twitter-troll-arrested-election-interference.

134 *U.S. v. Douglas Mackey*, Complaint 23 (E.D.N.Y. January 22, 2021), https://www.justice.gov/opa/press-release/file/1360816/download.

135 Nicole Hong, "Twitter Troll Tricked 4,900 Democrats in Vote-by-Phone Scheme, U.S. Says," *New York Times*, January 27, 2021, https://www.nytimes.com/2021/01/27/nyregion/douglass-mackey-arrested-far-right-twitter.html.

136 52 U.S.C. § 10307(b); *see also* 52 U.S.C. § 10101(b).

137 *United States v. North Carolina Republican Party*, 5:92-cv-00161 (E.D.N.C. 1992).

138 Michael Weingartner, "Remedying Intimidating Voter Disinformation Through § 1985(3)'s Support or Advocacy Clauses," Georgetown L. Rev. (forthcoming 2021), *available at* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3914719; *citing* 42 U.S.C. § 1985.

139 Colo. Rev. Stat. § 1-13-109.

140 Colorado Attorney General Phil Weiser, "Public Advisory on Voter Intimidation Crimes and Poll Security," October 19, 2020, https://coag.gov/app/uploads/2020/10/Public-Advisory-Voter-Intimidation-10.19.2020.pdf.

141 Haw. Rev. Stat. § 19-3(12).

142 Va. Code Ann. § 24.2-1005.1(A).

143 Va. Code Ann. § 24.2-1005.1(C).

144 L. Brandeis, Other People's Money 62 (National Home Library Foundation ed. 1933).

145 *Buckley v. Valeo*, 424 U.S. 1, 67 (1976).

146 *Buckley*, 424 U.S. at 66-67.

147 *Buckley*, 424 U.S. at 67.

148 *Buckley*, 424 U.S. at 67.

149 *Gaspee Project v. Mederos*, Case No. 20-1944 (1st Cir. September 14, 2021).

150 52 U.S.C. § 30104(a)-(b).

151 52 U.S.C. § 30120.

152 52 U.S.C. § 30101(17).

153 52 U.S.C. § 30104(f).

154 52 U.S.C. § 30120.

155 52 U.S.C. § 30124(a).

156 52 U.S.C. § 30124(b).

157 Policy Statement of Commissioner Lee E. Goodman on the Fraudulent Misrepresentation Doctrine, February 16, 2018, https://www.fec.gov/resources/cms-content/documents/Commissioner_Lee_E._Goodman_Policy_Statement_-_Fraudulent_Misrepresentation.pdf.

158 Policy Statement of Commissioner Lee E. Goodman on the Fraudulent Misrepresentation Doctrine, February 16, 2018, https://www.fec.gov/resources/cms-content/documents/Commissioner_Lee_E._Goodman_Policy_Statement_-_Fraudulent_Misrepresentation.pdf.

159 By contrast, the FEC has on occasion taken enforcement action when the required disclaimer was omitted entirely. For example, in *FEC v. Novacek*, 739 F. Supp. 2d 957 (N.D. Tex. 2010), a federal court at the commission's urging ordered a defendant to pay a $47,414 civil penalty for fraudulently misrepresenting themselves as acting on behalf of a political party when soliciting funds and failing to include in their communications the required disclaimer. A summary of the case and litigation documents can be found on the FEC website here: https://www.fec.gov/legal-resources/court-cases/fec-v-novacek/.

160 Alaska applies disclosure requirements to "nongroup entities," defined to mean "a person, other than an individual, that takes action the major purpose of which is to influence the outcome of an election" and that "cannot participate in business activities," "does not have shareholders who have a claim on corporate earnings," and is "independent from the influence of business corporations." Alaska Stat. § 15.13.400(13).

161 Alaska Stat. § 15.13.040(j).

162 Cal. Gov. Code § 84222.

163 Cal. Assembly Bill 2188, signed September 26, 2018, https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB2188.

164 Md. Code, Elec. Law § 13-309.2.

165 Minn. Stat. Ann. § 10A.27 Subd.15(b).

166 17 R.I. Gen. Laws § 17-25.3-1.

167 Wash. Rev. Code § 42.17A.320.

168 Wash. Admin. Code § 390-18-030.

169 Wash. Admin. Code § 390-18-050.

170 47 U.S.C. § 230.

171 John Bergmayer, "What Section 230 Is and Does—Yet Another Explanation of One of the Internet's Most Important Laws," Public Knowledge, May. 14, 2019, https://www.publicknowledge.org/blog/what-section-230-is-and-does-yet-another-explanation-of-one-of-the-internets-most-important-laws/.

172 47 U.S.C § 230(c)(1) (stating that platforms and users may not "be treated as the publisher or speaker of any information provided by another").

173 47 U.S.C. § 230(c)(2) (stating that platforms may not be held liable for "good faith" attempts "to restrict access to or availability of material that the provider or user considers to be obscene, lewd, lascivious, filthy, excessively violent, harassing, or otherwise objectionable").

174 *See* Jesse Littlewood and Emma Steiner, "Trending in the Wrong Direction: Social Media Platforms' Declining Enforcement of Voting Disinformation," Common Cause, September 2, 2021, https://www.commoncause.org/wp-content/uploads/2021/09/Disinfo_WhitePaperv3.pdf; "Facebook: From Election to Insurrection, Avaaz, March 18, 2021, https://avaazimages.avaaz.org/facebook_election_insurrection.pdf; Rachel Lerman, "Facebook Says It Has Taken Down 7 Million Posts for Spreading Coronavirus Misinformation," *Washington Post*, August 11, 2020, https://www.washingtonpost.com/technology/2020/08/11/facebook-covid-misinformation-takedowns/.

175 "Facebook: From Election to Insurrection, Avaaz, March 18, 2021, https://avaazimages.avaaz.org/facebook_election_insurrection.pdf.

176 Jesse Littlewood and Emma Steiner, "Trending in the Wrong Direction: Social Media Platforms' Declining Enforcement of

Voting Disinformation," Common Cause, September 2, 2021, https://www.commoncause.org/wp-content/uploads/2021/09/Disinfo_WhitePaperv3.pdf.

177 Harold Feld, "What Is Next for Section 230 Reform," *Pro Market*, November 18, 2020, https://promarket.org/2020/11/18/what-next-for-section-230-reform-court-fcc/.

178 Kiran Jeevanjee et al., "All the Ways Congress Wants to Change Section 230," *Slate*, March 23, 2021, https://slate.com/technology/2021/03/section-230-reform-legislative-tracker.html.

179 H.R. 874, Abandoning Online Censorship Act, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/house-bill/874/text?q=%7B%22search%22%3A%5B%22section+230%22%5D%7D&r=1&s=3.

180 S. 1384, 21st Century Foundation for the Right to Express and Engage Speech Act, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/senate-bill/1384/text.

181 Senator Bill Hagerty, "Hagerty Introduces Bill to Combat Big Tech Censorship; Treat Big Tech Corporations as Common Carriers," news release, April 27, 2021, https://www.hagerty.senate.gov/press-releases/2021/04/27/hagerty-introduces-bill-to-combat-big-tech-censorship-treat-big-tech-corporations-as-common-carriers/.

182 S. 797, Platform Accountability and Online Transparency Act, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/senate-bill/797/text.

183 House Committee on Energy and Commerce, "E&C Leaders Announce Legislation to Reform Section 230," news release, October 14, 2021, https://energycommerce.house.gov/newsroom/press-releases/ec-leaders-announce-legislation-to-reform-section-230.

184 Letter from public interest and civil rights groups to White House and 117th Congress, January 27, 2021, https://cdn.vox-cdn.com/uploads/chorus_asset/file/22261662/Section_230_Letter__Jan_27__2021___1_.pdf.

185 Federal Trade Commission," Our History," https://www.ftc.gov/about-ftc/our-history.

186 15 U.S.C. § 45.

187 Federal Trade Commission, "FTC Policy Statement on Unfairness, Federal Trade Commission," December 17, 1980, https://www.ftc.gov/public-statements/1980/12/ftc-policy-statement-unfairness.

188 Federal Trade Commission, "Enforcement Policy Statement on Deceptively Formatted Advertisements," https://www.ftc.gov/system/files/documents/public_statements/896923/151222deceptiveenforcement.pdf.

189 Federal Trade Commission, "Federal Trade Commission 2020 Privacy and Data Security Update," 2020, https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-2020-privacy-data-security-update/20210524_privacy_and_data_security_annual_update.pdf.

190 Federal Trade Commission, "Federal Trade Commission 2020 Privacy and Data Security Update," 2020 https://www.ftc.gov/system/files/documents/reports/federal-trade-commission-2020-privacy-data-security-update/20210524_privacy_and_data_security_annual_update.pdf.

191 Rosalie Chan, "The Cambridge Analytica Whistleblower Explains How the Firm Used Facebook Data to Sway Elections," *Business Insider*, October 5, 2019, https://www.businessinsider.com/cambridge-analytica-whistleblower-christopher-wylie-facebook-data-2019-10.

192 Christopher Wylie, "How I Helped Hack Democracy," *New York Intelligencer*, October 4, 2019, https://nymag.com/intelligencer/2019/10/book-excerpt-mindf-ck-by-christopher-wylie.html.

193 Rosalie Chan, "The Cambridge Analytica Whistleblower Explains How the Firm Used Facebook Data to Sway Elections," *Business Insider*, October 5, 2019, https://www.businessinsider.com/cambridge-analytica-whistleblower-christopher-wylie-facebook-data-2019-10.

194 Federal Trade Commission, "FTC Imposes $5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook," news release, July 24, 2019, https://www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions.

195 Federal Trade Commission, "FTC Imposes $5 Billion Penalty and Sweeping New Privacy Restrictions on Facebook," news release, July 24, 2019, https://www.ftc.gov/news-events/press-releases/2019/07/ftc-imposes-5-billion-penalty-sweeping-new-privacy-restrictions.

196 Federal Trade Commission, "FTC Issues Opinion and Order Against Cambridge Analytica for Deceiving Consumers About the Collection of Facebook Data, Complain with EU-U.S. Privacy Shield," news release, December 6, 2019, https://www.ftc.gov/news-events/press-releases/2019/12/ftc-issues-opinion-order-against-cambridge-analytica-deceiving.

197 Federal Trade Commission, "FTC Issues Opinion and Order Against Cambridge Analytica for Deceiving Consumers About the Collection of Facebook Data, Complain with EU-U.S. Privacy Shield," news release, December 6, 2019, https://www.ftc.gov/news-events/press-releases/2019/12/ftc-issues-opinion-order-against-cambridge-analytica-deceiving.

198 Chris Jay Hoofnagle, Woodrow Hartzog and Daniel J. Solove, "The FTC Can Rise to the Privacy Challenge, but not Without Help from Congress," Brookings Institution, August 8, 2019, https://www.brookings.edu/blog/techtank/2019/08/08/the-ftc-can-rise-to-the-privacy-challenge-but-not-without-help-from-congress/.

199 Federal Trade Commission, "Congressional Budget Justification Fiscal Year 2022," May 28, 2021, https://www.ftc.gov/

system/files/documents/reports/fy-2022-congressional-budget-justification/fy22cbj.pdf.

200 Federal Trade Commission, "FTC Report on Resources Used and Needed for Protecting Consumer Privacy and Security," 2020, https://www.ftc.gov/system/files/documents/reports/reports-response-senate-appropriations-committee-report-116-111-ftcs-use-its-authorities-resources/p065404reportresourcesprivacydatasecurity.pdf.

201 Chris Jay Hoofnagle, Woodrow Hartzog and Daniel J. Solove, "The FTC Can Rise to the Privacy Challenge, but not Without Help from Congress," Brookings Institution, August 8, 2019, https://www.brookings.edu/blog/techtank/2019/08/08/the-ftc-can-rise-to-the-privacy-challenge-but-not-without-help-from-congress/.

202 Nilay Patel, "Facebook's $5 billion FTC Fine Is an Embarrassing Joke," *The Verge*, July 12, 2019, https://www.theverge.com/2019/7/12/20692524/facebook-five-billion-ftc-fine-embarrassing-joke.

203 Rohit Chopra, "Dissenting Statement of Commissioner Rohit Chopra," July 24, 2019, https://www.ftc.gov/system/files/documents/public_statements/1536911/chopra_dissenting_statement_on_facebook_7-24-19.pdf.

204 Joel Breakstone et. al, "Students' Civic Online Reasoning: A National Portrait," Stanford History Education Group, November 14, 2019, https://purl.stanford.edu/gf151tb4868.

205 Michael Copps, "Learning Digital Literacy Is Key," Democracy Wire (blog), March 11, 2021, https://www.commoncause.org/democracy-wire/learning-digital-literacy-is-key/.

206 Paula Span, "Getting Wise to Fake News," *New York Times*, October 14, 2020, https://www.nytimes.com/2020/09/11/health/misinformation-social-media-elderly.html.

207 Paula Span, "Getting Wise to Fake News," *New York Times*, October 14, 2020, https://www.nytimes.com/2020/09/11/health/misinformation-social-media-elderly.html.

208 Michael Copps, "Learning Digital Literacy Is Key," Democracy Wire (blog), March 11, 2021, https://www.commoncause.org/democracy-wire/learning-digital-literacy-is-key/.

209 Sarah Schwartz, "More States Say They're Teaching Media Literacy, but What That Means Varies," *Education Week*, January 08, 2020, https://www.edweek.org/teaching-learning/more-states-say-theyre-teaching-media-literacy-but-what-that-means-varies/2020/01.

210 Peter Medlin, "Illinois State Law Is the First to Have High Schools Teach News Literacy, *NPR*, August 12, 2021, https://www.npr.org/2021/08/12/1026993142/illinois-is-the-first-state-to-have-high-schools-teach-news-literacy.

211 Media Literacy Advisory Committee, "Media Literacy Advisory Committee Report," Colorado Department of Education, December 2019, https://www.cde.state.co.us/cdedepcom/medialiteracyadvisorycommitteereport.

212 Matthew Ingram, "Section 230 Critics Are Forgetting About the First Amendment," *Columbia Journalism Review*, July 29, 2021, https://www.cjr.org/the_media_today/section-230-critics-are-forgetting-about-the-first-amendment.php.

213 Alex Campbell, "How Data Privacy Laws Can Fight Fake News," *Just Security*, August 15, 2019, https://www.justsecurity.org/65795/how-data-privacy-laws-can-fight-fake-news/.

214 Alex Campbell, "How Data Privacy Laws Can Fight Fake News," *Just Security*, August 15, 2019, https://www.justsecurity.org/65795/how-data-privacy-laws-can-fight-fake-news/.

215 Thorin Klosowski, "The State of Consumer Data Privacy Laws in the US (And Why It Matters)," *New York Times*, September 6, 2021, https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us/.

216 Attorney General Bob Bonta, "California Consumer Privacy Act (CCPA)," State of California Department of Justice, https://oag.ca.gov/privacy/ccpa.

217 Attorney General Bob Bonta, "California Consumer Privacy Act (CCPA)," State of California Department of Justice, https://oag.ca.gov/privacy/ccpa.

218 Thorin Klosowski, "The State of Consumer Data Privacy Laws in the US (and Why It Matters)," *New York Times*, September 6, 2021, https://www.nytimes.com/wirecutter/blog/state-of-privacy-laws-in-us/.

219 Saja Hindi, "New Data Privacy Laws Lets Coloradans Choose Whether Companies Can Collect Information," *Denver Post*, July 7, 2021, https://www.denverpost.com/2021/07/07/data-privacy-colorado-new-law/.

220 Saja Hindi, "New Data Privacy Laws Lets Coloradans Choose Whether Companies Can Collect Information," *Denver Post*, July 7, 2021, https://www.denverpost.com/2021/07/07/data-privacy-colorado-new-law/.

221 Sarah Rippy, "Colorado Privacy Act Becomes Law," International Association of Privacy Professionals (IAPP), July 8, 2021, https://iapp.org/news/a/colorado-privacy-act-becomes-law/.

222 Sarah Rippy, "Colorado Privacy Act Becomes Law," International Association of Privacy Professionals (IAPP), July 8, 2021, https://iapp.org/news/a/colorado-privacy-act-becomes-law/.

223 Sarah Rippy, "Colorado Privacy Act Becomes Law," International Association of Privacy Professionals (IAPP), July 8, 2021, https://iapp.org/news/a/colorado-privacy-act-becomes-law/.

224 Kate Cox, "Virginia Is About to Get a Major California-Style Data Privacy Law," *Ars Technica*, February 11, 2021, https://arstechnica.com/tech-policy/2021/02/virginia-is-about-to-get-a-major-california-style-data-privacy-law/.

225 Kate Cox, "Virginia Is About to Get A Major California-Style Data Privacy Law," *Ars Technica*, February 11, 2021, https://

arstechnica.com/tech-policy/2021/02/virginia-is-about-to-get-a-major-california-style-data-privacy-law/.

226 Kate Cox, "Virginia Is About to Get A Major California-Style Data Privacy Law," *Ars Technica*, February 11, 2021, https://arstechnica.com/tech-policy/2021/02/virginia-is-about-to-get-a-major-california-style-data-privacy-law/.

227 Hayley Tsukayama, "Improving Enforcement in State Consumer Privacy Laws," Electronic Frontier Foundation, July 7, 2021, https://www.eff.org/deeplinks/2021/07/improving-enforcement-state-consumer-privacy-laws.

228 *See* Twitter, "Civic Integrity Policies," https://help.twitter.com/en/rules-and-policies/election-integrity-policy.

229 *See* Facebook, "Community Standards," https://transparency.fb.com/policies/community-standards/.

230 YouTube, "How Does YouTube Support Civic Engagement and Stay Secure, Impartial, and Fair During Elections?," https://www.youtube.com/howyoutubeworks/our-commitments/supporting-political-integrity/.

231 Twitter, "Civic Integrity Policies," https://help.twitter.com/en/rules-and-policies/election-integrity-policy.

232 Facebook, "Community Standards," https://transparency.fb.com/policies/community-standards/

233 Jon Lloyd et al., "Misinformation in the 2020 US Elections: A Timeline of Platform Changes," Mozilla Foundation, March 8, 2021, https://foundation.mozilla.org/en/blog/misinformation-in-the-2020-us-elections-a-timeline-of-platform-changes/.

234 Jon Lloyd et al., "Misinformation in the 2020 US Elections: A Timeline of Platform Changes," Mozilla Foundation, March 8, 2021, https://foundation.mozilla.org/en/blog/misinformation-in-the-2020-us-elections-a-timeline-of-platform-changes/.

235 Jon Lloyd et al., "Misinformation in the 2020 US Elections: A Timeline of Platform Changes," Mozilla Foundation, March 8, 2021, https://foundation.mozilla.org/en/blog/misinformation-in-the-2020-us-elections-a-timeline-of-platform-changes/.

236 Jon Lloyd et al., "Misinformation in the 2020 US Elections: A Timeline of Platform Changes," Mozilla Foundation, March 8, 2021, https://foundation.mozilla.org/en/blog/misinformation-in-the-2020-us-elections-a-timeline-of-platform-changes/.

237 Jesse Littlewood and Emma Steiner, "Trending in the Wrong Direction: Social Media Platforms' Declining Enforcement of Voting Disinformation," Common Cause, September 2, 2021, https://www.commoncause.org/wp-content/uploads/2021/09/Disinfo_WhitePaperv3.pdf.

238 Jon Lloyd et al., "Misinformation in the 2020 US Elections: A Timeline of Platform Changes," Mozilla Foundation, March 8, 2021, https://foundation.mozilla.org/en/blog/misinformation-in-the-2020-us-elections-a-timeline-of-platform-changes/.

239 Jesse Littlewood and Emma Steiner, "Trending in the Wrong Direction: Social Media Platforms' Declining Enforcement of Voting Disinformation," Common Cause, September 2, 2021, https://www.commoncause.org/wp-content/uploads/2021/09/Disinfo_WhitePaperv3.pdf.

240 Jesse Littlewood and Emma Steiner, "Trending in the Wrong Direction: Social Media Platforms' Declining Enforcement of Voting Disinformation," Common Cause, September 2, 2021, https://www.commoncause.org/wp-content/uploads/2021/09/Disinfo_WhitePaperv3.pdf.

241 Jeff Horwitz, "Facebook Has Made Lots of New Rules This Year. It Doesn't Always Enforce Them.," *Wall Street Journal*, October 15, 2020, https://www.wsj.com/articles/facebook-has-made-lots-of-new-rules-this-year-it-doesnt-always-enforce-them-11602775676.

242 Alex Heath, "Facebook to End Special Treatment for Politicians After Trump Ban," *The Verge*, June 3, 2021, https://www.theverge.com/2021/6/3/22474738/facebook-ending-political-figure-exemption-moderation-policy.

243 Facebook, "Our Approach to Newsworthy Content," July 29, 2021, https://transparency.fb.com/features/approach-to-newsworthy-content/.

244 Alex Heath, "Facebook to End Special Treatment for Politicians After Trump Ban," *The Verge*, June 3, 2021, https://www.theverge.com/2021/6/3/22474738/facebook-ending-political-figure-exemption-moderation-policy.

245 Alex Heath, "Facebook to End Special Treatment for Politicians After Trump Ban," *The Verge*, June 3, 2021, https://www.theverge.com/2021/6/3/22474738/facebook-ending-political-figure-exemption-moderation-policy.

246 Mike Isaac and Cecilia Kang, "Facebook Says It Won't Back Down from Allowing Lies in Political Ads," *New York Times*, September 4, 2020, https://www.nytimes.com/2020/01/09/technology/facebook-political-ads-lies.html.

247 Mike Isaac and Cecilia Kang, "Facebook Says It Won't Back Down from Allowing Lies in Political Ads," *New York Times*, September 4, 2020, https://www.nytimes.com/2020/01/09/technology/facebook-political-ads-lies.html.

248 Mike Isaac, "Whistler-Blower to Accuse Facebook of Contributing to Jan. 6 Riot, Memo Says," *New York Times*, October 13, 2021, https://www.nytimes.com/2021/10/02/technology/whistle-blower-facebook-memo.html.

249 Mike Isaac, "Whistler-Blower to Accuse Facebook of Contributing to Jan. 6 Riot, Memo Says," *New York Times*, October 13, 2021, https://www.nytimes.com/2021/10/02/technology/whistle-blower-facebook-memo.html.

250 Jesse Littlewood and Emma Steiner, "Trending in the Wrong Direction: Social Media Platforms' Declining Enforcement of Voting Disinformation," Common Cause, September 2, 2021, https://www.commoncause.org/wp-content/uploads/2021/09/Disinfo_WhitePaperv3.pdf.

251 Nick Clegg, "In Response to Oversight Board, Trump Suspended for Two Years; Will Only Be Reinstated if Conditions Permit," Facebook, June 4, 2021, https://about.fb.com/news/2021/06/facebook-response-to-oversight-board-recommendations-trump/.

252 Nick Clegg, "In Response to Oversight Board, Trump Suspended for Two Years; Will Only Be Reinstated if Conditions Permit,"

Facebook, June 4, 2021, https://about.fb.com/news/2021/06/facebook-response-to-oversight-board-recommendations-trump/.

253 Common Cause, "Common Cause and Over 20 Organizations Demand Facebook Close Loophole That Allows Trump to Stay on Platform Despite Ban," news release, July 26, 2021, https://www.commoncause.org/press-release/common-cause-and-over-20-organizations-demand-facebook-close-loophole-that-allows-trump-to-remain-on-platform-despite-ban/.

254 Parker Molloy, "Twitter's Enforcement Inconsistency Undermines Its Efforts at Policy Reforms," Media Matters for America, August 7, 2020, https://www.mediamatters.org/twitter/twitters-enforcement-inconsistency-undermines-its-efforts-policy-reforms.

255 Parker Molloy, "Twitter's Enforcement Inconsistency Undermines Its Efforts at Policy Reforms," Media Matters for America, August 7, 2020, https://www.mediamatters.org/twitter/twitters-enforcement-inconsistency-undermines-its-efforts-policy-reforms.

256 Twitter, "About Public-Interest Exceptions on Twitter," https://help.twitter.com/en/rules-and-policies/public-interest.

257 Sara Morrison, "Facebook and Twitter Made Special World Leader Rules for Trump. What Happens Now?," *Vox*, January 20, 2021, https://www.vox.com/recode/22233450/trump-twitter-facebook-ban-world-leader-rules-exception.

258 Twitter, "About Public-Interest Exceptions on Twitter," https://help.twitter.com/en/rules-and-policies/public-interest.

259 Courtney Subramanian, "A Minute-by-Minute Timeline of Trump's Day as the Capitol Siege Unfolded on Jan. 6," *USA Today*, February 11, 2021, https://www.usatoday.com/story/news/politics/2021/02/11/trump-impeachment-trial-timeline-trump-actions-during-capitol-riot/6720727002/.

260 Kate Conger, "Jack Dorsey Says Twitter Played a Role in U.S. Capitol Riot," *New York Times*, March 25, 2021, https://www.nytimes.com/2021/03/25/business/jack-dorsey-twitter-capitol-riot.html.

261 Daniel Dale, "Fact Check: 11 False Claims Rep. Marjorie Taylor Greene Has Tweeted in the Past Month," CNN, January 21, 2021, https://www.cnn.com/2021/01/21/politics/fact-check-marjorie-taylor-greene-twitter-election-capitol/index.html.

262 "YouTube Regrets," Mozilla Foundation, July 2021, https://foundation.mozilla.org/en/campaigns/regrets-reporter/findings/.

263 Casey Newton, "How YouTube Failed the 2020 Election Test," *The Verge*, March 4, 2021, https://www.theverge.com/2021/3/4/22313213/youtube-2020-election-misinformation-report-long-fuse.

264 Aaron Sankin, "YouTube Said It Was Getting Serious About Hate Speech. Why Is It Still Full of Extremists?," *Gizmodo*, July 25, 2019, https://gizmodo.com/youtube-said-it-was-getting-serious-about-hate-speech-1836596239.

265 "YouTube Regrets," Mozilla Foundation, July 2021, https://foundation.mozilla.org/en/campaigns/regrets-reporter/findings/.

266 "YouTube Regrets," Mozilla Foundation, July 2021, https://foundation.mozilla.org/en/campaigns/regrets-reporter/findings/.

267 "YouTube Regrets," Mozilla Foundation, July 2021, https://foundation.mozilla.org/en/campaigns/regrets-reporter/findings/.

268 Tripp Mickle, "Trump Is Still Banned on YouTube. Now the Clock Is Ticking.," *Wall Street Journal*, May 6, 2021, https://www.wsj.com/articles/trump-is-still-banned-on-youtube-that-could-change-11620293402.

269 Katie Canales, "YouTube Says Videos That Violate Its Policies Will Now Receive a 'Strike.' Channels That Receive 3 Strikes in a 90-Day Period Will Be Permanently Removed," *Business Insider*, January 7, 2021, https://www.businessinsider.com/youtube-three-strikes-permanently-delete-channels-policy-violations-trump-capitol-2021-1.

270 *See* Section 2 "Federal Voter Intimidation & False Election Speech Laws"; *see also* 52 U.S.C. § 10307(b); 52 U.S.C. § 10101(b); *United States v. North Carolina Republican Party*, 5:92-cv-00161 (E.D.N.C. 1992).

271 Nicole Hong, "Twitter Troll Tricked 4,900 Democrats in Vote-by-Phone Scheme, U.S. Says," *New York Times*, January 27, 2021, https://www.nytimes.com/2021/01/27/nyregion/douglass-mackey-arrested-far-right-twitter.html.

272 S. 1840, Deceptive Practices and Voter Intimidation Prevention Act of 2021, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/senate-bill/1840/text.

273 S. 2747, Freedom to Vote Act, 117th Cong. (2021), Sec. 3201-3206, "Deceptive Practices and Voter Intimidation Prevention Act of 2021," https://www.congress.gov/bill/117th-congress/senate-bill/2747/text.

274 S. 1, For the People Act, 117th Cong. (2021), Sec. 1301-1304, "Prohibiting Deceptive Practices and Preventing Voter Intimidation," https://www.congress.gov/bill/117th-congress/senate-bill/1/text.

275 S. 1840, Deceptive Practices and Voter Intimidation Prevention Act of 2021, 117th Cong. (2021), Sec. 3, "Prohibition on Deceptive Practices in Federal Elections," https://www.congress.gov/bill/117th-congress/senate-bill/1840/text.

276 S. 1840, Deceptive Practices and Voter Intimidation Prevention Act of 2021, 117th Cong. (2021), Sec. 3, "Prohibition on Deceptive Practices in Federal Elections," https://www.congress.gov/bill/117th-congress/senate-bill/1840/text.

277 S. 1840, Deceptive Practices and Voter Intimidation Prevention Act of 2021, 117th Cong. (2021), Sec. 4, "Corrective Action," https://www.congress.gov/bill/117th-congress/senate-bill/1840/text.

278 S. 1840, Deceptive Practices and Voter Intimidation Prevention Act of 2021, 117th Cong. (2021), https://www.congress.gov/

bill/117th-congress/senate-bill/1840/text.

279 Va. Code Ann. § 24.2-1005.1(A).

280 Amendment to Rules Comm. Print 117-13, Offered by Congressman Brad Sherman (CA-30), September 14, 2021, https://amendments-rules.house.gov/amendments/SHERMA_056_xml%20(Revised%20NDAA%20Amendment%20746)210917132434404.pdf.

281 S. 443, Democracy Is Strengthened by Casting Light On Spending in Elections (DISCLOSE) Act of 2021, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/senate-bill/443/text.

282 S. 1356, Honest Ads Act, 116th Cong. (2019), https://www.congress.gov/bill/116th-congress/senate-bill/1356/text.

283 S. 2747, Freedom to Vote Act, 117th Cong. (2021), Sec. 6001-6022, "DISCLOSE Act" and Sec. 6101-6110 "Honest Ads Act," https://www.congress.gov/bill/117th-congress/senate-bill/2747/text.

284 S. 1, For the People Act, 117th Cong. (2021), Sec. 4100-4122, "DISCLOSE Act" and Sec. 4201-4210 "Honest Ads Act," https://www.congress.gov/bill/117th-congress/senate-bill/1/text.

285 S. 443, Democracy Is Strengthened by Casting Light on Spending in Elections (DISCLOSE) Act of 2021, 117th Cong., Sec. 202-203, https://www.congress.gov/bill/117th-congress/senate-bill/443/text.

286 S. 443, Democracy Is Strengthened by Casting Light on Spending in Elections (DISCLOSE) Act of 2021, 117th Cong. (2021), Sec. 303, https://www.congress.gov/bill/117th-congress/senate-bill/443/text.

287 Honest Ads Act, S. 1356, 116th Cong. (2019), https://www.congress.gov/bill/116th-congress/senate-bill/1356/text.

288 Honest Ads Act, S. 1356, 116th Cong. (2019), https://www.congress.gov/bill/116th-congress/senate-bill/1356/text.

289 Federal Election Commission, "Legislative Recommendations of the Federal Election Commission 2021," Adopted May 6, 2021, https://www.fec.gov/resources/cms-content/documents/legrec2021.pdf.

290 H.R. 1272, Restoring Integrity to America's Elections Act, 116th Cong. (2019), https://www.congress.gov/bill/116th-congress/house-bill/1272/text.

291 H.R. 1, For the People Act, 117th Cong. (2021), Sec. 6001-6011, "Restoring Integrity to America's Elections," https://www.congress.gov/bill/117th-congress/house-bill/1/text.

292 S. 2747, Freedom to Vote Act, 117th Cong. (2021), Sec. 7101-7110, "Restoring Integrity to America's Elections," https://www.congress.gov/bill/117th-congress/senate-bill/2747/text.

293 *See* Alaska Stat. §§ 15.13.400(13), 15.13.040(j).

294 *See* Cal. Gov. Code § 84222.

295 *See* Cal. Assembly Bill 2188, signed September 26, 2018, https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB2188.

296 *See* NY Elec. Law §§ 14-107(2), 14-107(5-a), 14-106(4) and 14-107-b.

297 *See* Wash. Rev. Code § 42.17A.320; Wash. Admin. Code §§ 390-18-030 and 390-18-050.

298 Media Literacy Toolkit, PEN America, December 21, 2020, https://pen.org/media-literacy-toolkit/.

299 Eric McNeill, "Media Literacy Now launches next phase of media literacy campaign with model bill and new national coalition partners," Media Literacy Now, January 2, 2017, https://medialiteracynow.org/media-literacy-now-launches-next-phase-of-media-literacy-campaign-with-model-bill-and-new-national-coalition-partners/.

300 N.Y. Assembly Bill 6042.

301 N.Y. Assembly Bill 6042.

302 Katherine J. Wu, "Radical Ideas Spread through Social Media. Are the Algorithms to Blame?" *PBS*, March 28, 2019, https://www.pbs.org/wgbh/nova/article/radical-ideas-social-media-algorithms/.

303 Sheera Frenkel, "How Misinformation 'Superspreaders' Seed False Election Theories," *New York Times*, November 23, 2020, https://www.nytimes.com/2020/11/23/technology/election-misinformation-facebook-twitter.html.

304 S. 1896, Algorithmic Justice and Online Transparency Act, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/senate-bill/1896/text; H.R. 3611, Algorithmic Justice and Online Platform Transparency Act, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/house-bill/3611/text?r=2&s=1.

305 S. 1896, Algorithmic Justice and Online Transparency Act, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/senate-bill/1896/text; H.R. 3611, Algorithmic Justice and Online Platform Transparency Act, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/house-bill/3611/text?r=2&s=1.

306 Channel 4 News Investigations Team, "Revealed: Trump Campaign Strategy to Deter Millions of Black Americans from Voting in 2016," Channel 4, September 28, 2020, https://www.channel4.com/news/revealed-trump-campaign-strategy-to-deter-millions-of-black-americans-from-voting-in-2016.

307 "The Leadership Conference on Civil and Human Rights Transition Priorities," Leadership Conference on Civil and Human Rights, November 24, 2020, http://civilrightsdocs.info/pdf/policy/task-force-priorities/Transition-TaskForceTopPriorities-The%20Leadership%20Conference-November2020-FINAL.pdf.

308 "Letter from Common Cause, Free Press and PEN America to Congress," April 8, 2020, https://www.commoncause.org/press-release/congress-must-include-local-news-funding-in-next-covid-19-stimulus/.

309 Yosef Getachew and Jonathan Walter, "Restore Local Journalism with US Funding, Oversight Promoting Inclusiveness," *The Press of Atlantic City*, May 19, 2021, https://pressofatlanticcity.com/opinion/columnists/restore-local-journalism-with-us-funding-oversight-promoting-inclusiveness-by-yosef-getachew-and-jonathan-walter/article_602f794e-b416-11eb-ae1d-2f96f242750e.html.

310 Senator Brian Schatz, "Schatz, Veasey, Bennet, Klobuchar Reintroduce Legislation to Bolster Local Journalism," news release, May 13, 2021, https://www.schatz.senate.gov/news/press-releases/schatz-veasey-bennet-klobuchar-reintroduce-legislation-to-bolster-local-journalism.

311 "The Disinformation Black Box: Researching Social Media Data," Hearing Before the Subcomm. on Investigations and Oversight of the H. Comm. on Science, Space, and Technology, 117th Cong. (2021) (testimony of Laura Edelson, NYU Cybersecurity for Democracy).

312 H.R. 3451, Social Media DATA Act, 117th Cong. (2021), https://www.congress.gov/bill/117th-congress/house-bill/3451/text?r=1&s=1.

313 Vivek H. Murthy, "Confronting Health Misinformation: The U.S. Surgeon General's Advisory on Building a Healthy Information Environment," 2021, https://www.hhs.gov/sites/default/files/surgeon-general-misinformation-advisory.pdf.

314 Vivek H. Murthy, "Confronting Health Misinformation: The U.S. Surgeon General's Advisory on

315 Building a Healthy Information Environment," 2021, https://www.hhs.gov/sites/default/files/surgeon-general-misinformation-advisory.pdf.

316 "Letter from Public Interest Groups to White House," April 29, 2021, https://pen.org/letter-white-house-must-establish-disinformation-defense-and-free-expression-task-force/.

317 "Letter from Public Interest Groups to White House," April 29, 2021, https://pen.org/letter-white-house-must-establish-disinformation-defense-and-free-expression-task-force/.

318 *See, e.g.*, *United States v. North Carolina Republican Party*, 5:92-cv-00161 (E.D.N.C. 1992) (DOJ lawsuit against the North Carolina Republican Party, which had mailed disinformation postcards to 125,000 Black voters throughout the state incorrectly stating that recipients could not vote if they had moved within 30 days of the election and also threatening criminal prosecution).

319 *U.S. v. Douglas Mackey*, Complaint 1 (E.D.N.Y. January 22, 2021), https://www.justice.gov/opa/press-release/file/1360816/download.

320 Nicole Hong, "Twitter Troll Tricked 4,900 Democrats in Vote-by-Phone Scheme, U.S. Says," *New York Times*, January 27, 2021, https://www.nytimes.com/2021/01/27/nyregion/douglass-mackey-arrested-far-right-twitter.html.

321 Letter from Senate Democrats to Lina Khan, Chair, Federal Trade Commission, September 20, 2021, https://www.blumenthal.senate.gov/imo/media/doc/2021.09.20%20-%20FTC%20-%20Privacy%20Rulemaking.pdf.

322 Letter from Public Interest Groups to Lina Khan, Chair, Federal Trade Commission, August 4, 2021, https://www.publicknowledge.org/documents/public-interest-group-ftc-privacy-letter/.

323 Letter from Senate Democrats to Lina Khan, Chair, Federal Trade Commission, September 20, 2021, https://www.blumenthal.senate.gov/imo/media/doc/2021.09.20%20-%20FTC%20-%20Privacy%20Rulemaking.pdf; Letter from Public Interest Groups to Lina Khan, Chair, Federal Trade Commission, August 4, 2021, https://www.publicknowledge.org/documents/public-interest-group-ftc-privacy-letter/.

324 Federal Election Commission, "Advance Notice of Proposed Rulemaking: Internet Communication Disclaimers," 76 Fed. Reg. 63567, October 13, 2011.

325 Common Cause, Comments in Response to FEC Notice of Proposed Rulemaking 2018-06, May 24, 2018, https://www.commoncause.org/wp-content/uploads/2018/05/CC-Comments_NPRM-2018-06-Internet-Disclaimers_FINAL-5.24.18.pdf.

326 Jesse Littlewood and Emma Steiner, "Trending in the Wrong Direction: Social Media Platforms' Declining Enforcement of Voting Disinformation," Common Cause, September 2, 2021, https://www.commoncause.org/wp-content/uploads/2021/09/Disinfo_WhitePaperv3.pdf.

327 Craig Silverman and Ryan Mac, "Facebook Promised to Label Political Ads, But Ads for Biden, the Daily Wire, and Interest Groups Are Slipping Through," *BuzzFeed News*, October 22, 2020, https://www.buzzfeednews.com/article/craigsilverman/facebook-biden-election-ads.

328 Jeremy B. Merrill and Jamiles Lartey, "Trump's Crime and Carnage Ad Blitz Is Going Unanswered on Facebook," The Marshall Project, September 23, 2020, https://www.themarshallproject.org/2020/09/23/trump-s-crime-and-carnage-ad-blitz-is-going-unanswered-on-facebook.

329 Keach Hagey and Jeff Horwitz, "Facebook Tried to Make Its Platform a Healthier Place. It Got Angrier Instead," *Wall Street Journal*, September 15, 2021, https://www.wsj.com/articles/facebook-algorithm-change-zuckerberg-11631654215?mod=article_inline.

330 Jeff Horwitz, "Facebook Says Its Rules Apply to All. Company Documents Reveal a Secret Elite That's Exempt.," *Wall Street Journal*, September 13, 2021, https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353?mod=article_inline.

331 Common Cause, "Facebook Shutdown of NYU Ad Observatory Project Researchers Undermines Democracy," news release, August 4, 2021, https://www.commoncause.org/press-release/facebook-shutdown-of-nyu-ad-observatory-project-researchers-undermines-democracy/.

332 Kari Paul, "Facebook Has a Blind Spot': Why Spanish-Language Misinformation Is Flourishing," *The Guardian*, March 3, 2021, https://www.theguardian.com/technology/2021/mar/03/facebook-spanish-language-misinformation-covid-19-election.

333 Ariz. Rev. Stat. § 16-1013.

334 Ariz. Rev. Stat. § 16-1012.

335 Ariz. Rev. Stat. § 16-515.

336 Other states offering candidate pledges related to false statements include Arkansas (Ark. Code § 7-6-102), Illinois (10 ILCS § 5/29B-10), Maine (Me. Stat. tit. 21-A, § 1101 et seq.), Minnesota (Minn. Stat. § 211B), Nevada (Nev. Rev. Stat. § 294A.290), and Texas (Tex. Elec. Code § 255.004-06).

337 Cal. Elec. Code § 20440.

338 Colo. Rev. Stat. § 1-13-713.

339 Colo. Rev. Stat. § 1-13-109.

340 Colorado Attorney General Phil Weiser, "Public Advisory on Voter Intimidation Crimes and Poll Security," October 19, 2020, https://coag.gov/app/uploads/2020/10/Public-Advisory-Voter-Intimidation-10.19.2020.pdf.

341 Fla. Stat. § 104.0615(2).

342 Fla. Stat. § 104.0615(3).

343 Ga. Code Ann. § 21-2-566(4).

344 Ga. Code Ann. § 21-2-567(a).

345 Haw. Rev. Stat. § 19-3(12).

346 Haw. Rev. Stat. § 19-3(4).

347 Haw. Rev. Stat. § 11-132(5).

348 Me. Stat. tit. 21, § 674(1)(B).

349 Me. Stat. tit. 21, § 674(3)(A).

350 Md. Code, Com. Law § 16-201(a)(5)-(6).

351 Md. Code, Com. Law § 16-201(a)(7).

352 Mich. Elec. Laws § 168.932(a).

353 Mich. Elec. Laws § 168.931(3).

354 Minn. Stat. § 211B.07.

355 Minn. Stat. § 211B.07.

356 Nev. Rev. Stat. § 293.710(1)(d).

357 Nev. Rev. Stat. § 293.710(1)(a)-(b).

358 N.M. Stat. § 1-20-14.

359 New Mexico Secretary of State Maggie Toulouse Oliver, "Guidance on Voter Intimidation and Discriminatory Conduct," https://www.sos.state.nm.us/voting-and-elections/voter-information-portal/guidance-on-voter-intimidation-and-discriminatory-conduct/.

360 N.C. Gen. Stat. § 163-274(7).

361 25 Pa. Cons. Stat. § 1711.

362 Pa. Department of State, "Guidance on Voter Intimidation and Discriminatory Conduct," October 2020, https://www.dos.pa.gov/VotingElections/OtherServicesEvents/Documents/DOS%20Voter%20Intimidation%20Guidance%2010.14.16.pdf.

363 Va. Code Ann. §§ 24.2-607 and 24.2-1005.

364 Va. Code Ann. § 24.2-1002.

365 Va. Code Ann. § 24.2-1005.1(A).

366 Va. Code Ann. § 24.2-1005.1(C).

367 Wis. Stat. § 12.09.

368 Wis. Stat. § 12.05.

369 Alaska applies disclosure requirements to "nongroup entities," defined to mean "a person, other than an individual, that takes action the major purpose of which is to influence the outcome of an election" and that "cannot participate in business activities," "does not have shareholders who have a claim on corporate earnings" and is "independent from the influence of business corporations." Alaska Stat. § 15.13.400(13).

370 Alaska Stat. § 15.13.040(j).

371 Cal. Gov. Code § 84222.

372 Cal. Assembly Bill 2188, signed September 26, 2018, https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB2188.

373 Md. Code, Elec. Law § 13-309.2.

374 Minn. Stat. Ann. § 10A.27 Subd.15(b).

375 NY Elec. Law §§ 14-107(2), 14-107(5-a) and 14-107-b.

376 NY Elec. Law § 14-106(4).

377 17 R.I. Gen. Laws § 17-25.3-1.

378 Wash. Rev. Code § 42.17A.320.

379 Wash. Admin. Code § 390-18-030.

380 Wash. Admin. Code § 390-18-050.

381 "U.S. Media Literacy Policy Report 2020," Media Literacy Now, January 2020, https://medialiteracynow.org/wp-content/uploads/2020/01/U.S.-Media-Literacy-Policy-Report-2020.pdf.

382 Media Literacy Advisory Committee, "Media Literacy Advisory Committee Report," Colorado Department of Education, December 2019, https://www.cde.state.co.us/cdedepcom/medialiteracyadvisorycommitteereport.

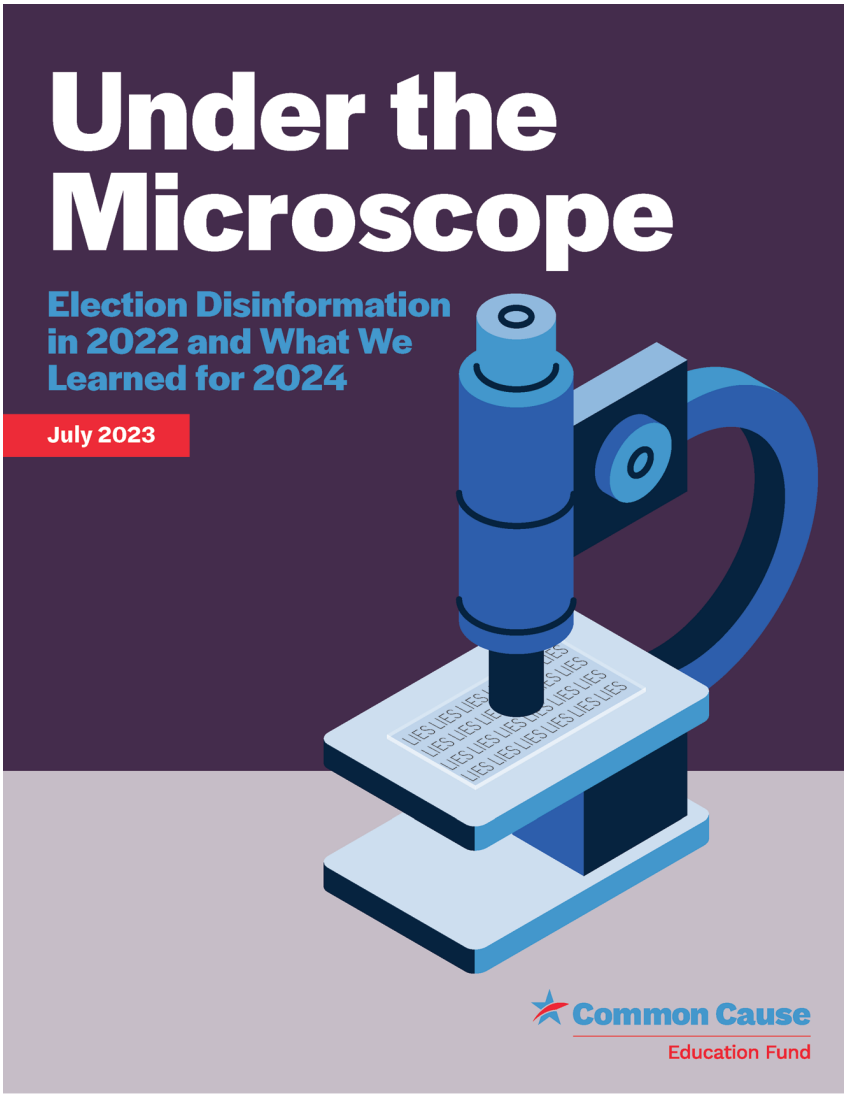
383 "U.S. Media Literacy Policy Report 2020," Media Literacy Now, January 2020, https://medialiteracynow.org/wp-content/uploads/2020/01/U.S.-Media-Literacy-Policy-Report-2020.pdf.

384 Peter Medlin, "Illinois State Law is the First to Have High Schools Teach News Literacy," *NPR*, August 12, 2021, https://www.npr.org/2021/08/12/1026993142/illinois-is-the-first-state-to-have-high-schools-teach-news-literacy.

385 "U.S. Media Literacy Policy Report 2020," Media Literacy Now, January 2020, https://medialiteracynow.org/wp-content/uploads/2020/01/U.S.-Media-Literacy-Policy-Report-2020.pdf.

386 "U.S. Media Literacy Policy Report 2020," Media Literacy Now, January 2020, https://medialiteracynow.org/wp-content/uploads/2020/01/U.S.-Media-Literacy-Policy-Report-2020.pdf.

Common Cause
Education Fund
805 15th Street, NW, Suite 800
Washington, DC 20005
202.833.1200
commoncause.org/education-fund/


Under the Microscope
Election Disinformation in 2022 and What We Learned for 2024
July 2023
LIES LIES LIES
Common Cause
Education Fund

About Common Cause Education Fund

The Common Cause Education Fund is the research and public education affiliate of Common Cause, founded by John Gardner in 1970. We work to create open, honest, and accountable government that serves the public interest; promote equal rights, opportunity, and representation for all; and empower all people to make their voices heard in the political process.

Acknowledgments

This report was authored by Emma Steiner with key assistance and editorial guidance from Jesse Littlewood, and published by Common Cause Education Fund. Common Cause Education Fund's disinformation work has been provided by the FJC- A Foundation of Philanthropic Funds, the Minami Tamaki Yamauchi Kwok & Lee Foundation, the Trusted Elections Fund, the Marks Family Foundation, and the Leadership Conference on Civil and Human Rights. We are incredibly grateful for their support.

We thank Allegra Chapman, Marilyn Carpinteyro, Sylvia Albert, Kathay Feng, Susannah Goodman, Liz Iacobucci, Sam Vorhees, Stephen Spaulding, and Aaron Scherb for guidance and editing; Melissa Brown Levine for copy-editing; Kerstin Vogdes Diehn for design; Jack Mumby and Ashlee Keown for digital support; and Katie Scally and David Vance for strategic communications support. Special thanks are due to the volunteers of Common Cause Education Fund's Social Media Monitoring Program, led by Raelyn Roberson, and the volunteers of the Algorithmic Transparency Institute's Civic Listening Corps, led by John Schmidt, who spent thousands of hours monitoring social media for disinformation that could disenfranchise voters. Many of the examples in this report came from volunteers in these programs.

EXECUTIVE SUMMARY

Election disinformation looms large with the 2023 elections underway and the high-profile 2024 races already unfolding. As we enter these new disinformation threat environments, we take a look back at what we learned about election disinformation and its continuing evolution from previous election cycles through to today.

In our 2021 report on election disinformation, *As a Matter of Fact*, we illustrated how election disinformation causes digital and social harms and included recommendations for lawmakers and platforms in the lead-up to the 2022 elections.

The 2022 elections presented a wide array of new challenges—and a new starting point for election disinformation. Now, a significant portion of the electorate and the people running to represent them believe—and find it convenient to say—that our elections are illegitimate. We're reporting back on the challenges voters faced, our efforts to disrupt disinformation in 2022, and what challenges lie ahead for the 2024 elections.

First, this report details our key findings from the 2022 election cycle and provides definitions for key terms. The report is then broken into sections detailing the lead-up and state of play heading into 2024, findings from our nonpartisan Election Protection coalition's successes and lessons learned, and what lies ahead, as well as recommendations for legislative proposals to protect voters.

Section One, "The Lead-Up to 2022," covers the role of profit in election denial, the outsized influence of social media platforms, threats of political violence, and information gaps.

Section Two, "Common Cause Education Fund's Work in 2022," covers our work identifying, flagging, and removing election disinformation, inoculating voters against disinformation, the role of partnerships in our Election Protection work, our work with media, and two case studies demonstrating successful interventions.

Section Three, "Looking Ahead," covers the role of election disinformation in candidacies, how tech platforms are backing down from enforcing their policies against disinformation, the remaining threat of political violence, how election disinformation fuels voter suppression through attacks on the voting process, and legislative solutions.

KEY FINDINGS/POINTS

The year 2022 was a more challenging environment for voters than 2020, despite growing familiarity with new tools and methods of voting in a pandemic.

- Right-wing partisan influencers, politicians, and activists alike used the popularity of election denial to build audiences and profit off of lies about voting and elections.
- Due to profit incentives and competition, the tech industry further relaxed already-inadequate standards for content moderation around election disinformation and showed no interest in changing algorithmic design to tackle the issue.
- The 2022 elections were the first federal election since the January 6th insurrection, and, as such, advocates feared a resurgence of political violence and domestic violent extremism—which continue to be a threat today.
- The potential for vulnerable populations who exist within information voids and news deserts to be targeted by disinformation was greater than ever.

Nevertheless, the nonpartisan Election Protection community scored some key successes as a coalition to protect voters from voter suppression.

- Common Cause Education Fund's Social Media Monitoring program discovered emerging disinformation narratives about elections and voting and pushed real-time intelligence about disinformation to the Election Protection community.
- We worked as a coalition to promote positive, pro-voter inoculation content about the importance of election workers, counting every vote, and other hot-button voting and election issues.
- As a coalition, we implemented lessons learned from 2020 to build on messaging and make sure it reached more audiences and touched on a wider range of subjects than before, keeping in mind the populations specifically targeted by disinformation.
- We helped educate the media on how to accurately and responsibly report on election disinformation and saw noted changes in how information was conveyed to voters.

Despite these successes, we know that the road ahead won't be easy. The 2024 election cycle will present unique challenges in addition to the ones that are now standard in elections.

- Election disinformation is now essentially obligatory for nearly all Republican primary candidates and opportunists seeking financial benefit.
- It will be even more difficult to rely on tech platforms acting responsibly and voluntarily enforcing their policies.
- There is remaining potential for political violence to flare up in the wake of potential indictments of a presidential candidate and a primary focused on rehashing lies about 2020.
- More legislative norms are being eaten away by lawmakers seeking to cultivate support from a base steeped in conspiracy theories—who introduce legislation premised on lies about elections.

It will be even more difficult to rely on tech platforms acting responsibly and voluntarily enforcing their policies.

- Any institution, tool, or practice, even ones with bipartisan support and buy-in, can become a target of disinformation.
- New legislation points a way forward for grappling with both emerging and existing threats to voters.

DISINFORMATION is false rhetoric used to mislead. In elections, it's used to dampen turnout among some voters, mobilize others based on lies, or call into question the results if an opponent wins in an attempt to either overturn the election or profit off of the chaos. Disinformation can alter voter participation, potentially causing voters to miss their opportunity to vote if they are confused about the voting process (the time, place, and manner of the election) or choose to stay home ("self-suppress") due to worries about intimidation, violence or other consequences. Election disinformation also alters public perceptions about elections and their security, thereby impacting legislation and democratic norms in the long run.

ELECTION DENIAL: New research by the Massachusetts Institute of Technology delves into the roots of election denialism, and finds that it is largely motivated by racial resentment: "Among Republicans, conspiracism has a potent effect on embracing election denialism, followed by racial resentment. Among independents, the strongest influences on denialism are Christian nationalism and racial resentment. And, although election denialism is rare among Democrats, what variation does exist is mostly explained by levels of racial resentment."

Election denial is motivated by the belief that others' votes are lesser and shouldn't count, and that the only way forward is to overturn undesired electoral outcomes. This belief is racist at its very core and the forms it takes target members of marginalized populations. The fact that regardless of partisan affiliation, election denial is rooted in racial resentment is a reminder that any attempt to combat disinformation must acknowledge and uplift those most affected by it.

SECTION ONE: THE LEAD-UP TO 2022

The Role of Profit in Election Denial

Right-wing partisan influencers, politicians, and activists alike used the popularity of election denial to build audiences and profit off of lies about voting and elections.

A small group of mostly right-wing personalities is responsible for "super-spreading" voter fraud myths, spawning millions of online interactions around false and misleading stories. For example, in the four-week period from mid-October to mid-November of 2020, then-president Donald Trump and the "top 25 superspreaders of voter fraud misinformation accounted for 28.6 percent of the interactions people had with that content."

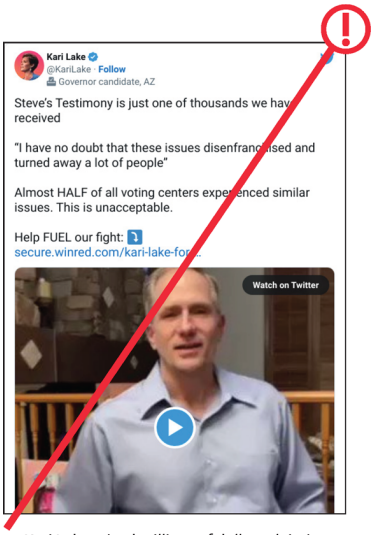
This trend holds true with podcasts too. A new analysis from the Brookings Institution shows that political podcasts are a consistent vector of disinformation. In their review of 79 different political podcasts, Brookings analysts found that "10 prominent podcasters were responsible…for more than 60% of all the dataset's unsubstantiated and false claims." Whether it's COVID-19 vaccine disinformation or election disinformation, time and time again, research has shown that the bulk of engagement on disinformation is driven by a few superspreaders across social media platforms and other forms of media.

Not only is it lucrative to sell ads on channels that promote popular conspiracy theories, but there are endless opportunities to fundraise off of an election loss.

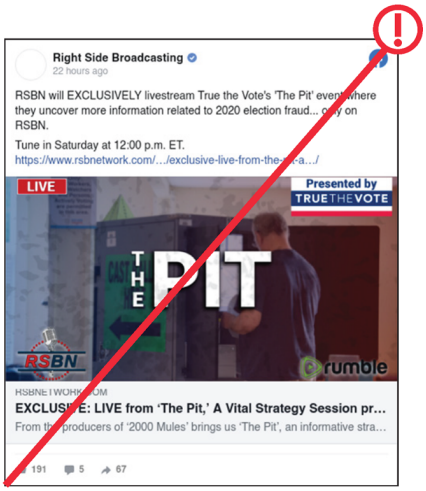
Many social media influencers who spread disinformation start with election denial and use that to build an audience before moving on to spreading disinformation about other issues, such as COVID-19 denial and climate change denial. They are then able to use the audience they have cultivated to promote more and more false claims. Meanwhile, a cottage industry of election-denying influencers, activists, and fake analysts has benefited from online election disinformation.

There are multiple revenue streams an election denier can tap into with election disinformation. Not only is it lucrative to sell ads on channels that promote popular conspiracy theories, but there are endless opportunities to fundraise off of an election loss. For example, failed Arizona gubernatorial candidate Kari Lake raised $2.5 million after her loss, pledging to fight her defeat in the courts and overturn the results of the election. Trump himself fundraised off of election denial, raising over $250 million for recounts, legal fights, and promises to "Stop the Steal," most of which was spent on unrelated expenses. Issue One recently found an entire network of political consultants and companies who profited from supporting the campaigns of secretary of state candidates who promoted election denial.

Other vote suppressor groups like "True the Vote" fundraise for both themselves and for sham election reviews—the cost of which was over $9 million raised by Trump backers and given to contractors, who themselves profited from election denial. Some figures continue to tour the country giving talks as "experts," continually raising funds to keep their sideshow going. In what Accountable.US describes as the "election denial industrial complex," just a small group of attorneys, politicians, influencers, and vote suppressors are able to raise vast sums of money for their forays into election denial—and the opportunities to grift just keep coming.



Kari Lake raised millions of dollars claiming widespread disenfranchisement of Republican voters in 2022.



Groups like True the Vote raise their profile and fundraise through the organization of events that promise evidence of widespread election fraud.

Social Media's Outsized Role in Election Disinformation

Because of profit incentives and competition, the tech industry further relaxed already-inadequate standards for content moderation around election disinformation and showed no interest in changing algorithmic design to tackle the issue.

Major social media platforms such as Twitter, Facebook, TikTok, and YouTube play an integral role in spreading information about our democracy. More than 70% of U.S. residents use social media, and half of the adults in the United States "often" or "sometimes" get their news from social media. Social media has provided platforms for diverse voices and viewpoints, allowing users to find information from voices they trust. However, while social media has no doubt provided access to news and information and provided a voice and platform for many voices, research shows that social media has its darker side: it heightens polarization in this country, fuels white nationalism and racism by providing a space to organize and radicalize, and contributes to racialized disinformation and organized hate against marginalized groups. The internet affects who is targeted by election disinformation and who has access to reliable information online.

More than 70% of U.S. residents use social media, and half of the adults in the United States "often" or "sometimes" get their news from social media.

Existing moderation standards on social media networks have gaps that disinformers exploited and continue to exploit to profit from false claims. It doesn't help that negative incentives pressure social media platforms to

decline to take action—after all, vitriolic engagement is still engagement, and it brings clicks and eyeballs to sites that are in fierce competition against each other. There's political, social, and financial pressure not to remove bad actors for fear of impacting revenue streams and inciting backlash.

The January 6th Select Committee also found in their unreleased report on social media that "platforms' lax enforcement against violent rhetoric, hate speech and the big lie stemmed from longstanding fear of scrutiny from elected officials and government regulators." This situation persisted past 2020—in Common Cause Education Fund's 2021 report *Trending in the Wrong Direction,* we found that in many cases, platforms backed down on their existing civic integrity policies without saying anything, as opposed to their announcements when they instituted the policies. While in the immediate aftermath of the 2020 election, social media platforms acted swiftly to react and remove inciting content, within months, these policies were relaxed—and posts that would have been actioned were allowed to gain massive engagement once more. The problem was exacerbated in languages other than English, with ever fewer resources dedicated to, for example, Spanish-language moderation and fact-checks.

As shown by Common Cause Education Fund's 2021 report and our 2021 white paper, moderation at major tech platforms has been inadequate at best and backsliding at worst. Civil society groups thought that if we could point out places where moderation wasn't happening, social media companies would engage with us, fix it, and learn to prevent gaps in enforcement in the future. But during the 2022 election cycle, recent tech layoffs made it difficult for civil society advocates to even know where to reach out—and made it harder for platforms themselves to conduct the basic functions of moderation. To account for this, we raised our concerns more publicly.

For the 2022 midterm, given platforms' backsliding and impenetrable moderation standards, Common Cause led 130 public interest organizations to draft and submit a letter to leading social media platforms, advising them to monitor and reduce mis- and disinformation through implementing the following: "auditing algorithms that look for disinformation, downranking known falsehoods, creating full-time civic integrity teams, ensuring policies are applied retroactively—i.e., to content posted before the rule was instituted—moderating live content, sharing data with researchers and creating transparency reports on enforcement's effectiveness." We know what platforms need to do to reduce the spread of disinformation—they just refused, and continue to refuse, to make better choices for user safety.

Threats of Political Violence

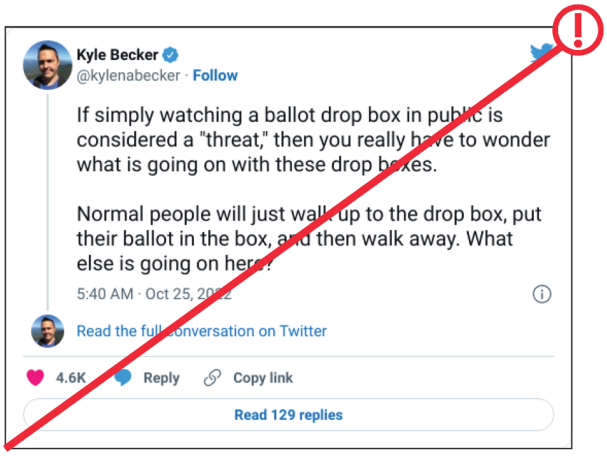
The 2022 elections were the first federal election since the January 6th insurrection, and, as such, advocates feared a resurgence of political violence and domestic violent extremism—which continue to be a threat today.

Going into 2022, we had reason to be nervous about the potential for political violence. It was unknown how vote suppressors and election deniers would react to electoral outcomes, and there were new trends of targeting voters and continued targeting of election workers.

One telling example from 2020: after the Trump campaign took video footage of Fulton County, Georgia, election workers Ruby Freeman and Shaye Moss out of context to claim that they were engaged in fraud, the two women were targeted by mass harassment and threats. This was egged on not only by Trump's lawyer Rudy Giuliani but also by Trump himself in widely viewed social media posts alleging crimes. Trump even mentioned Freeman over a dozen times in his infamous call asking Georgia officials to overturn the election. This campaign of harassment led to death threats and visits to the women's homes, and resulted in Moss and Freeman having to flee their residences. Trump even amplified attacks on Freeman after the release of her testimony, years later, to the January 6th Select Committee, asking "What will the Great State of Georgia do with the Ruby Freeman MESS?"

Voters also feared intimidation at polling locations: A poll from fall 2022 showed that "35% of Black Americans believe violence is likely or very likely at their polling place in November." Reuters reported that 40% of voters were concerned about intimidation at the polls. One particular trend of concern was drop box surveillance, organized by election deniers on Truth Social. Election deniers with militia ties announced their intent to stand guard at ballot drop boxes and ceased action in Maricopa County only in response to a court order.

A poll from fall 2022 showed that "35% of Black Americans believe violence is likely or very likely at their polling place in November."



Ballot drop boxes became a target for vigilante surveillance in 2022.

Information Gaps and Vulnerable Voters

The potential for vulnerable populations who exist within information voids and news deserts to be targeted by disinformation was greater than ever.

As we entered the 2022 midterms, researchers like those at the Brennan Center for Justice at NYU Law warned that information gaps—"when there is high demand for information about a topic, but the supply of accurate and reliable information is inadequate to meet that demand"—would present an issue for voters. This was further exacerbated by the fact that disinformers were relying on disinformation from 2020 to set a foundation for disinformation in 2022. The Brennan Center cited declining voter trust in elections and lack of public outreach about changes to voting procedures.

Heading into the midterms, only 47% of Americans polled had a "great deal" of confidence that 2022 votes would be counted properly, and Election Protection advocates had to thread a difficult needle of reassuring voters who had been exposed to election disinformation while also encouraging turnout.

Meanwhile, nonincumbent political candidates were adding to the problem. A study last year from NYU's Center for Social Media and Politics found that "politicians in the 2022 election are sharing more links to unreliable news sources than they did in 2020, and the increase appears to be driven by nonincumbent Republican candidates." The partisan difference in usage of unreliable sources was staggering: "36 percent of news that Republican candidates shared came from unreliable sites, while that was true for only 2 percent of news shared by Democratic candidates each day."

Another study found that YouTube's algorithm shows more election-fraud content to accounts already "skeptical" of elections, creating a feedback loop of conspiracy content. As the 2022 threat framework from the Election Integrity Partnership detailed, social media disinformation was particularly suited for viral spread because of factors like the potential for massive engagement. This meant that people in news deserts, people targeted by disinformation, and people who rely on social media for news would potentially be more exposed to disinformation about the election.

SECTION TWO: COMMON CAUSE EDUCATION FUND'S WORK IN 2022

Common Cause Education Fund's Social Media Monitoring program discovered emerging disinformation narratives about elections and voting and pushed real-time intelligence about disinformation to the Election Protection community.

As one of the co-leads of the Election Protection coalition, Common Cause Education Fund took a leadership role in developing the strategy to combat election disinformation in 2022. We were uniquely positioned to reduce the impact and spread of disinformation through our Stopping Cyber Suppression program and through engagement with our national and state partners. Our interventions on social media protected voters and helped train grassroots volunteers to defend themselves and their communities from disinformation.

Identifying, Flagging, and Removing Election Disinformation

Over the 2022 primaries and general November election, we **recruited and trained 2,202 monitors** who in total **submitted 3,825 items of potential social media disinformation for review** to our team. On Election Day itself, **156 Common Cause volunteers** (plus an additional **44 youth volunteers**) gathered over **750 items** of potential social media disinformation. Hundreds of additional volunteer monitors, whom we helped train, worked with our state and local partners on the ground. We were additionally involved in the Georgia Senate runoff with **33 social media monitoring volunteers**.

We recruited and trained 2,202 monitors who in total submitted 3,825 items of potential social media disinformation for review to our team.

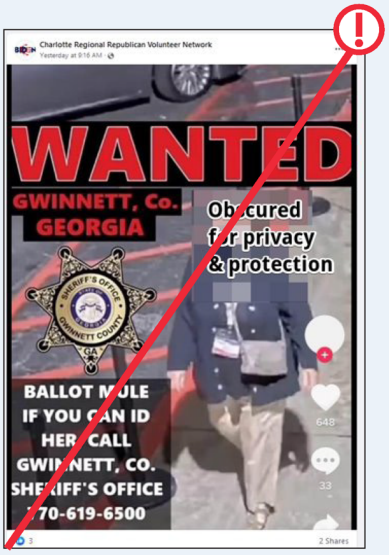
When we receive potential disinformation, we triage it based on importance and impact. For content that might violate social media platform policies, we immediately report and request removal of the content. We also review content to determine if it is a growing trend or narrative of disinformation, looking at the content gathered by our volunteers and by partners. We then create pro-voter talking points and social media posts that push back on the disinformation narratives and circulate those to our national and state partners.

Removing disinformation from social media platforms was challenging and is only becoming more so. However, we were able to remove over 300 social media posts across Facebook and Twitter in the 2022 election cycle. Some of these posts were threatening in nature, not only targeting specific individuals but creating a climate of fear around voter participation.

CASE STUDY: Stopping Threatening "Ballot Mule" Disinformation

In early 2022, a film named *2,000 Mules* by known disinformation spreader Dinesh D'Souza was released. The film was created in collaboration with voter suppression and disinformation group True the Vote and falsely alleges mass "ballot trafficking" by so-called ballot mules and makes explicit false claims about election and voting procedures. We were on alert that these false claims could become a viral disinformation narrative on Facebook, Twitter, and other platforms given the high-profile nature of the movie production team involved and the continued salience of election conspiracy theories as we headed into the

2022 midterms. And given the increase in intimidation and threats to elections officials and voter protection groups created by election disinformation, we knew that this movie would turn up the temperature and increase the conditions that lead to threats, intimidation, and political violence.



Common Cause reached out to our contacts at Meta (Facebook/Instagram) before the movie was released with an early warning, alerted them to the issues that this would likely create, and again after the release, highlighted to social media companies the viral nature of this disinformation narrative and fact-checks from PolitiFact and the Associated Press.

Our monitoring program soon identified trending viral content on Facebook and Twitter related to this "ballot mule" narrative. It was being posted by multiple users across multiple social media platforms, either inspired by or coordinated with the film. Much of the content was threatening in nature—making accusations about individuals or groups that they were involved in "ballot trafficking" as a "ballot mule"—with no evidence (and none has since been presented).

Our monitoring did identify some "ballot mule" content that was manipulated media—posts that falsely invoked law enforcement action targeting an individual involved in elections. We took immediate action, contacting Meta and Twitter about posts on their platforms that contained this manipulated media that directly targeted an individual—posts that had racked up tens of thousands of likes, comments, and shares despite the threat.

While Twitter took action over the next few days to remove these posts and similar posts, it took outside pressure with two media stories about this issue until Meta finally responded that this content violated their policies. Even after that point, we could find examples live on the platform—and only when we reported them were they removed. In total, over 200 pieces of this kind of threatening content were removed from Facebook and Twitter.

In the end, our advocacy resulted in Twitter and Facebook both enforcing and updating their policies to make this kind of intimidation content prohibited on their platform, a major success that will keep this content off the platforms. That said, as our experience shows, we have to remain vigilant and continue to monitor to ensure appropriate action is taken.

Inoculating Voters Against Disinformation

We worked as a coalition to promote positive, pro-voter inoculation content about the importance of election workers, counting every vote, and other hot-button voting and election issues.

A second key component of our work to combat election disinformation is to stop disinformation from taking root in the first place—to "inoculate" audiences against potential disinformation. Numerous studies have shown that when individuals are provided accurate information about a topic from a trusted messenger, it reduces the impact of disinformation. While this is broadly true across different issue areas, in voting and elections, it is especially critical, as voters can miss their opportunity to participate if they fall for disinformation or choose to "self-suppress" based on false narratives.

A second key component of our work to combat election disinformation is to stop disinformation from taking root in the first place.

Stopping election disinformation is imperative to achieving true multiracial democracy with equal participation. Voters most at risk from election disinformation are new voters and infrequent voters who don't have as much experience navigating our elections system, voters with limited English proficiency (as the bulk of voting information is in English), and students and other transient populations who are not as likely to be engaged by the parties. Often it is voters of color (especially those in immigrant communities) and young voters who don't have the information needed or experience with voting, which compounds the impact of election disinformation. Black and Latinx Americans are three times more likely than white Americans to be told they lack correct voting identification, to be unable to locate a polling place, or to miss a registration deadline. And more than half of voters under the age of 35 (who are more diverse than voters over 35) do not have the resources or knowledge they need to vote by mail and are therefore more susceptible to mis- and disinformation.

Stopping election disinformation is imperative to achieving true multiracial democracy with equal participation.

Common Cause worked with the Leadership Conference on Civil and Human Rights to create content that combines voter information and messaging with "prebunking" to stop disinformation before it takes root in a community. These messaging guides and example content were created and distributed to the entire Election Protection network.

One key element is localized content for specific states. Similarly, we coordinated the translation of inoculation content into the languages of communities that are targeted by disinformation. The state-level organizations that are trusted messengers in their communities must have the resources, strategy, and capacity to effectively inoculate their communities against disinformation.

Election officials are important sources of trusted information, and the National Association of Secretaries of State has a public education campaign designed to lift up and amplify the voices of elections officials. However, elections officials are often underfunded, understaffed, and have limitations on the reach of their content.

Thanks to the support from funders and partners, we invested additional resources in 2022 to update and expand our inoculation content that could be communicated by the diversity of trusted nonpartisan sources. In 2020, we found that the most shared and spread content used bright, engaging illustrations reflecting democratic values while sharing our key messages. We worked with illustrators to create dozens of images specific to the disinformation narratives we needed to combat (based on the intelligence we gathered from our monitoring).

In 2022, we also put a premium on translating our content into Spanish given the prevalence of disinformation in Spanish and the limited resources social media companies put into combating non-English disinformation.

Using trusted messengers to communities is key for successful inoculation, and our partner network was critical to this effort. To make it easier to share inoculation content, we created an online searchable database for partner organizations. As disinformation narratives and threats changed, we continued to add content to this database as the calendar proceeded (moving to post-election inoculation content in the week before the election) and as new narratives came up.

Our content was mentioned in other posts 15,000 times, engaged with by 60 million social media users (who clicked or engaged with the content), and viewed a total of 298 million times

Tracking reach is challenging, but through the use of the #OurElections hashtag and the analytics provided by partners, we believe we had millions of views of our content on Facebook, Twitter, and TikTok. In fact, our content was mentioned in other posts 15,000 times, engaged with by 60 million social media users (who clicked or engaged with the content), and viewed a total of 298 million times.

We also found that creating vibrant graphics that celebrated voter participation from all types of voters was necessary if partner groups were going to share these graphics and messages.

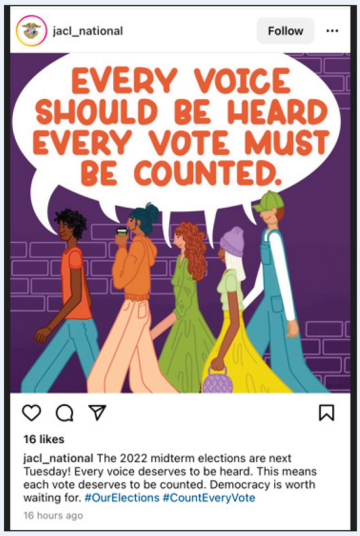
CASE STUDY: Elections Night Is (Still) Not Results Night

In any election, the results reported on election night are unofficial. Currently, no state requires that official results be certified on election night itself. States have different rules on how and when to process ballots, and often it can take time to count military and overseas ballots, as well as validated provisional ballots. At the same time, it is often clear who the winner of an election is on election night, and news organizations have often "projected" a winner based on their analysis of unofficial reporting from elections officials and projections of where outstanding ballots are coming from.

Many of us have grown accustomed to hearing these projections and "knowing" the winner of the election soon after polls close. But as states change their rules on when and how to count ballots (which can delay the time it takes to release unofficial results), and as more voters use vote by mail or have an issue at the ballot box and need to use a provisional ballot, it shouldn't be expected to always have these unofficial results on election night.

Disinformation purveyors have used this expectation to create false narratives that claim that elections are somehow rigged or fraudulent if the unofficial results are not projected on election night or if these unofficial results (or media projections) change as counting—and the certification and verification process—plays out.

In 2020, a broad coalition of voter protection organizations joined forces to communicate this reality and push back on disinformation, highlighting that we should "count every vote" and that "election night isn't results night." This included grassroots communication through partner groups, traditional and social media outreach, media advisories, and much more.



The 2020 election was a unique election environment during the height of public concern around COVID-19. Elections officials and state officials expanded vote by mail and other voting options, and as a result, the use of vote by mail and other nontraditional methods increased dramatically from previous elections. That made the communications push to inoculate against "election night" disinformation more salient in 2020. By the time the November 2022 elections came up, there was a societal push to "return to normal," rollback vote by mail in states, and the reluctance of the media to run the same story twice about "election night isn't results night." Yet we knew to expect disinformation attacks in this narrative and needed to again mobilize to communicate this inoculation message.

We held nearly a dozen briefings with national media outlets to highlight key disinformation narratives to expect. A key part of these briefings was to ensure that journalists with large platforms did not amplify disinformation when publishing related stories and that they understood the importance and impact of election disinformation.

In addition to leveraging media outreach, we mobilized our coalition to push pro-voter messages that defuse disinformation, including the "election night" messages.

Partnerships

As a coalition, we implemented lessons learned from 2020 to build on messaging and make sure it reached more audiences and touched on a wider range of subjects than before, keeping in mind the populations specifically targeted by disinformation.

To reach the most vulnerable voters and have the most impactful interventions on election disinformation, a large nonpartisan pro-voter coalition is necessary. The Election Protection coalition, led by Common Cause and the Lawyers' Committee for Civil Rights Under Law, is the largest nonpartisan voter protection coalition with over 100 national and state partners that has been active for over two decades. While new efforts and coalitions are formed in each cycle, Election Protection maintains a presence throughout the "off years" (and in fact, engages in state, local, and special elections every year). This makes for a robust network of organizations with a shared history and experience working with each other and produces remarkable results, including field efforts in dozens of states (where nonpartisan Election Protection volunteers assist voters at polling places), year-round engagement with elections officials, a nonpartisan voter protection hotline, and our Election Protection Anti-Disinformation Working Group, which Common Cause co-chairs and is the hub of strategy and information sharing between organizations. Our work would be significantly hampered if it was necessary—due to funding, staffing, or organizational decisions—to start a new election anti-disinformation effort from scratch each year.

Partnerships: States

Instead, because of our ongoing engagement with Election Protection and the network of state leaders, we have built-in communications pathways into and between states. Backbone organizations like Common Cause can drive research, analysis, and strategy—including messages and message framing—maintaining coordination with the full coalition of state and national organizations. Leading up to, on, and after Election Day, we regularly updated and convened with people in the field doing voter protection work, journalists covering democracy issues, and state voter protection organization leaders.

In addition to our inoculation content, Common Cause shared coalition-wide updates on how to counter specific disinformation narratives (in coordination with state partners) that we witnessed in real time. Many of these were directed to groups working with limited-English persons, like CASA Pennsylvania, APIA Vote Michigan, Voces de la Frontera (Wisconsin), and our core national partners, Arab American Institute Foundation, NALEO, and APIA Vote.

To use one group's experience, the Arab American Institute Foundation was able to track an emerging disinformation narrative in Michigan and connect with leaders on the ground in California and Texas when it emerged. They noted that "this disinformation effort surprised local groups who were not familiar with this type of disinformation and how to counter-message. We were able to quickly gather key local allies and build their confidence with counter-messaging and reassurance." Because of our partnership, they were able to help other states and local groups anticipate and work to disarm rapidly spreading narratives.

Another group, Voces de la Frontera, which is based in Wisconsin and works primarily in Spanish, was "surprised by how quickly we were able to detect and report fake content on all social media platforms. In the future this will allow us to navigate the web and media in a better way in upcoming elections."

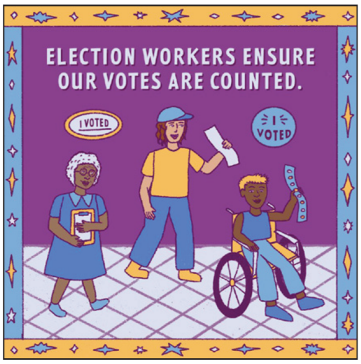
We also worked with partner groups to focus on communicating to voters in their communities with accessible information, inoculation content, and resources. As one subgrantee, Election Protection Arizona/Arizona Democracy Resource Center reported, "I think it made some of our community members feel like they are being heard instead of just being targets in GOTV campaigns."

Partnerships: Election Workers

As one of the largest national and state-based grassroots organizations advancing democracy and transparency, Common Cause Education Fund works on the ground in 30+ states alongside local and state elections officials and administrators. Because we're in constant communication with these government actors, we both inform them of worrisome trends requiring fixes and pass along their information and updates to the public at large.

Common Cause Education Fund works on the ground in 30+ states alongside local and state elections officials and administrators.

Since 2020, there has been a disturbing trend of rising threats and intimidation of election workers. Our anti-disinformation work removes threatening content and provides inoculation messaging that helps "turn down the temperature," but additional support for elections officials is needed. Our sister 501(c)(4) organization, Common Cause, has helped successfully implement strong elections policies and practices across the states—in Colorado, for example, Common Cause helped pass the strongest set of reforms in the country—that not only ensure greater access for voters but additionally protect voters, and administrators, from disinformation, intimidation, and political violence.



For the 2022 midterm elections, we intentionally crafted and encouraged the spread of social media and offline content that contained positive messaging about election workers. We know that even some conspiracy theorists who promulgate false narratives about a "rigged election" believe that their own local elections are secure. In other words, when election workers are humanized as people like us and from our community, they are more likely (though not always) believed to be working to fairly administer elections. Our content and narrative-building work in 2022 helped spread the narrative that elections are run by us, and not by nameless bureaucrats imported from somewhere else.

Our Work with the Media

We helped educate the media on how to accurately and responsibly report on election disinformation and saw noted changes in how information was conveyed to voters.

For the 2022 midterms, our team dedicated itself to engaging an even higher number of journalists than in 2020 to inform them on how to responsibly report disinformation. This is an essential part of the puzzle because, unless done appropriately, writing about disinformation can exacerbate, rather than ameliorate, the problem. As such, last summer and fall, we held individual briefings for key reporters from ProPublica, the *New York Times*, the *Washington Post*, NPR, the Associated Press, Reuters, Bloomberg, and other outlets to provide them with background on our work to combat disinformation. A key part of these briefings was to ensure that these journalists—with large platforms—did *not amplify* disinformation when publishing stories we brought to their attention.

More tactically, we informed journalists of disinformation posts lingering on platforms to force action from social media platforms through highly publicized media pressure. In June, for example, we got posts targeting a specific individual for unproven ballot fraud successfully removed. And by placing stories in Bloomberg and ProPublica about a proliferation of posts targeting election workers, we successfully pressured platforms to both remove the posts and implement new rules.

In 2022, we saw a significant amount of "inoculation" stories from the media—particularly about certification and the voting process. We helped mobilize the Election Protection community to ensure that we were speaking with one voice on inoculation messaging. We didn't want people to forget that results would take time in 2022 as well, so we highlighted the importance for the media to set expectations for election timelines. Inoculation stories from the media included positive stories about democracy in action, such as our nonpartisan voter protection work and the volunteers who make it happen.

Our emphasis on inoculation work in non-English languages also made an impact. In addition to creating content, providing strategy, and sharing resources with our nonprofit partners, as detailed earlier, our ongoing partnership with PolitiFact created **32 articles in Spanish on election issues**, published in three major Spanish-language outlets. This filled the information gap about the election process in Spanish, prevented disinformation from taking root in limited-English-proficiency communities, and provided our Spanish-language partners with key resources to rebut or inoculate against Spanish-language disinformation.

SECTION THREE: LOOKING AHEAD

The year 2024 will present new challenges. As such, we'll both continue the work we've honed over the last few cycles and move into new territory too.

Many Candidates Continue to Endorse Election Denial

Election disinformation is now essentially obligatory for nearly all Republican primary candidates and opportunists seeking financial benefit.

The same polling that demonstrates increased trust in elections still shows the continued impact of disinformation. Voters have more confidence in their local elections than in national elections, and 51% of Republicans "say they think people submit too many ballots in drop boxes either very or somewhat often." This illustrates how even though the gap in trust is decreasing since 2020, election disinformation myths and perceptions of widespread fraud still persist.

Still, "one third of the [Republican] party's 85 candidates for governor, secretary of state and attorney general"—officials who would be responsible for election oversight—"embrac[ed] Trump's efforts to overturn his 2020 loss." Half of them, incumbents in particular, secured seats in 2022. And 220 election skeptics who "cast doubt on the 2020 election," three dozen of whom denied the 2020 results outright, won seats in the U.S. House of Representatives. Despite their positions in power, election deniers still can't provide the evidence they claim they have. In one telling example, Arizona State Senator Wendy Rogers claimed that she couldn't give evidence to agents because "she was waiting to see the 'perp walk' of those who committed fraud during the election." Rogers was able to give further oxygen to conspiracies with this remark while also refusing to elaborate further.

The University of California San Diego's Yankelovich Center finds that there's a partisan gap in trust in elections: "Democrats are more than twice as likely as Republicans (85% versus 39%) to view the results of this [2022] November's election as accurate, while Republicans are more than five times as likely (43% versus 8%) to suspect significant fraud." There's some hope, though: Bright Line Watch finds that "public confidence that votes were counted accurately at the local, state, and national levels increased after the election and beliefs in voter and election fraud decreased. The changes were generally largest among Republicans." Additionally, a Monmouth poll finds that 55% of Republicans surveyed claim that Biden's 2020 win was illegitimate (down from 69% in their last poll). Finally, the Carnegie Endowment for International Peace has a new report detailing why political violence didn't materialize in the wake of the 2022 election: Trump may be unique at mobilizing supporters for it, among other reasons.

Despite these encouraging trends, election disinformation will remain an issue as long as it is lucrative, popular, and profitable for disinformers to promote it.

Tech Platforms Are Backing Down on Civic Integrity

It will be more difficult to rely on tech platforms to act responsibly and enforce their policies.

In dealing with platforms for the 2022 election, we experienced inconsistent applications of policy, conflicting information on violative content, and instances where we were simply ignored. This is all in accordance with a general trend: tech platforms are cutting down on staff dedicated to misinformation. For example, there is just one person left to handle misinformation policy at YouTube, and YouTube recently announced that they will no longer enforce civic integrity policies around 2020 election disinformation. Other platforms have similarly reduced their policy staff. As people continue to seek news from social media, the problem of disinformation will persist and intensify—and dwindling staff will be on hand at major platforms to deal with it.

One example of inconsistent application of policy is how major platforms have treated the restoration of Donald Trump's accounts. Major platforms like YouTube, Meta, and Twitter deciding to restore Donald Trump's accounts shows that either they don't understand that the threat of incitement isn't over, or they've chosen potential profit over people. Meta claims that there are new guardrails to prevent Trump from inciting further violence, such as "heightened penalties" for future violations.

In dealing with platforms for the 2022 election, we experienced inconsistent applications of policy, conflicting information on violative content, and instances where we were simply ignored.

They also say that the risk has "sufficiently receded." It's worth noting in response to that claim that we are still experiencing Trump-incited political violence, whether it's an election denial shooting in New Mexico or an attempt to assassinate the Speaker of the U.S. House. It also appears that Trump is amplifying more and more extremist content on his social media site (while his accounts were restored, he is rarely posting on Meta and has not yet posted on Twitter)—Accountable Tech found more than 350 posts that would violate Facebook's standards.

Meta says that "in the event that Mr. Trump posts further violating content [on Facebook or Instagram], the content will be removed and he will be suspended for between one month and two years." If he posts content that isn't violating, like "content that delegitimizes an upcoming election or is related to QAnon," the spread of the post will be limited. Trump has amplified QAnon accounts more than 400 times on Truth Social, and the election denial movement as a whole is getting closer to QAnon. It's concerning that delegitimizing the 2020 election or posting QAnon content isn't considered violative, given the very real potential for violence.

If the hands-off approach to Donald Trump's accounts is any indication, the Election Protection community will have new challenges to face due to social media platforms' inaction.

The Remaining Threat of Political Violence

There is remaining potential for political violence to flare up in the wake of indictments of a presidential candidate and a primary focused on rehashing lies about 2020.

Violent rhetoric online is still motivating political violence offline: Paul Pelosi's assailant, David DePape, made claims of a stolen election to police after being arrested for his October 2022 assault on the former Speaker's husband. Further review of his online activity shows that he was steeped in conspiracy theories.

Election denial and conspiracy theories were also key motivators in the shootings of Democratic officials and elected lawmakers' houses in New Mexico last year. Losing New Mexico House GOP candidate Solomon Pena orchestrated, and even participated in, shootings at the homes of elected officials he believed rigged his, and other, elections. Pena's campaign website, which is still live, contains alarming rhetoric about the 2020 election, including the claim that the "offenders are not criminal defendants, they are enemy combatants." Pena was also inspired, as *Talking Points Memo* notes, by the work of election deniers like David Clements. Texts from his phone included messages about certification and claims that "they sold us out to the highest bidder," as well as the addresses of the officials targeted. Per the information of a confidential informant, Pena intended the shootings to cause harm and even participated in one himself.

The presidential primaries will begin soon, and one major candidate is facing several potential indictments. In response, Trump has called for his supporters to instigate violence on his behalf as indictments loom. Recent posts by Trump on his social media network, Truth Social, claim that "the Democrats used Covid inspired Mail In Ballots to CHEAT.... Now they are using PROSECUTORS to CHEAT," and "the Democrats are using Prosecutors

for purposes of Election Interference. It is their new way of CHEATING on Elections!" In this way, Trump is trying to connect his claims of a rigged election in 2020 to his new troubles—and trying to incite the same response from his supporters.

There were also attempts by Trump to incite violence against Manhattan District Attorney Alvin Bragg—culminating in Bragg receiving a suspicious envelope with white powder. Trump's rhetoric aimed at Bragg contained claims of Soros ties and featured thinly veiled racist remarks. This type of inciting language will likely continue to be used and amplified by partisan disinformers as the election grows closer and has the potential to inspire further political violence.

Election Disinformation Fuels Voter Suppression

More legislative norms are being eaten away by lawmakers seeking to cultivate support from a base steeped in conspiracy theories—and introducing legislation premised on lies about elections.

It's easier for elected officials to pass restrictive voting laws under the guise of election integrity if voters believe in any number of unfounded conspiracy theories keeping the idea of widespread partisan voter fraud at the forefront.

Election disinformation is thus the "tail that wags the dog" as states pass laws restricting voter access.

Election disinformation is thus the "tail that wags the dog" as states pass laws restricting voter access. Some politicians even fund specifically designed law enforcement units to find "voter fraud," creating a vicious cycle of headlines about arrests for election crimes—despite the fact that most individuals prosecuted were in fact given wrong information by state employees. The goal of these voter intimidation squads is to depress the vote, especially in communities of color. And the idea is catching on, in states ranging from Virginia to Arkansas, which have proposed similar units. The Florida Secretary of State has even proposed an increase in size for its "election crimes unit" from 15 employees to 27, with a corresponding budget increase to $3.15 million.

Other states continue to introduce new legislation to restrict access to mail voting and access to drop boxes. Voting Rights Lab counts hundreds of bills introduced so far this year that reduce access to the vote and criminalize actions of election administrators. The Brennan Center for Justice has counted 150 restrictive voting bills introduced this year, ranging from bills that restrict vote by mail to bills that criminalize errors from election officials. Election conspiracies continue to provide the foundation for further voter restrictions, and even many new decisions made about election administration can be based on the myth of widespread fraud. For example, the majority of the Shasta County, California, Board of Supervisors voted to end their contract with Dominion Voting Systems over fraud claims and disinformation about voting machines: "Just because we're all sitting up here and elected does not mean we had free and fair elections every single time," said one supervisor.

Election-denier politicians are already starting to do what they did in 2020: coordinate to introduce vote-suppressive and anti-administrative legislation across the states. To do so, they're resurfacing old rhetoric about voter fraud and election rigging to push photo ID laws, cut reforms that facilitate voting, and criminalize elections officials' work. Look no further than one proposed bill in Kansas where drop box access would be highly restricted—for fear of "mules." Not only would drop boxes be limited in this proposed bill but also video cameras would record the faces of voters dropping off ballots.

Lawmakers in Kansas cited the debunked documentary *2,000 Mules* as motivation for the bill: "I think part of the concern that's kind of driven bills like this has been partly the whole notion of what are called mules, as far as that somehow somebody's going to stuff a ballot box akin to, you know, there was a documentary called '2,000

Mules' that came out a year ago." And in Nebraska, a legislator who introduced a voter suppression bill didn't endorse a belief in widespread fraud, but said "the perception is—there is. ... And perception is reality." Election disinformation, even when acknowledged by its proponents as false, is used to fuel legislative voter suppression under the guise of protecting elections.

Attacks on the Voting Process

Any institution, tool, or practice, even ones with bipartisan support and buy-in, can become a successful target of disinformation.

Not only are new restrictions proposed almost daily across the country, but existing rules and procedures are newly challenged. In Kansas, a grace period allowing mail-in ballot return of up to three days after an election came under fire, even as state legislators fighting it conceded that widespread fraud isn't real: "I mean, people do question the fraud all the time. Is there fraud? I think actually we're a fairly good state. But we can always make things better." Even when lawmakers acknowledge that there's nothing to the conspiracies they base bad bills on, they still cite the perceptions of election insecurity—that they themselves created—as a reason to advance them.

This isn't solely limited to legislation, either. Election deniers and disinformers are able to take any voting process out of context and portray it as something nefarious to their audiences, and the limit is the disinformers' own creativity. Parts of the process as mundane as what type of pen is used at the polls, how signatures are checked, how ballot tabulators work, and even how long it takes to announce results can and have been targeted for disinformation—and used to erode confidence in our election systems.

Legislative Recommendations:

New legislation points a way forward for grappling with both emerging and existing threats to voters.

As mentioned in Common Cause's 2021 report on election disinformation, there are a number of federal and state laws that already exist to help protect the freedom to vote without intimidation. There are also several legislative proposals that would further aid in the fight against election denialism and help protect voters through their focus on protecting election workers, tackling disinformation in political advertising, and fighting deceptive practices. While no bill will solve every issue we face, there are several bills that will protect access to the vote.

In addition to the bill recently introduced by Senator Amy Klobuchar and Representative Yvette Clarke, which would regulate AI-generated content in political advertising, the Freedom to Vote Act, recently reintroduced, provides a number of solutions for problems of voter intimidation and access. Not only does it increase access to the vote by promoting online registration and allowing for same-day registration, but it also establishes further protections for disabled voters and election workers. The Freedom to Vote Act also includes provisions against deceptive practices, such as prohibiting false statements about federal elections 60 days before an election that would prevent someone from exercising their right to vote.

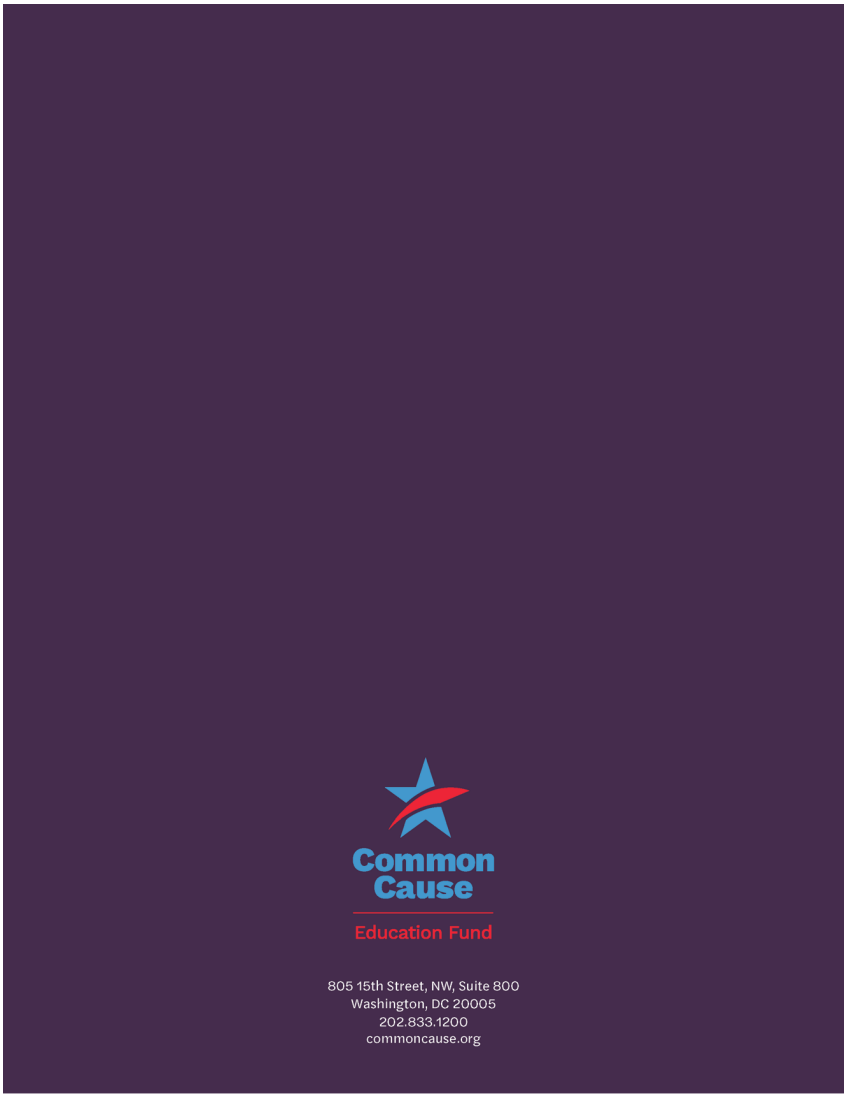
CONCLUSION

Through Common Cause's years of experience tracking, analyzing, and disrupting election disinformation, we've learned how it is made, how it is spread, and how to combat it. Disinformers operate by taking advantage of pre-existing narratives and using social media to amplify them, seeking out audiences who may be more susceptible to disinformation about voting and elections or unable to identify the correct information on the subject.

At Common Cause, we see protecting our democracy, ensuring that voters can vote without barriers to the ballot box, and fighting back against all types of voter suppression as central to our work. That's why our Election Protection efforts take place 365 days a year, not just in the weeks surrounding Election Day. As the threats to voter participation have grown, we have expanded our work to include the new frontier of voter suppression—disinformation, political violence, and election sabotage.

Understanding major disinformation narratives as they arise allows us to better prepare for and respond to attacks on the right to vote. We will no doubt see more acts of political violence, threats, and intimidation fueled by election disinformation in 2023 and 2024. We will publish a memo later this year that outlines the disinformation narratives we anticipate in 2024 and how we intend to combat them.

As these threats to democracy are linked, so is our response. The recommendations made here, if implemented, would have an appreciable impact on the threat of election denialism and disinformation.


Common
Cause
Education Fund
805 15th Street, NW, Suite 800
Washington, DC 20005
202.833.1200
commoncause.org

Statement

of

Jennifer Huddleston

Research Fellow
Cato Institute

before the
Senate Rules Committee
September 27, 2023
"AI and the Future of Our Elections"

Statement for the record regarding AI and Elections

Chair Klobuchar, Ranking Member Fischer, and distinguished members of the Rules Committee,

My name is Jennifer Huddleston, and I am a technology policy research fellow at the Cato Institute. My research focuses primarily on the intersection of law and technology, including issues related to speech and the governance of emerging technologies, such as artificial intelligence (AI). Therefore, I welcome the opportunity to submit a statement regarding the potential need for and effect of a government intervention into requirements around the use of AI in election-related speech.

In this statement for the record, I will focus on two key points:

- AI is a general-purpose technology and overly broad regulatory actions based in fear of potential misuse by bad actors are likely to have unintended consequences that could prevent numerous beneficial and benign uses, as well as the overall development of this technology, including in elections and campaigns.
- Before presuming new law or rules are necessary, policymakers should carefully define the perceived harm they are trying to address and consider if existing regulations may already address that harm. They should also examine if there are cases where outdated interpretations or regulations may be preventing a better technological solution. This is true for the use of AI in election campaigns.

The most basic definition of AI is a computer or robot that can perform tasks associated with human intelligence and discernment. Most of us have been encountering AI much longer than we realize in tools like autocorrect, autocomplete, chatbots, and translation software. While

generative AI has garnered recent attention, the definitions of AI typically found in most policy proposals would impact far more than the use of ChatGPT.

The last few election cycles in the United States have seen rising fears about misinformation, disinformation, and deepfakes. For all the fears around technology and elections, however, we shouldn't presume only the worst-case scenarios. Fears about the potential manipulation of new forms of media and technology have merged, and society has had to evolve their awareness of potential deception or manipulation of that media and redefine what makes information "real" or "true."[1] Some may bemoan the decreasing trust in media more generally, but this distrust existed before AI, and even before the creation of social media.[2] This skepticism and awareness may actually become a positive when it comes to concerns about the use of AI, as it may render AI deepfakes more akin to the annoyance of spam emails and prompt greater scrutiny of certain types of content more generally.[3]

While the internet has increased the popularity and speed with which an individual piece of content can be shared, it also has developed its own norms around understanding the veracity of certain claims. These norms have developed without government dictates and will likely continue to develop in the face of new technologies like AI. As Jeffrey Westling, Director of Technology and Innovation Policy at the American Action Forum, writes, "These societal norms can and will continue to drive trust in video as the viewer will understand that these institutions investigated the claims beyond just what appears on screen. And to the extent that videos become

---

[1] Jeffrey Westling, *Deception & Truth: A Deep Look At Deep Fakes*, 2019, https://www.techdirt.com/2019/02/28/deception-trust-deep-look-deep-fakes/
[2] Megan Brenan, *Americans' Trust In Media Remains Near Record Low*, Gallup, 2022, *https*://news.gallup.com/poll/403166/americans-trust-media-remains-near-record-low.aspx#:~:text=Just%207%25%20of%20Americans%20have,in%20newspapers%2C%20TV%20and%20radio.
[3] Taylor Barkley, *How Much Should We Worry about Deep Fakes?*, Human Progress, 2019, https://humanprogress.org/how-much-should-we-worry-about-deep-fakes/

more consistently faked, society will shift back towards looking at the context behind the video."[4]

With all this considered, we should be cautiously optimistic that history shows a societal ability to adapt to new challenges in understanding the veracity of information put before us and to avoid overly broad rushes to regulate everything but the kitchen sink for fear of what could happen. While the saying "a lie may travel halfway round the world before the truth puts on its shoes" may have some concern with how quickly a fraudulent or manipulated image may spread, there are a variety of non-government responses that often come into play. For example, many online platforms now provide further context around certain types of media, including both election-related speech and manipulated media. These precise rules vary, allowing different platforms to come to different decisions around the same piece of material.[5]

It should be emphasized that not all uses of AI in election advertisements should be presumed to be manipulative or fraudulent. In fact, even when it comes to election advertising, there are beneficial and non-manipulative uses of technologies like AI. For example, AI could be used to translate an existing ad in English to the native language of a group of voters that might not otherwise be reached or add subtitles to reach communities of individuals with disabilities. It could also be used to lower the costs of production and post-production, such as removing a disruption in a shot. Even these examples are more direct interactions that may be more visible than the countless examples of AI that may be used in spell-checking a script or using an algorithm in a search engine to conduct research or promote an ad. These actions are not

---

[4] Jeffrey Westling, *Are Deep Fakes a Shallow Concern? A Critical Analysis of the Likely Societal Reaction to Deep Fakes,* TPRC47, 2019, https://dx.doi.org/10.2139/ssrn.3426174

[5] *See, e.g.* Subramaniam Vincent, *Why Facebook Left up the "Drunk Pelosi: Video but YouTube Took it Down,* Markkula Center for Applied Ethics at Santa Clara University, 2019, https://www.scu.edu/ethics/all-about-ethics/why-facebook-left-up-the-drunk-pelosi-video-but-youtube-took-it-down/

manipulative or deceptive nor do they give rise to concerns about mis or disinformation. However, under many definitions, some or all of these actions would result in labeling requirements that an advertisement used AI. Given the broad use of AI , such a "warning label" could become meaningless as it applies to both benign and manipulative uses. Existing law does not get tossed out the window just by the appearance of new technologies, and actions by bad actors must be addressed in existing FEC rules. New technologies should not change the underlying rules of the road. Similarly, potential regulations must also consider the impact on issues such as speech.

It is important to note that "truth in advertising" laws do not apply to current political advertising.[6] While it may be distasteful to some, current law does not require political ads to be only truthful or factual.[7] Political speech, including that contained in election advertising, is protected by the First Amendment. As a result, there are significant limits on when or how the government can intervene in narrowly tailored ways when responding to a compelling government interest. This does not mean that there is no redress if harm occurs, as existing laws such as defamation can still apply.

As generative AI remains rather new, many societal norms around its use are still evolving; however, already existing standards around election advertising online are incorporating the use of AI into their policies. For example, Google tweaked its policy earlier this month to mandate AI disclosures in political ads.[8] As has been seen with current policies around election advertising, different platforms may reach different specifics on these disclosures, like they have

---

[6] Edgar B. Herwick III, *Why Don't Truth In Advertising Laws Apply To Political Ads?,* GBH, 2019, https://www.wgbh.org/news/politics/2019-11-06/why-dont-truth-in-advertising-laws-apply-to-political-ads

[7] Domenico Montanaro, *The truth in political advertising: 'You're allowed to lie',* NPR, 2022, https://www.npr.org/2022/03/17/1087047638/the-truth-in-political-advertising-youre-allowed-to-lie

[8] Gerrit De Vynck, *Google to require politicians to disclose use of AI in election ads,* Washington Post, 2023, https://t.co/b5yhKWRQh2

with other concerns around election advertising. Gradually, general norms and best practices will evolve and adapt with changing understandings, much like other technologies, in a far quicker way than top-down law could.

As mentioned above, it is important to remember that existing laws did not disappear with the emergence of AI. FEC rules can still apply to actions by AI, and the recourse will remain the same for violations by new technology. Many of the expressed concerns about potential AI generated disinformation or misinformation are not unique to AI but rather a new manifestation of existing concerns. Instead of rushing to create problematic licensure regimes or regulations that are likely to become quickly outdated, agencies and Congress should clearly articulate the harm that they are trying to solve and why it is not addressed by existing regulations. Any such regulations should be narrowly tailored to specific applications — even within an area such as election law — and not broadly applied to all uses of technology. Policymakers should recognize that AI, like other technologies, is ultimately a tool. Like any tool, it can be used for both productive and disruptive purposes; however, many of the potential disruptive concerns are likely already addressed. Rather than rush to create new regulations, policymakers should examine if their concerns are addressed by existing regulations or if they are truly novel concerns. Additionally, particularly around issues such as election advertising, policymakers must also consider the impact regulation could have on already utilized forms of AI as well as important values such as free speech.

Townhall
TIPSHEET

Biden Allies Spread Photoshopped Pictures of the President to 'Prove' He's Fit for Second Term

John Hasson
August 01, 2023 9:00 AM

As a growing number of Americans believe President Biden is too old for a second term in office, White House allies are employing a desperate strategy to convince voters that the President is fit enough to serve:

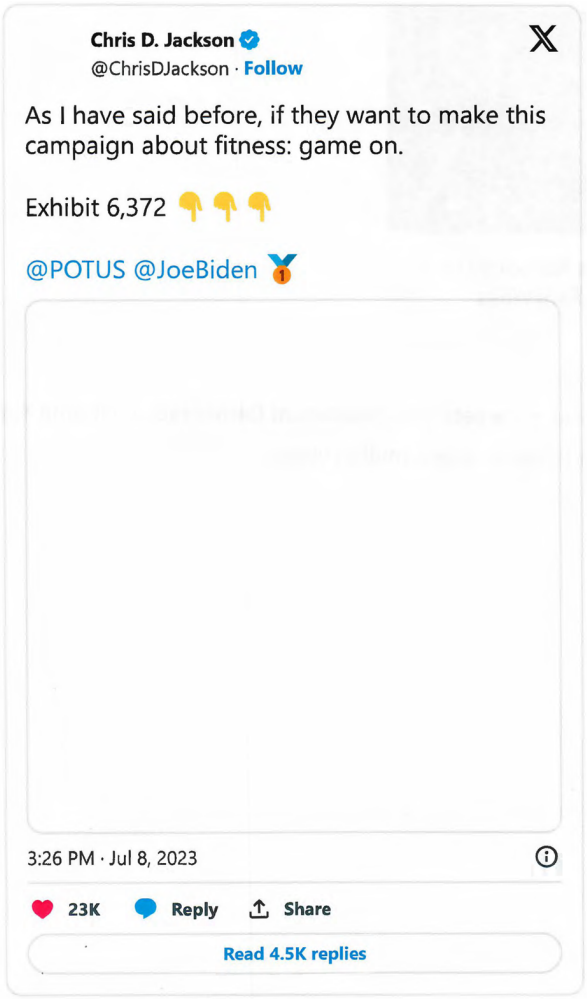
They are photoshopping Biden's pictures.

Their efforts appear to be part of a months-long campaign to sway the 68% of voters who think Biden is too old for a second term.

Since Biden launched his re-election campaign on April 25, White House allies have shared the same four digitally altered photos of Biden repeatedly—without acknowledging that the photos have been edited. One edited photo even retained its original Getty Images watermark, adding to its supposed authenticity.

In each case, Democrat influencers contrast the images with unflattering pictures of Trump, arguing that the comparison proves Biden is not only fit for office, but also more fit than Donald Trump is.

Their posts have been amplified by A-list Democrats, including Malcolm Nance, Jon Cooper, and former White House Chief of Staff Ron Klain, who promised in his resignation letter that he would do whatever he could to help Biden's reelection campaign.



Recommended

But the image in Jackson's tweet was fake.

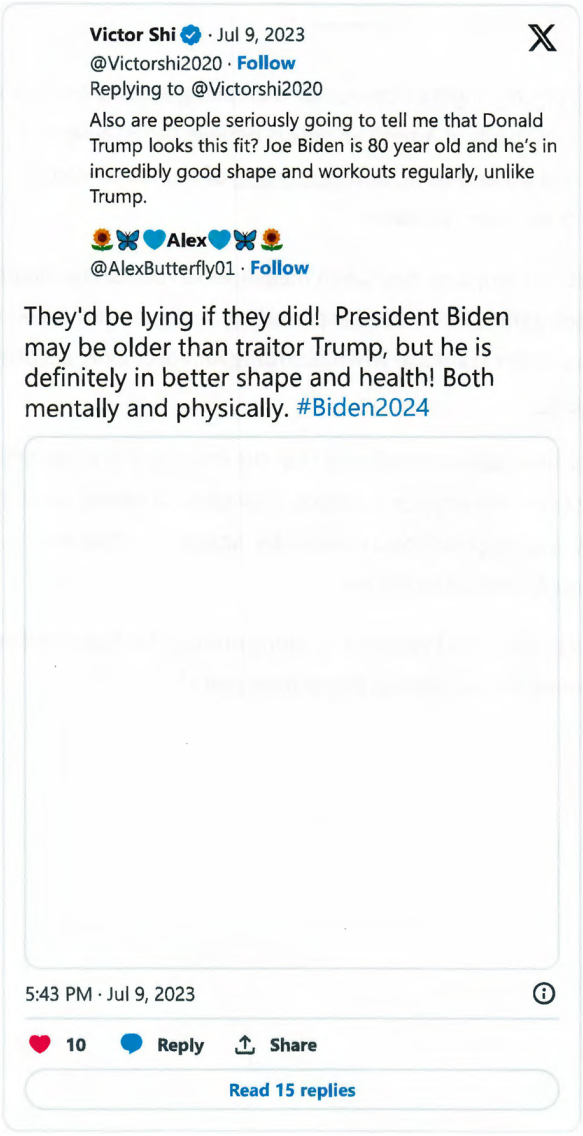
2. An original image, from a video posted by Josh Wingrove vs an edited version, posted by Chris Jackson on May 14.

4. An original image from Susan Walsh, (AP) vs an edited version, posted by Chris Jackson on July 13.

Navarro's image went even more viral than Jackson's did, topping out at 3.6 million views and 57K "likes."

Navarro's post was easily the most successful in generating engagement, but it was hardly unique in content. Multiple Democrat accounts, with followings


Andrew—Author of America Rises Newslet...
@AmoneyResists · Follow
Mentally, there is no person on earth more unfit for office than Donald Trump. Biden could lose more than half of his cognitive abilities and he'd still be infinitely superior to the malignantly narcissistic sociopath with no grip on reality. Physically, there's also no contest.
12:18 AM · May 15, 2023
7.4K
Reply
Share
Read 831 replies



Victor Shi · Jul 9, 2023
@Victorshi2020 · **Follow**
Replying to @Victorshi2020

Also are people seriously going to tell me that Donald Trump looks this fit? Joe Biden is 80 year old and he's in incredibly good shape and workouts regularly, unlike Trump.

🌻🦋💙**Alex**💙🦋🌻
@AlexButterfly01 · **Follow**

They'd be lying if they did! President Biden may be older than traitor Trump, but he is definitely in better shape and health! Both mentally and physically. #Biden2024

5:43 PM · Jul 9, 2023

10 **Reply** **Share**

Read 15 replies

TECHFREEDOM
LAW FOR A DYNAMIC FUTURE

771.200.4997
Washington, DC
techfreedom.org

October 2, 2023

The Honorable Amy Klobuchar
Chair, Committee on Rules & Administration
United States Senate
448 Russell Senate Office Building
Washington, D.C. 20510

The Honorable Deb Fischer
Ranking Member, Committee on Rules & Administration
United States Senate
425 Dirksen Senate Office Building
Washington, D.C. 20510

Re: AI and the Future of our Elections

Dear Chair Klobuchar, Ranking Member Fischer, and members of the committee:

During my testimony on September 27, I referenced a study of political deepfakes finding, among other things, that political deepfakes are no more credible or emotionally manipulative as other forms of deceptively edited videos, and that political knowledge and digital literacy were factors that increased viewer discernment across subgroups. I am attaching that study to this letter, for the record, in hopes that it provides valuable insight into what actions are necessary and effective with respect to safeguarding our elections.

Thank you once more for the opportunity to testify on this important issue, and I remain at your disposal as you carry out this committee's vital role.

Sincerely,

Ari Cohn
Free Speech Counsel
TechFreedom

Encl.

Political Deepfakes Are As Credible As Other Fake Media And (Sometimes) Real Media*

Soubhik Barari† Christopher Lucas‡ Kevin Munger§

First Draft: January 13, 2021
This Draft: April 16, 2021

Abstract

We demonstrate that fabricated videos of public officials synthesized by deep learning ("deepfakes") are credible to a large portion of the American public – up to 50% of a representative sample of 5,750 subjects – however no more than equivalent misinformation in extant modalities like text headlines or audio recordings. Moreover, there are no meaningful heterogeneities in these credibility perceptions nor greater affective responses relative to other mediums across subgroups. However, when asked to discern real videos from deepfakes, partisanship explains a large gap in viewers' detection accuracy, but only for real videos, not deepfakes. Brief informational messages or accuracy primes only sometimes (and somewhat) attenuate deepfakes' effects. Above all else, broader literacy in politics and digital technology increases discernment between deepfakes and authentic videos of political elites. Our findings come from two experiments testing exposure to a novel collection of deepfakes created in collaboration with tech industry partners.

*For excellent research assistance, we thank Jordan Duffin Wong. We thank the Wiedenbaum Center at Washington University in St. Louis for generously funding this experiment. For helpful comments, we thank the Political Data Science Lab and the Junior Faculty Reading Group at Washington University in St. Louis; the Imai Research Group; the Enos Research Design Happy Hour; the American Politics Research Workshop at Harvard University; the Harvard Experiments Working Group; and Jacob Brown, Andy Guess, Connor Huff, Yphtach Lelkes, Jacob Montgomery, and Steven Webster for helpful comments. We thank Hany Farid for sharing video clips used in this project. We are especially grateful to Sid Gandhi, Rashi Ranka, and the entire Deepfakeblue team for their collaboration on the production of videos used in this project. All replication data and code is publicly available here.

†Ph.D. Candidate, Harvard University; URL: soubhikbarari.org, Email: sbarari@g.harvard.edu

‡Assistant Professor, Washington University in St. Louis; URL: christopherlucas.org, Email: christopher.lucas@wustl.edu

§Assistant Professor, Pennsylvania State University; URL: kevinmunger.com, Email: kmm7999@psu.edu

1 Introduction

Studies of democratic politics have long emphasized the importance of a well-informed electorate for bolstering democratic accountability (Lippmann, 1922; Berelson, Lazarsfeld and McPhee, 1954; Downs, 1957; Snyder Jr and Strömberg, 2010; Herman and Chomsky, 2010). Information allows voters to accurately judge attributes of electoral candidates such as leadership, expertise, competence, character, and values in order to make principled decisions at the ballot-box (Popkin, 1991; Pierce, 1993; Alexander and Andersen, 1993; Alvarez, 1998; Strömberg, 2004; Caprara et al., 2006). Political misinformation, then, threatens the electorate's ability to credibly evaluate public officials and elect competent leaders (Carpini and Keeter, 1996; Kuklinski et al., 2000; Hollyer, Rosendorff and Vreeland, 2019; Aral and Eckles, 2019; Jerit and Zhao, 2020).

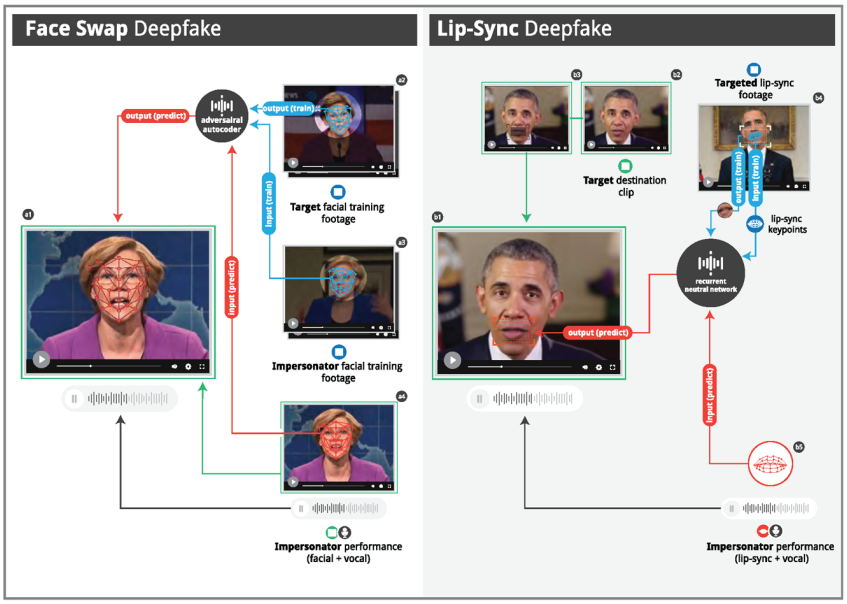
Societal concerns about misinformation have recently centered on novel deep learning technologies capable of synthesizing realistic videos of politicians making statements that they never said, colloquially termed *deepfakes*. Unlike previously available video manipulation tools, contemporary deepfake tools are open source, and thereby unlicensed, unregulated, and can be harnessed by hobbyists (rather than visual effect specialists) with relatively basic computational skills and resources (Government Accountability Office: Science and Analytics, 2020). Figure 1 graphically summarises the two major technologies for the production of deepfakes, which, by many counts, are responsible for the production of the vast majority of political deepfakes at the time of writing (Lewis, 2018; Davis, 2020; Ajder et al., 2019). Since the advent of open-source deepfake technologies[1], political elites around the world have been targeted in deepfake video scandals. Notable examples include a 2018 deepfake of Gabon president Ali Bongo, which triggered a coup attempt, and a viral deepfake reporting on a sex scandal involving a Malaysian cabinet minister (Harwell, 2019).

Consequently, legislators (Gazis and Becket, 2019; Brown, 2019; Lum, 2019; Galston, 2020), news outlets (Harwell, 2019; Parkin, 2019; Frum, 2020; Hwang and Watts, 2020; Schick, 2020; Toews, 2020) and civil society groups (Lewis, 2018; Davis, 2020; Ajder et al., 2019; Bateman, 2020) have all emphasized the potential harm that deepfakes may cause to democracy, and legislation exists in more than a dozen states to regulate the production and dissemination of deepfake videos (Prochaska, Grass and West, 2020). Philosophers claim that deepfake technology removes our ability to verify testimony, undercutting the credibility of both honest and dishonest claims (Rini, 2020).

This article evaluates whether or not these concerns are warranted by answering a series of fundamental research questions. First, are deepfake videos of salient public officials more credible (i.e. not appearing fake or doctored) than equivalent information faked in existing media modalities such as textual headlines or audio recordings? We denote this question as

[1]According to many reports (Lewis, 2018; Davis, 2020; Ajder et al., 2019), the earliest deep learning face-swap tool to receive popular press was `Face2Face` in 2016; see also Suwajanakorn, Seitz and Kemelmacher-Shlizerman (2017) in 2017, `FakeApp` in 2018, `Faceswap` and `DeepFaceLab` in 2019.

Figure 1: **How Deepfake Videos are Generated**



Notes: Shown are two major methods of producing deepfakes. The left illustrates the production of a *face-swap deepfake* which requires: a full clip featuring the impersonator's performance including the audio (**black**) and the background context for the clip (**green**) where the facial features are swapped (**red**) via a trained (**blue**) deep learning model called an autoencoder. The right illustrates a *lip-sync deepfake* which requires a destination clip of the target (**green**) and a vocal impersonator's performance including their audio (**black**) and lip sync keypoints (**red**); these keypoints are transferred into a matching synthetic lip-sync video of the target via a deep convolutional neural network model trained on the target (**blue**).

Research Question 1 — **RQ1** — throughout the text. Second, are deepfakes differentially more credible than other modalities to certain subgroups (**RQ2**)? Third, are deepfake videos more or less credible than authentic videos of political elites (**RQ3**)? Although the scope of possible deepfakes, political or non-political, is vast, our questions chiefly concern deepfake *scandal* videos of political elites, given their attention in contemporary society and imminent election regulations. Before we conducted any research, we pre-registered our hypotheses for these questions, drawing on media effects studies across several disciplines, which we briefly summarise.

On **RQ1**, we expected that deepfake videos would be deceptively perceived as more credible than equivalent information in text or audio formats.[2] In addition to the motivating concerns

[2]In this article, we elicit direct credibility evaluations of media upon exposure, rather than asking if the

of contemporary political observers, this prediction squares with a literature that documents how audiovisual information is the prima facie format for persuasive communication on a variety of topics, including emotionally charged or traumatic events (Christianson and Loftus, 1987; Kassin and Garfield, 1991), courtroom testimony (Kassin and Garfield, 1991), presidential election campaigns (Grabe and Bucy, 2009), support for medical aid (Yadav et al., 2011), and belief in climate change (Goldberg et al., 2019).

On **RQ2**, we pre-registered a number of "at-risk" subgroups defined by characteristics previously shown to exacerbate or reduce susceptibility to misinformation in other contexts (Table 1). The first category draws on the observation that older adults, Guess, Nagler and Tucker (2019) report that "users over 65 shared nearly 7 times as many articles from fake news domains as the youngest age group" during the 2016 US Presidential election. Next, we believe that Bayesian reasoning, or the selective acceptance of information based on consistency with prior beliefs, may powerfully shape how at least two groups respond to deepfakes. Thus, we hypothesize that *partisan group identity* and *sexist attitudes* will influence credibility perceptions of a deepfake video. A large literature documents how partisan identity – either by way of strong directional motivations to reject new evidence or differing priors about the credibility of new evidence – directs voters' attitudes about events, issues, and candidates (Kahan, 2012; Druckman and McGrath, 2019; Leeper and Slothuus, 2014; Enders and Smallpage, 2019). Moreover, voters' evaluations of candidates or events can be driven by prior negative stereotypes towards *sexist attitudes* towards women (Jamieson, Hall et al., 1995; Teele, Kalla and Rosenbluth, 2017; Cassese and Holman, 2019). A recent survey finds that, next to partisanship, holding ambivalent sexist views[3] most strongly predicted support for Donald Trump in the 2016 election (Schaffner, MacWilliams and Nteta, 2018). Another set of subgroups may be especially susceptible to deepfakes due to constraints on cognitive resources or knowledge. Performance in cognitive reflection tasks measures reliance on "gut" intuition which may preclude careful examination of video evidence (Pennycook and Rand, 2019; Pennycook et al., 2019). Similarly, those with little political knowledge may have little prior exposure to the targetted political figure, rendering them unable to discern "uncanny" deepfake artifacts that resemble, but do not perfectly replicate their intended facial features (Mori, MacDorman and Kageki, 2012; Brenton et al., 2005). Finally, the last two categories describe traits that we can intervene on via direct information provision – or raising the salience of deepfakes conceptually or by example – and accuracy priming – or raising the salience or normative value of engaging with accurate news – each of which we expect to reduce deepfakes' credibility (Pennycook et al., 2020, 2019). Echoing our expectations for **RQ1**, we pre-registered the prediction that all subgroups in Table 1 would be differentially

depicted events truly occurred which may be evaluated on their perceived plausibility independent of the information presented.

[3]Ambivalent sexism describes a bundle of both outright hostile (e.g. "women are physically inferior to men") and deceptively benevolent views about women (e.g. "women are objects of desire") (Glick and Fiske, 1996)

Table 1: **Subgroups Hypothesized to Percieve Deepfakes As Credible**

	Subgroup	Mechanism(s) of Credibility
Non-Intervenable in Survey	Older adults ($\geq$ 65 y.o.)	Inability to evaluate accuracy of digital information (Guess, Nagler and Tucker, 2019; Barbera, 2018; Osmundsen et al., 2020)
	Partisans (with out-partisan target)	• Directional motivated reasoning about out-partisans (Kahan, 2012; Leeper and Slothuus, 2014; Enders and Smallpage, 2019) • Accuracy motivated reasoning about out-partisans (Druckman and McGrath, 2019; Tappin, Pennycook and Rand, 2020)
	Sexists (with female target)	• Consistency with prior hostile beliefs about women (Glick and Fiske, 1996; Schaffner, MacWilliams and Nteta, 2018; Cassese and Holman, 2019) • Consistency with prior benevolent beliefs about women (Glick and Fiske, 1996; Schaffner, MacWilliams and Nteta, 2018; Cassese and Holman, 2019)
	Low cognitive reflection	Overreliance on intuition over analytical thinking in making judgments (Pennycook and Rand, 2019; Pennycook et al., 2019)
	Low political knowledge	• Inability to evaluate plausibility of political events • Inability to recognize real facial features of target (Brenton et al., 2005; Mori, MacDorman and Kageki, 2012; Lupia, 2016; Tucker et al., 2018)
	Low digital literacy	• Inability to evaluate accuracy of digital information • Limited/no recognition of deepfake technology (Guess et al., 2020; Munger et al., 2020)
Intervenable	Low accuracy salience	Limited/no attention to factual accuracy of media (Pennycook et al., 2020, 2019)
	Uninformed about deepfakes	Limited/no recognition of deepfake technology

Notes: This list is neither exhaustive nor mutually exclusive. We clarify possible mechanisms for each groups' susceptibility, but proving these and not alternative mechanisms is beyond the scope of this paper.

susceptible to deepfake misinformation over text and audio misinformation.

Lastly, on **RQ3** – as with **RQ1** – if popular claims about deepfakes are correct, they should be nearly indistinguishable from authentic video clips in a shared context (e.g. a news feed

about politics). Thus, we expected that deepfakes should be perceived as equally credible as authentic video clips in the same context.

To test our hypotheses, we employed two experiments embedded in a survey fielded to a nationally representative sample of 5,750 respondents on the Lucid survey research platform.[4] The first experiment (incidental exposure) presents respondents with a news feed of apparently authentic video clips, audio clips, and text headlines about candidates in the 2020 Democratic presidential primary, in which a deepfake video of one of the candidates may or may not be embedded. The second experiment (detection task) asks the same respondents to scroll through a feed of eight news videos – randomized to contain either no deepfakes (dubbed the no-fake feed), two deepfakes (low-fake), or six deepfakes (high-fake) – and discern deepfakes from the authentic video clips. Table 2 describes our overall design and Appendix Figure A6 provides a graphical illustration of the survey flow.

2 Results

Figure 2-5 summarise our main results which robustly reject our hypotheses for **RQ1** and **RQ2**, but shed a nuanced light on **RQ3**. Figure 2 compares baseline and relative subgroup credibility evaluations across all clips in the incidental exposure news feed[5] alongside baseline and relative subgroup affect towards Warren from respondents in all the Warren clip conditions Figure 3 compares performance in the detection task across environments and subgroups, while Figure 4 and Figure 5 break down performance differences by our pre-registered subgroup traits and by clips respectively. We organize our results into three main findings, each of which we discuss in detail in relation to our original hypotheses, and conclude with a brief discussion of external validity.

For all of results involving multiple group-wise comparisons or estimating multiple substantive coefficients, we adjust p-values according to the Benjamini-Hochberg "step-up" procedure which bounds each group of tests' false discovery rate at $\alpha = 0.05$ without as strict of a correction as the Bonferroni procedure which assumes no dependence between hypotheses (Benjamini and Hochberg, 1995). Additionally, where we find unexpected null results, we conduct equivalence tests to verify that the estimated null effects are meaningfully null in magnitude (Wellek, 2010). For consistency, we deem an effect "meaningfully null" if it fails to explain half of a standard deviation or more of the outcome, i.e. falls within the equivalence bounds of $\pm 0.5\sigma$, although we show our null effects are null according to other reasonable interpretations as well.

[4] At the time of fielding, Aronow et al. (2020) noted systematic trends in inattentive survey respondents on Lucid. We describe the battery of attention checks we employ to maintain a high-quality sample in Appendix E. All findings are consistent across samples, though slightly smaller in magnitude for less attentive respondents.

[5] Excluded are credibility comparisons with the satirical SNL skit clipping of Bernie Sanders following the randomized stimuli due to the same differential item functioning concerns expressed in the previous section.

2.1 Deepfake scandal videos are no more credible or emotionally appealing than comparable fake media

In the incidental exposure experiment, just under half of subjects (42%) found our deepfake videos of Warren at least somewhat credible (top left of Figure 2). However, the videos were, on average, less credible than the faked audio (44%) and comparable in credibility to the fake text (42%). Both the fake audio and video clippings not only fail to reject a traditional null hypothesis of no effect relative to the fake text headline, but also reject the null hypothesis of a minimal change of $\pm 0.5\sigma$ (≈ 0.64) in credibility confidence, let alone a full point step between confidence categories. Appendix Tables F6 and F7 show that these differences are robust to a variety of model-based adjustments. Our best answer to **RQ1** is, thus, "no".

Even if deepfakes are not more credible than comparable fake media, are they more emotionally appealing? Relative to no exposure, videos do slightly increase negative affect towards Elizabeth Warren as measured by the 0-100 feeling thermometer, though this still fails to clear our equivalence bounds for a null effect. However, there are demonstrably null effects of the deepfake video on affect when compared to `text` and `audio`, as seen in the top-right cell in Figure 2. Deepfake videos are also at least as affectively triggering as negative attack advertisements, a decades-old technology, of the same target. Appendix Table F10 produces this same null effect with model-based controls.

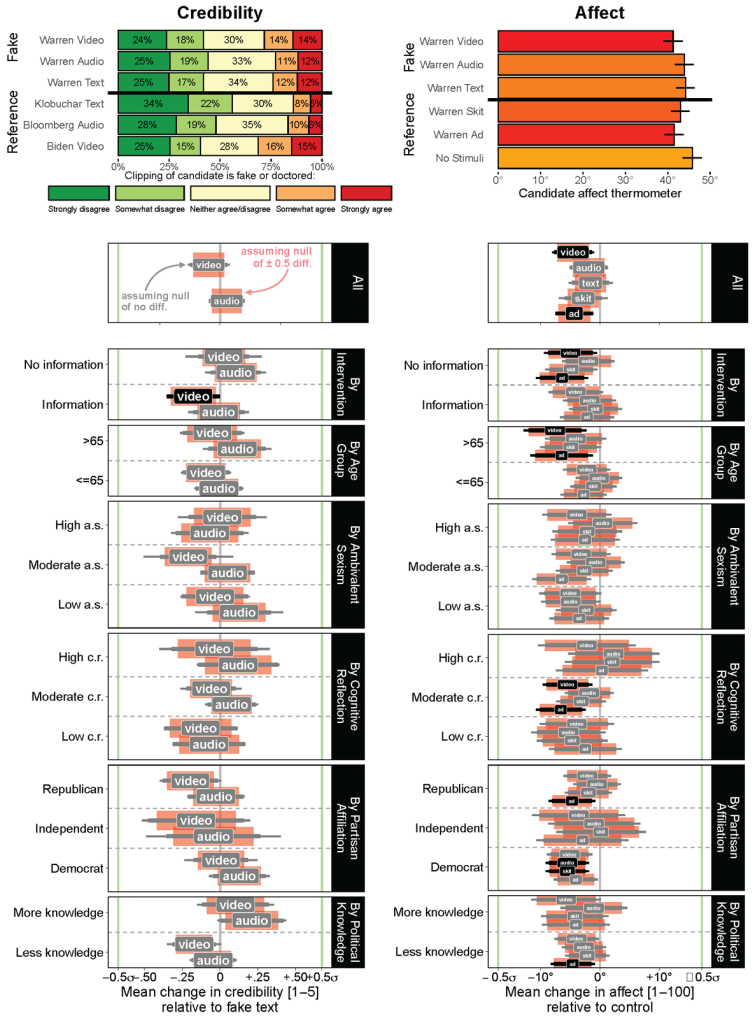
Investigating whether the previous null results mask any credibility or affect heterogeneities for subgroups specified in Table 1 (panels 2–7 in Figure 2), we find few. Verbal information, as intended, provides a targeted, but small reduction of the `video` stimuli's credibility relative to the `text` stimuli, which is partially robust to different model-based adjustments (Appendix Tables F8 and F9). The answer we give to **RQ2** is then also "no", however with one caveat. Sexist attitudes and out-partisan identification predict increases in the credibility (substantively large in the latter case) of the scandal deepfake (Appendix Figure G25, Tables F17–F18, Tables F20–F21), just not relative to the same information in text or audio format.

2.2 Digital literacy and political knowledge improve discernment more than information

Our respondent groups' baseline performance accuracy (Figure 3) in the detection task (53–61% across all groups) and error rates of less than 50% suggest that their discernment capabilities are better than random. Though, notably, the false negative rate for our clips is consistently larger than the false positive rate, meaning deepfakes appear to be slightly more credible than authentic videos (more likely to be misclassified as authentic than vice versa). A little more than a third of all deepfakes in our feed are undetected while a little under a third of authentic clips are falsely flagged across all subgroups.

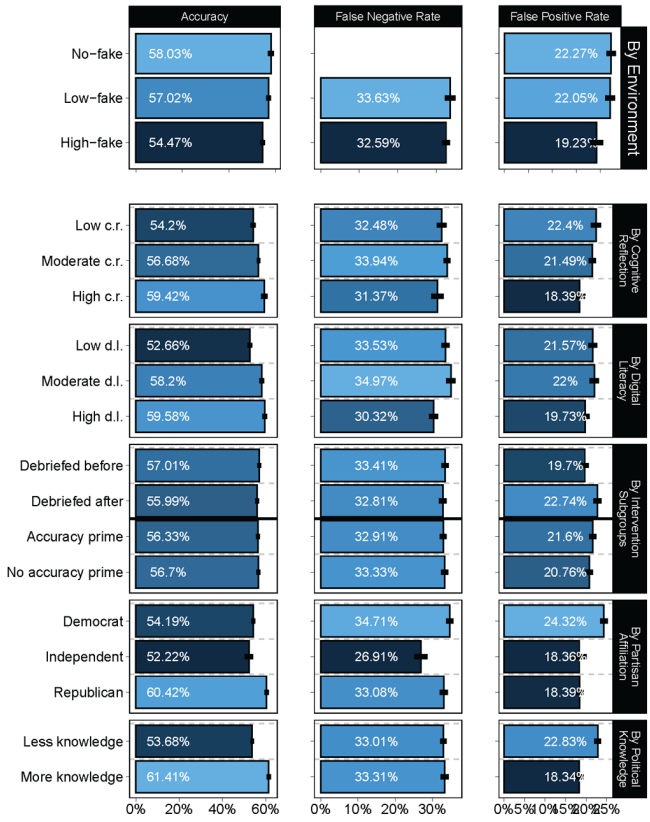
Examining whether Table 1 traits explain performances, we find that neither of our interventions improves discernment accuracy during the detection task (see estimated marginal effects on accuracy in Figure 4). While information and accuracy salience fail, Figure 4 shows

Figure 2: **Minimal Overall and Subgroup Effects of Incidental Exposure to a Deepfake Video**



Notes: Categories for ambivalent sexism and cognitive reflection constructed as equal-sized percentiles from sample values. Thicker lines denote 95% confidence intervals (CIs), thinner lines denote 95% CIs adjusted for multiple subgroup comparisons (Benjamini and Hochberg, 1995), **red** blocks indicate 95% CIs from two one-sided equivalence t-tests with equivalence bounds indicated by the vertical **red** lines (Wellek, 2010). For brevity, the `text` condition is not shown in the subgroup results for affect, however in no subgroup condition did `text` produce a significant effect relative to `control`.

Figure 3: **Performance Comparisons in Deepfake Detection Task**

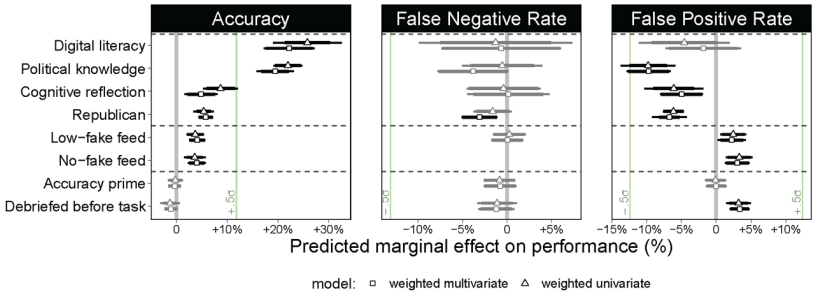


Notes: Shown are three different measures for n=5,497 (99%) of respondents who provide a response to at least one video in the detection experiment task. Coefficient estimates are given in Appendix F and are robust to the choice of missing-ness threshold. Accuracy is the % of all videos in the task correctly classified as either fake or real. False negative rate is the % of deepfakes in the task incorrectly classified as authentic (as such, this quantity is degenerate in the no-fake condition). False positive rate is the % of authentic videos in the task incorrectly classified as deepfakes.

that respondent traits – specifically digital literacy, political knowledge and, to a lesser extent, cognitive reflection – predict the most substantively meaningful improvements according to our standards of evaluation. A unit increase in digital literacy predicts that a subject will provide 15% fewer false positives, a reduction of roughly half a standard deviation in the sam-

ple. Republicans also appear to marginally outperform Democrats and Independents, though scoring little less than a full clip higher in correct classifications than the rest.

Figure 4: **Predictors of Detection Task Performance**



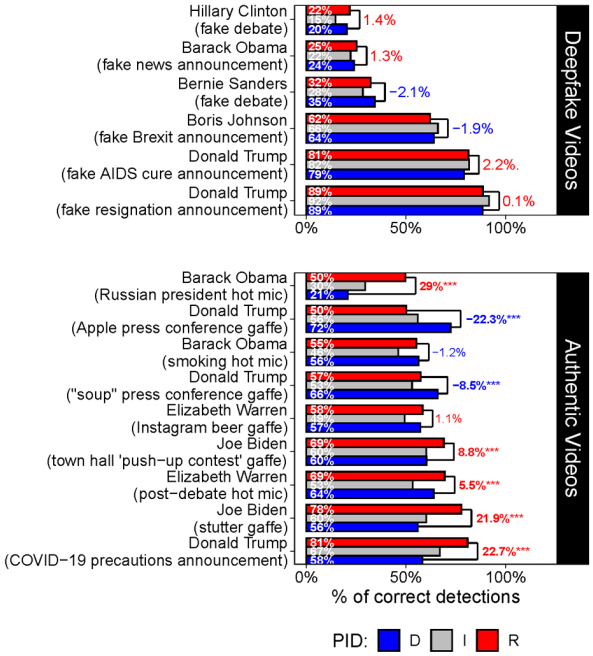
Notes: Predictors are grouped by dashed grey lines into respondent traits (all re-scaled to the [0,1] range), detection environment (relative to high-fake), and intervention assignment. Predictors include all group indicators from Table 1 excluding age which has no significant effects on performance (see Appendix Tables F23, F24, F24). The multivariate model estimates the effects of all predictors jointly and additionally controls for age group, education, and internet usage. Both models are weighted via a post-stratification model (see Appendix Section E). Appendix Figure F24 shows that the cut-off for non-responses in the task does not change the substantive interpretation of detection experiment results.

2.3 Skepticism of authentic videos, but not deepfakes, varies significantly by partisanship

Remarkably, although partisanship overall predicts small effects on performance relative to other traits, an examination of individual clips (Figure 5) reveals some massive performance gaps between Democrats and Republicans, but *only* for real videos. 58% of Republicans believed that real leaked footage of Obama caught insinuating a post-election deal with the Russian president was authentic compared to 22% of Democrats, a highly significant differential according to a simple Chi-squared test ($\chi^2 = 333.34, p < 0.01$). Performance is flipped for the clip of Donald Trump's public misnaming of Apple CEO Tim Cook which was correctly identified by 87% of Democrats, but only 51% of Republicans ($\chi^2 = 75.15, p < 0.01$). Most striking is that for an authentic clip from a presidential address of Trump urging Americans to take cautions around the COVID-19 pandemic, the finding holds in the opposite direction: although a positive portrayal[6], at least for Democrats who by and large hold similarly cautionary attitudes towards COVID-19 (Clinton et al., 2020), only 60% of Democratic viewers flagged it as authentic whereas fully 81% of Republicans believed it to be real ($\chi^2 = 169.96, p < 0.01$). Controlling for both clip and respondent characteristics, Appendix Figure G31 shows that Re-

[6]Positive portrayal, here, means depiction of valence traits or characteristics that, all else equal, voters should unanimously prefer more of rather than less of (Bartels, 2002).

Figure 5: **Detection Performance Comparisons Across Partisanship and Clip Authenticity**



Notes: p-values for differences in correct detection proportions between Democrats and Republicans (*** indicates $p' < 0.001$) derived from two-proportions z-tests where p' is the transformed larger p-value after adjusting for multiple comparisons via (Benjamini and Hochberg, 1995).

publican identity only predicts a boost in performance when asked to corroborate real scandal video clippings of Obama. Thus, individual clips' performance suggests that partisans fare much worse in correctly identifying real clips, but not deepfakes, portraying their own party's elites in a scandal. In contrast, digital literacy, political knowledge, and cognitive reflection bolster correct detections roughly evenly for all clips (Appendix Figure G30).

Taken together with the previous finding, this provides a nuanced answer to **RQ3**. Baseline discernment accuracy is not particularly high for any subgroup, however performance varies significantly by subgroup. Literacy (both political and technological) reduces false skepticism, while partisanship increases skepticism about real scandal videos of in-party elites.

3 Discussion

We have demonstrated that deepfakes, even when designed specifically to defame a prominent politician, are not uniquely credible or emotionally manipulative: they are no more effective than the same misinformation presented as text or audio or the same target attacked via a campaign ad or mocked in a satirical skit. Partisan group identity appears to harm discernment performance, but for authentic videos more than for deepfakes. Our experiments reveal that several cognitive characteristics are essential components of how citizens process both authentic and fake video media. In particular, the respondents with the highest levels of general knowledge about politics, literacy in digital technology, and propensity for cognitive reflection performed the best in the detection experiment. Our study contributes to a growing body of research on manipulated video media (Vaccari and Chadwick, 2020; Wittenberg et al., 2020; Dobber et al., 2020; Ternovski, Kalla and Aronow, 2021), which, taken together, cast doubt on the fear that deepfakes themselves will directly deceive the public of false events on a mass scale. In light of this, policy-makers should devote more time and resources to bolster the credibility of real news videos and curb the development and spread of deepfake videos that perpetrate psychological or social damages against their targets. A recent count of deepfakes on the Internet finds that most are non-consensual pornographic clips of women, suggesting that perhaps the greater, more novel harm of deepfakes is the harassment of its targets, not the deception of its viewers (Harris, 2018).

3.1 External Validity

We conclude with a discussion of four external validity considerations about our results. First, it is possible that deepfakes of other, less salient elites may produce larger effects relative to text or audio than the ones seen here. However, thusfar, deepfakes of this kind (at least accessible to the public) have been exceedingly rare, possibly for technical limitations: as we describe in Appendix B, deepfakes require a large training set of high-definition facial images, which may be unavailable for a city councilor or a low-profile Congressman. Hence, we believe our effects are representative of the kind likely to be seen in the present population of deepfakes, though research on 'downballot deepfakes' would be valuable.

In the detection task experiment, we selected a diverse set of publicly accessible deepfakes of elites that vary in source logos, elite targets, input technology and setting. However, we cannot control for all idiosyncratic features of each clip (e.g. title, tone) nor show that either our deepfakes or authentic clips are exactly representative of these features in the news environment. One thing we can consider is how our results might differ if elites in our detection task were shown in proportion to how often they were actually involved in scandals. For example, according to journalistic (Leonhardt and Thompson, 2017; Quealy, 2021) and scholarly (Bode et al., 2020) accounts of President Trump's behavior, it is possible that news consumers during this period would encounter many more authentic scandal videos of Trump than of other elites. A back-of-the-envelope correction of this under-representation

by simply double-counting each subjects' response to each Trump clip complete nullifies the aggregate effect of Republican partisanship on accuracy shown in Figure 4. More realistically, Republicans and Democrats disproportionately encounter favorable media coverage of their party's elites to begin with; if all Democrats' and Republicans' false positive rates are re-graded by dropping non-congenial clips in their detection task, Democrats reduce false positives from 24.3% to 13.7%, while Republicans improve from 18% to 17.8%. Thus, both the supply of real scandals and selective exposure in real-world news-seeking (to the extent it occurs) could attenuate the false skepticism we have observed.

In this study, we elicited credibility perceptions of clippings ("is this clip real?"), which may be distinct from belief in the occurrence of the depicted event ("did X happen?"). In theory, someone could flag a video as a deepfake, yet believe that the event still occurred. However, manipulation checks on two clips in our detection task suggest that respondents who believe the video is fake generally believe the event did not occur and vice versa (Figure G32 and Figure G33 in Appendix G). Exploring the theoretical and empirical distinctions between these outcomes is a research agenda of its own.

Finally, we recognize that deepfake technology will continue to improve beyond the scope of this experiment. Although we have faithfully replicated the deepfake production process using the best available technology at the time of fielding, readers may live in a world where open-source deepfake technology is capable of generating photorealistic deepfakes completely indistinguishable from authentic videos. In this case, reactions to deepfakes may more closely resemble the responses to real videos we have seen here, where cognitive effort and literacy still improve discernment, while partisanship still continues to drive false beliefs depending on what is shown. Thus, while we encourage technological solutions to constrain the spread of manipulated video, there will never be a substitute for an informed, digitally literate, and reflective public for the practice of democracy.

4 Materials and Methods

Our design is motivated by a number of considerations. Firstly, the two experiments capture different quantities of interest by way of comparing different types of randomized media exposure. The incidental exposure experiment measures the perceived credibility of a single, carefully masked deepfake video relative to the equivalent scandal depicted via other formats, or similar reference stimuli about the candidate in question (**RQ1**, **RQ2**). In the incidental exposure experiment, we also compare affect toward the politicians in each clip as an auxiliary outcome. In contrast, the detection task captures the credibility of deepfakes relative to authentic videos (**RQ3**) measured by overall discernment accuracy and errors due to false positives.

Second, the experiments both inherently and by their ordering allow us to test credibility perceptions across differing levels of information provision. The first experiment simulates exposure to a deepfake "in the wild" with, at most, *verbal description* about deepfakes for those

Table 2: **Overview of Experiments Embedded in Survey**

	Exposure(s)	Pre-Exposure Interventions	Outcomes
(1) Incidental Exposure	1. Pre-exposure authentic coverage of 2020 Democratic Primary candidates 2. Randomized exposure to `text`, `audio`, `video`, `skit` clip of Elizabeth Warren scandal, attack `ad`, or `control` (no stimuli) 3. Post-exposure authentic coverage of 2020 Democratic Primary candidates	• Information about deepfakes	• Belief that candidate clippings are not fake/doctored (credibility) • Favorability of candidates (affect)
(2) Detection Task	Randomized task environment: • No-fake feed: eight authentic clips of political elites • Low-fake feed: six authentic clips, two deepfakes of political elites • High-fake feed: two authentic clips, six deepfakes of political elites	• Debrief of deepfakes exposed to in (1) before task • Accuracy prime	• Deepfake detection accuracy • Deepfake false positive rate • Deepfake false negative rate

randomized to receive information. All participants in the detection task, on the other hand, are explicitly told about deepfakes and some are even provided *visual examples* of deepfakes if randomly assigned to be debriefed about their incidental exposure before the task.

Third, and arguably most important for external validity, our two experiments allow us to test credibility perceptions across multiple deepfakes that differ in their targets, quality, and technology. In the first experiment, as we will describe in the next section, we hired a professional firm to produce several novel deepfakes of a single politician depicted in several realistic scandals via *face swap*. In the second experiment, we used a representative set of pre-existing deepfakes of many different elites made by experts and amateurs alike never existed before made via *lip sync* and *face swap*. Although, in either experiment, we cannot causally attribute credibility due to particular aspects of each video, we draw our overall conclusions using a realistic, externally valid set of deepfake stimuli.

To adjust for observable demographic skews in our respondent pool, all analyses are replicated using post-stratification weights estimated from the U.S. Census in Appendix F. Details of this post-stratification and other characteristics of the sample are given in Appendix E.

4.1 Incidental exposure experiment

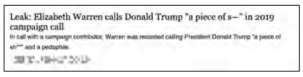
In the first experiment, we implement a 2 x 6 factorial design pairing a randomized informational message about deepfakes with randomization into one of six conditions – a deepfake `video` (presented as a leaked mobile phone recording), or alternatively `audio`, `text`, or `skit` of a political scandal involving a 2020 Democratic primary candidate Elizabeth Warren, a campaign attack `ad` against Warren, or a `control` condition of no clip at all – after which we measure several outcomes. In the incidental exposure experiment, we selected Elizabeth Warren because she was both a salient politician during the primary election, and (at the time of fielding) had not been the target of any visible deepfake online. Thus credibility perceptions would not be contaminated from prior exposure as would be the case if we recycled an existing deepfake.

To create a natural environment for media consumption, we surround the experimentally manipulated media exposure with four media clips, two before and two after. These reports are all real coverage of different Democratic primary candidates, presented either in audio, textual, or video form. The order and content of these media are fixed, and primarily serve to mask the main manipulation, replicating the visual style of Facebook posts. The six conditions of our manipulation (`video`, `audio`, `text`, `skit`, `ad`, `control`) and their exact differences from each other are shown in Table 3, where `video` is the group assigned to the deepfake.

Participants in the `video`, `audio`, and `skit` conditions are randomly exposed to one of five different scandal events to reduce the possibility that our results are being driven by a single story. Each scandal is entirely fictitious, and the media associated with it was created in collaboration with a professional actor and a tech industry partner. Specifically, the `audio` condition consists of the audio recording of the actor making a scandalous statement. Participants in the `skit` condition are exposed to the original videos used in the creation of the deepfake video, prior to the modifications made by the neural network algorithm. That is, this condition displays the unaltered video of the paid actress hired to impersonate Elizabeth Warren which is clearly framed as a skit: the title of the corresponding deepfake in the `video` condition is shown, but "Leak" is replaced with "Spot-On Impersonation". Finally, the `video` condition employs a deepfake constructed from the footage used in the `skit` condition. Details on the production of these stimuli are provided in Appendix B and each of the five scripts are provided in Table B4. We do not register any hypotheses about heterogeneous effects across these stories within condition, but conduct exploratory analyses which show small differences across conditions (Appendix G).

Finally, in the `ad` condition, subjects are exposed to a real negative campaign ad titled, "Tell Senator Warren: No Faux Casino, Pocahontas!", which criticizes Senator Warren's supposedly illicit support for federally funding a local casino owned by an Indian tribe, despite her previous opposition to such legislation and her disputed claims of Cherokee heritage. Although the ad frames Warren as politically insincere, similar to script (e) and primes the viewer of her Cherokee heritage controversy, similar to script (c), it stylistically and informationally differs

Table 3: **Experimental Conditions in Incidental Exposure Experiment**

	Condition	Description of Variation	Example Clip
Scandal Clips (Script Held Constant)	`video` ($n = 872$)	Face-swap performed on video in skit condition; title and video edited to resemble leaked video footage.	
	`audio` ($n = 954$)	Visuals stripped from video condition; title edited to resemble leaked hot mic.	
	`text` ($n = 950$)	Visuals and sound stripped from video condition; title describes scandal as a leak; subtitle describes event captured on video.	
	`skit` ($n = 956$)	Filmed impersonator portraying a campaign scandal event.	
Reference Stimuli	`ad` ($n = 935$)	Campaign attack advertisement describing real scandal event.	
	`control` ($n = 916$)	No stimulus presented.	N/A

in many other ways, and thus is not an exact ad counterfactual of our deepfake. Nevertheless, the ad serves as a benchmark comparison for a deepfake's affective effect, since it is an actual campaign stimulus used in the primary election to activate negative emotions towards Warren.

Following the feed, respondents are asked to evaluate the credibility of each textual, audio, or video clip in the feed (the extent to which they believe the clip is "fake or doctored"

on a 5 point scale) in between other distraction evaluations (funny, offensive, informative). Consequently, respondents are also asked to evaluate how warmly or coldly they feel towards each of the Democratic candidates on a continuous 100 point feeling thermometer.

Our main counterfactuals of the deepfake `video` condition are the `text` and `audio` conditions. Importantly, we do not make a comparison of credibility ("is this fake or doctored?") of the `skit` and `ad` stimuli with the three scandal clippings, due to concerns about differential item functioning: it is possible that respondents say the ad or skit is "fake or doctored" because they correctly perceive the `skit` as a staged depiction or the `ad` as an edited video rather than because they incorrectly perceive it as depicting Warren participating in a fabricated event. However, we can still usefully compare affective responses towards Warren between the scandal clippings and these reference stimuli.

4.2 Detection task experiment

After completing the battery of questions in which we measure our primary outcomes of interest and ask another attention check question, the subjects begin the second experimental task that measures their ability to discriminate between authentic and deepfake videos.

Before this task, half of the subjects (in addition to all of the subjects not taking part in this task) are debriefed about whether or not they were exposed to a deepfake in the first experiment. The other half are debriefed after this final task. This randomization allows us to test for the effect of the debrief, which unlike the verbal information randomly provided in the first stage provides visual examples of deepfakes. Additionally, half of all respondents are provided an accuracy prime – an intervention designed to increase the salience of information accuracy (Pennycook and Rand, 2019).

In the task itself, we employ videos created by (Agarwal et al., 2019) and a mix of other publicly available deepfake videos of both lip-sync and face-swap varieties. To the extent that respondents have previously viewed these videos, we should expect detection performance to be biased upwards, although no respondent explicitly indicated as such in open feedback. For the pool of authentic videos, we primarily selected, where possible, real-world video scandals of the elites used in the deepfake pool. Unlike in the incidental exposure experiment, in both the deepfake and non-deepfake pools, we have clips of Republican (Donald Trump) and Democrats (Barack Obama, Joe Biden, Elizabeth Warren), creating both Democratic and Republican out-partisans in the detection task. Subjects were randomly assigned to one of three environmental conditions: the percentage of deepfakes in their video feed was either 75% (high-fake), 25% (low-fake) or 0% (no-fake). Appendix C displays screenshots and descriptions of each of these videos. Misclassifications (or reductions in accuracy) in the detection task can be decomposed into false negatives, or misclassifications of deepfakes as authentic, and false positives, or misclassification of authentic clips as deepfakes. We measure both, in addition to overall accuracy, to gauge our respondents' discernment abilities and the source of their errors.

4.3 Ethical considerations

Creating deepfakes raises important ethical concerns, which we aimed to address at every stage of our research design. First, given the risk of deepfakes disrupting elections, understanding their effects is of the utmost importance: this research has the potential to improve the resilience of democratic politics to this technological threat by better informing policy and consumer behavior. Second, we created deepfakes of a candidate who was not currently running for office to ensure that our experiment could not plausibly influence the outcome of an election. Third, we designed "active debriefs" that required subjects to affirm in writing whether they were exposed to false media. Finally, deepfakes are increasingly part of the standard media environment, so our study only exposes subjects to things they should be prepared to encounter online. We discuss these points in more detail in Appendix D.

References

Abadi, Martín, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard et al. 2016. Tensorflow: A System for Large-Scale Machine Learning. In *12th USENIX Symposium on Operating Systems Design and Implementation*. pp. 265–283.

Abram, Cleo. 2020. "The most urgent threat of deepfakes isn't politics. It's porn.".
URL: https://www.vox.com/2020/6/8/21284005/urgent-threat-deepfakes-politics-porn-krist

Agarwal, Shruti, Hany Farid, Yuming Gu, Mingming He, Koki Nagano and Hao Li. 2019. Protecting World Leaders Against Deep Fakes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. pp. 38–45.

Ajder, Henry, Giorgio Patrini, Francesco Cavalli and Laurence Cullen. 2019. "The State of Deepfakes: Landscape, Threats, and Impact." *Policy Brief* .
URL: http://regmedia.co.uk/2019/10/08/deepfake_report.pdf

Alexander, Deborah and Kristi Andersen. 1993. "Gender as a Factor in the Attribution of Leadership Traits." *Political Research Quarterly* 46(3):527–545.

Alvarez, R Michael. 1998. *Information and Elections*. University of Michigan Press.

Aral, Sinan and Dean Eckles. 2019. "Protecting Elections From Social Media Manipulation." *Science* 365(6456):858–861.

Aronow, Peter M., Josh Kalla, Lilla Orr and John Ternovsk. 2020. "Evidence of Rising Rates of Inattentiveness on Lucid in 2020." *Working Paper* .
URL: https://osf.io/preprints/socarxiv/8sbe4

Barbera, Pablo. 2018. Explaining the Spread of Misinformation on Social Media: Evidence From the 2016 US Presidential Election. In *Symposium: Fake News and the Politics of Misinformation. APSA*.

Bartels, Larry M. 2002. "The Impact of Candidate Traits in American Presidential Elections." *Leaders' Personalities and the Outcomes of Democratic Elections* pp. 44–69.

Bateman, Jon. 2020. "Deepfakes and Synthetic Media in the Financial System: Assessing Threat Scenarios.".

Benjamini, Yoav and Yosef Hochberg. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing." *Journal of the Royal Statistical Society: Series B (Methodological)* 57(1):289–300.

Berelson, Bernard R, Paul F Lazarsfeld and William N McPhee. 1954. *Voting: A Study of Opinion Formation in a Presidential Campaign*. University of Chicago Press.

Blum, Jeremy. 2020. "Trump Rant About 'Anarchist' Protesters Wielding Deadly 'Cans Of Soup' Goes Viral.".
URL: https://www.huffpost.com/entry/trump-deadly-cans-of-soup_n_5f4fbcc6c5b69eb5c0379f01

Bode, Leticia, Ceren Budak, Jonathan M Ladd, Frank Newport, Josh Pasek, Lisa O Singh, Stuart N Soroka and Michael W Traugott. 2020. *Words That Matter: How the News and Social Media Shaped the 2016 Presidential Campaign*. Brookings Institution Press.

Brenton, Harry, Marco Gillies, Daniel Ballin and David Chatting. 2005. The Uncanny Valley: Does It Exist? In *Proceedings of the Conference of Human Computer Interaction*.

Brown, Nina Iacono. 2019. "Congress Wants to Solve Deepfakes by 2020. That Should Worry Us.".
URL: https://slate.com/technology/2019/07/congress-deepfake-regulation-230-2020.html

Caprara, Gian Vittorio, Shalom Schwartz, Cristina Capanna, Michele Vecchione and Claudio Barbaranelli. 2006. "Personality and Politics: Values, Traits, and Political Choice." *Political Psychology* 27(1):1–28.

Carpini, Michael X Delli and Scott Keeter. 1996. *What Americans Know About Politics and Why It Matters*. Yale University Press.

Cassese, Erin C and Mirya R Holman. 2019. "Playing the Woman Card: Ambivalent Sexism in the 2016 US Presidential Race." *Political Psychology* 40(1):55–74.

Christianson, Sven-åke and Elizabeth F Loftus. 1987. "Memory for Traumatic Events." *Applied Cognitive Psychology* 1(4):225–239.

Clinton, J., J. Cohen, J. Lapinski and M. Trussler. 2020. "Partisan Pandemic: How Partisanship and Public Health Concerns Affect Individuals' Social Mobility During COVID-19." *Science Advances* .

Davis, Raina. 2020. "Technology Factsheet: Deepfakes." *Policy Brief* .
URL: https://www.belfercenter.org/publication/technology-factsheet-deepfakes

Dobber, Tom, Nadia Metoui, Damian Trilling, Natali Helberger and Claes de Vreese. 2020. "Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?" *The International Journal of Press/Politics* .

Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York.

Druckman, James N and Mary C McGrath. 2019. "The Evidence for Motivated Reasoning in Climate Change Preference Formation." *Nature Climate Change* 9(2):111–119.

Enders, Adam M and Steven M Smallpage. 2019. "Informational Cues, Partisan-Motivated Reasoning, and the Manipulation of Conspiracy Beliefs." *Political Communication* 36(1):83–102.

Frum, David. 2020. "The Very Real Threat of Trump's Deepfake.".
URL: *https://www.theatlantic.com/ideas/archive/2020/04/trumps-first-deepfake/610750/*

Galston, William A. 2020. "Is seeing still believing? The deepfake challenge to truth in politics.".
URL: *https://www.brookings.edu/research/is-seeing-still-believing-the-deepfake-challenge-to-truth-in-politics/*

Gazis, Olivia and Stefan Becket. 2019. "Senators Pressure Social Media Giants to Crack Down on "Deepfakes".".
URL: https://www.cbsnews.com/news/deepfakes-mark-warner-marco-rubio-pressure-social-me

Glick, Peter and Susan T Fiske. 1996. "The Ambivalent Sexism Inventory: Differentiating Hostile and Benevolent Sexism." *Journal of Personality and Social Psychology* 70(3):491.

Goldberg, Matthew H, Sander van der Linden, Matthew T Ballew, Seth A Rosenthal, Abel Gustafson and Anthony Leiserowitz. 2019. "The Experience of Consensus: Video as an Effective Medium to Communicate Scientific Agreement on Climate Change." *Science Communication* 41(5):659–673.

Goodman, J. David. 2012. "Microphone Catches a Candid Obama.".
URL: https://www.nytimes.com/2012/03/27/us/politics/obama-caught-on-microphone-telling-medvedev-of-flexibility.html

Government Accountability Office: Science, Technology Assessment and Analytics. 2020. "Science and Tech Spotlight: Deepfakes.".
URL: https://www.gao.gov/assets/710/704774.pdf

Grabe, Maria Elizabeth and Erik Page Bucy. 2009. *Image Bite Politics: News and the Visual Framing of Elections*. Oxford University Press.

Guess, Andrew, Jonathan Nagler and Joshua Tucker. 2019. "Less Than You Think: Prevalence and Predictors of Fake News Dissemination on Facebook." *Science Advances* 5(1).

Guess, Andrew M., Michael Lerner, Benjamin Lyons, Jacob M. Montgomery, Brendan Nyhan, Jason Reifler and Neelanjan 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):15536–15545.
URL: https://www.pnas.org/content/117/27/15536

Harris, Douglas. 2018. "Deepfakes: False Pornography Is Here and the Law Cannot Protect You." *Duke L. & Tech. Rev.* 17:99.

Harwell, Drew. 2019. "Top AI Researchers Race to Detect 'Deepfake' Videos: 'We are outgunned'.".
URL: https://www.washingtonpost.com/technology/2019/06/12/top-ai-researchers-race-detect-deepfake-videos-we-are-outgunned/

Herman, Edward S and Noam Chomsky. 2010. *Manufacturing Consent: The Political Economy of the Mass Media*. Random House.

Hollyer, James R, B Peter Rosendorff and James Raymond Vreeland. 2019. "Transparency, Protest and Democratic Stability." *British Journal of Political Science* 49(4):1251–1277.

Hwang, Tim and Clint Watts. 2020. "Opinion: Deepfakes are coming for American democracy. Here's how we can prepare.".
URL: https://www.washingtonpost.com/opinions/2020/09/10/deepfakes-are-coming-american-democracy-heres-how-we-can-prepare/

Jamieson, Kathleen Hall, Kathleen Hall et al. 1995. *Beyond the Double Bind: Women and Leadership*. Oxford University Press.

Jerit, Jennifer and Yangzi Zhao. 2020. "Political Misinformation." *Annual Review of Political Science* 23:77–94.

Kahan, Dan M. 2012. "Ideology, Motivated Reasoning, and Cognitive Reflection: An Experimental Study." *Judgment and Decision Making* 8:407–24.

Kassin, Saul M and David A Garfield. 1991. "Blood and Guts: General and Trial-Specific Effects of Videotaped Crime Scenes on Mock Jurors." *Journal of Applied Social Psychology* 21(18):1459–1472.

Ko, Allan, Merry Mou and Nathan Matias. 2016. "The Obligation to Experiment." *Medium* .

Krook, Mona Lena and Juliana Restrepo Sanín. 2020. "The Cost of Doing Politics? Analyzing Violence and Harassment Against Female Politicians." *Perspectives on Politics* 18(3):740–755.

Kuklinski, James H, Paul J Quirk, Jennifer Jerit, David Schwieder and Robert F Rich. 2000. "Misinformation and the Currency of Democratic Citizenship." *Journal of Politics* 62(3):790–816.

Leeper, Thomas J and Rune Slothuus. 2014. "Political Parties, Motivated Reasoning, and Public Opinion Formation." *Political Psychology* 35:129–156.

Leonhardt, David and Stuart Thompson. 2017. "Trump's Lies.".
URL: https://www.nytimes.com/interactive/2017/06/23/opinion/trumps-lies.html

Lewis, Rebecca. 2018. "Alternative Influence: Broadcasting the Reactionary Right on YouTube." *Data & Society* 18.

Lippmann, Walter. 1922. *Public Opinion*. New Cork.

Lum, Zi-Ann. 2019. "Obama Tells Canadian Crowd He's Worried About 'Deepfake' Videos.".
URL: https://www.huffingtonpost.ca/entry/obama-deepfake-video_ca_5cf29aafe4b0e8085e3ad233

Lupia, Arthur. 2016. *Uninformed: Why People Know So Little About Politics and What We Can Do About It*. Oxford University Press.

Makhzani, Alireza, Jonathon Shlens, Navdeep Jaitly, Ian Goodfellow and Brendan Frey. 2015. "Adversarial Autoencoders." *Working Paper* .

Mori, Masahiro, Karl F MacDorman and Norri Kageki. 2012. "The Uncanny Valley." *IEEE Robotics & Automation Magazine* 19(2):98–100.

Munger, Kevin, Mario Luca, Jonathan Nagler and Joshua Tucker. 2020. "The (Null) Effects of Clickbait Headlines on Polarization, Trust, and Learning." *Public Opinion Quarterly* .

Osmundsen, Mathias, Alexander Bor, Peter Bjerregaard Vahlstrup, Anja Bechmann and Michael Bang Petersen. 2020. "Partisan Polarization Is the Primary Psychological Motivation Behind "Fake News" Sharing on Twitter.".

Parkin, Simon. 2019. "The Rise of the Deepfake and the Threat to Democracy.".
URL: https://www.theguardian.com/technology/ng-interactive/2019/jun/22/the-rise-of-the-deepfake-and-the-threat-to-democracy

Pennycook, Gordon and David G Rand. 2019. "Lazy, Not Biased: Susceptibility to Partisan Fake News Is Better Explained by Lack of Reasoning Than by Motivated Reasoning." *Cognition* 188:39–50.

Pennycook, Gordon, Jonathon McPhetres, Yunhao Zhang, Jackson G Lu and David G Rand. 2020. "Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention." *Psychological Science* 31(7):770–780.

Pennycook, Gordon, Ziv Epstein, Mohsen Mosleh, Antonio A Arechar, Dean Eckles and David G Rand. 2019. "Understanding and Reducing the Spread of Misinformation Online." *Working Paper* .

Pierce, Patrick A. 1993. "Political Sophistication and the Use of Candidate Traits in Candidate Evaluation." *Political Psychology* pp. 21–35.

Popkin, Samuel L. 1991. *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns*. University of Chicago Press.

Prochaska, Stephen, Michael Grass and Jevin West. 2020. "Deepfakes in the 2020 Election and Beyond: Lessons From the 2020 Workshop Series." *Center for an Informed Republic* .
URL: https://cpb-us-e1.wpmucdn.com/sites.uw.edu/dist/6/4560/files/2020/10/CIP_Deepfakes_Report_Extended.pdf

Quealy, Kevin. 2021. "Trump's Lies.".
URL: https://www.nytimes.com/interactive/2021/01/19/upshot/trump-complete-insult-list.html

Rini, Regina. 2020. "Deepfakes and the Epistemic Backstop." *Philosopher's Imprint* 20(24).

Schaffner, Brian F, Matthew MacWilliams and Tatishe Nteta. 2018. "Understanding White Polarization in the 2016 Vote for President: The Sobering Role of Racism and Sexism." *Political Science Quarterly* 133(1):9–34.

Schick, Nina. 2020. "Deepfakes are jumping from porn to politics. It's time to fight back.".
URL: https://www.wired.co.uk/article/deepfakes-porn-politics

Snyder Jr, James M and David Strömberg. 2010. "Press Coverage and Political Accountability." *Journal of Political Economy* 118(2):355–408.

Strömberg, David. 2004. "Mass Media Competition, Political Competition, and Public Policy." *The Review of Economic Studies* 71(1):265–284.

Suwajanakorn, Supasorn, Steven M Seitz and Ira Kemelmacher-Shlizerman. 2017. "Synthesizing Obama: Learning Lip Sync From Audio." *ACM Transactions on Graphics (TOG)* 36(4):1–13.

Tappin, Ben M, Gordon Pennycook and David G Rand. 2020. "Rethinking the Link Between Cognitive Sophistication and Politically Motivated Reasoning." *Journal of Experimental Psychology: General* .

Teele, Dawn, Joshua Kalla and Frances McCall Rosenbluth. 2017. "The Ties That Double Bind: Social Roles and Women's Underrepresentation in Politics." *American Political Science Review* .

Ternovski, John, Joshua Kalla and Peter Michael Aronow. 2021. "Deepfake Warnings for

Political Videos Increase Disbelief but Do Not Improve Discernment: Evidence from Two Experiments.".

Toews, Rob. 2020. "Deepfakes Are Going To Wreak Havoc On Society. We Are Not Prepared.".
URL: https://www.forbes.com/sites/robtoews/2020/05/25/deepfakes-are-going-to-wreak-havoc-on-society-we-are-not-prepared/?sh=7fd212087494

Trump's bizarre "Tim/Apple" tweet is a reminder the president refuses to own tiny mistakes. N.d.
URL: https://www.vox.com/2019/3/11/18259996/trump-tim-cook-apple-tweet-time-and-words

Tucker, Joshua A, Andrew Guess, Pablo Barberá, Cristian Vaccari, Alexandra Siegel, Sergey Sanovich, Denis Stukal and Brendan Nyhan. 2018. "Social media, political polarization, and political disinformation: A review of the scientific literature." *Hewlett Foundation* .

Vaccari, Cristian and Andrew Chadwick. 2020. "Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News." *Social Media+ Society* 6(1).

Wellek, Stefan. 2010. *Testing Statistical Hypotheses of Equivalence and Noninferiority*. CRC press.

Wittenberg, Chloe, Jonathan Zong, David Rand et al. 2020. "The (Minimal) Persuasive Advantage of Political Video Over Text." *Working Paper* .

Yadav, Aman, Michael M Phillips, Mary A Lundeberg, Matthew J Koehler, Katherine Hilden and Kathryn H Dirkin. 2011. "If a Picture Is Worth a Thousand Words Is Video Worth a Million? Differences in Affective and Cognitive Processing of Video and Text Cases." *Journal of Computing in Higher Education* 23(1):15–37.

Zimmer, Ben. 2019. "Elizabeth Warren and the Down-to-Earth Trap.".
URL: https://www.theatlantic.com/entertainment/archive/2019/01/why-elizabeth-warrens-beer-moment-fell-flat/579544/

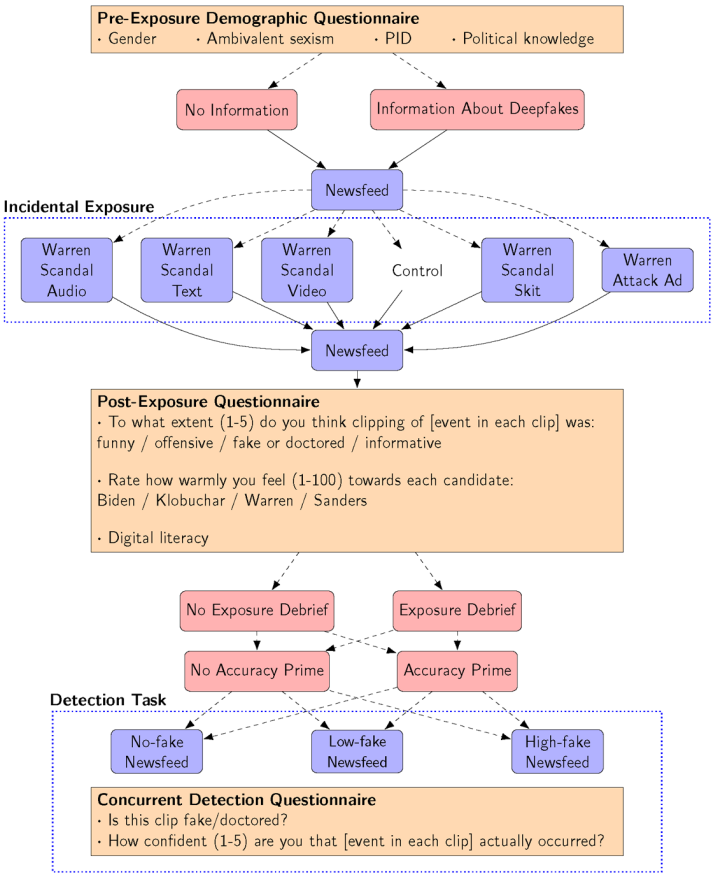
Political Deepfakes Are As Credible As Other Fake Media And (Sometimes) Real Media

Supplementary Information

Contents

A Experiment Overview

Figure A6: **Graphical Summary of Experimental Flow**



Notes: orange cells denote covariate and outcome measurements, **blue** cells denote exposures, and **red** cells denote treatments. In the `audio`, `text`, `skit`, and `video` cells, respondents are further randomized to one of the 5 clippings in Table B4. Subjects who do not receive an exposure debrief prior to the detection task receive it immediately after in overall debrief.

B Stimuli in Exposure Experiment

B.1 Production details

We discuss the ethical reasoning behind our research design in more detail in Appendix D, but we first highlight here our selection of Elizabeth Warren for on both ethical and practical grounds. Senator Elizabeth Warren is a salient politician, making our experiment more ecologically valid than one with a low-profile or hypothetical politician (nearly all political deepfakes target high-profile politicians), but she is not slated for re-election until 2024. We selected a female candidate because women are more likely to be the targets of non-political deepfakes (Ajder et al., 2019; Abram, 2020) and harassment more broadly (Krook and Sanín, 2020) and we specifically test whether pre-existing prejudice against women among subjects changes the effect of the deepfake.

Prior to production, we consulted *Buzzfeed* CEO Jonah Peretti, who produced the first viral deepfake video in 2018 of Barack Obama telling the world that "President Trump is a complete and utter dipsh*t". In the correspondence below, he explained how the deepfake, created via a professional actor's expert impersonation and synthesized via face-swap, came to exist, emphasizing the need for high-quality impersonator and post-production:

> "The idea was shaped by Jordan's ability to do a good Obama impersonation - so that part isn't fancy tech. Jordan is just better at impressions than other people making deep fakes and he did Obama as a character on Key & Peele.
>
> Then we worked with Jared who used a combination of deep fake software downloaded from Reddit and Adobe products we use to do video effects and post production work. It wasn't straightforward and required a combination of approaches and Jared's prodigious talents."

In collaboration with an industry partner and following the lessons from our correspondence with *Buzzfeed*, we produced a series of deepfake videos using target footage of 2020 presidential candidate and senator Elizabeth Warren and performances of a professional Elizabeth Warren impersonator. Warren's campaign disseminated a series of campaign video recordings of the senator in her home kitchen making personal thank-you calls to campaign donors and, in some cases, discussing policy matters and events during the campaign. We produced a series of videos performances of the impersonator in a similar kitchen performing several different sketches that each represented a potential "scandal" for Warren. To script these scandals, we carefully studied past controversial hot mic scandals of Democratic politicians as well as exact statements made by Warren in these campaign videos. We then scripted statements in Warren's natural tone and affliction that appeared plausible in the universe of political hot mic statements. As such, these statements are not meant to invoke extreme disbelief or incredulity, though testing the credulity threshold of deepfake scandals in a principled manner could the subject of future research. Table B4 describes the content of the final performances selected for our experiment. We used the audio from these sketches for the audio condition

Table B4: **Descriptions and Scripts of Scandal Performances**

Scandal Description	Title	Script
In-Party Incivility	LEAK: Elizabeth Warren calls Joe Biden "a piece of sh*t" and a pedophile in 2019 campaign call	"Why shouldn't you vote for Joe Biden in 2020? Well, I'll tell you why: because he's a sexist piece of shit who likes to grope young girls, that's why."
Out-Party Incivility	LEAK: Elizabeth Warren calls Donald Trump "a piece of sh*t" and a pedophile in 2019 campaign call	"Why shouldn't you vote for Donald Trump in 2020? Well, I'll tell you why: because he's a sexist piece of shit who likes to grope young girls, that's why."
Past Controversy (racialized comment)	LEAK: Elizabeth Warren re-claims Cherokee heritage in 2019 campaign call	"Well, you know, as someone who has Cherokee ancestry, who's proud of their Native heritage, I deeply identify with other indigenous people and people of color in this country and I will do everything I can to fight for you in Washington."
Novel Controversy (homophobic comment)	LEAK: Elizabeth Warren admits she doesn't "endorse the LGBTQ lifestyle" in 2019 campaign call	"Well, as a Christian woman of faith, I don't personally support the LGBTQ lifestyle, but I will try to do what I can for marriage equality in Washington."
Political Insincerity	LEAK: Elizabeth Warren flips stance on student loan debt in 2019 campaign call	"Well, I know I've said that before, but I don't really think that eliminating student loan debt for anyone is fair or realistic."

and the video plus audio for the parody skit. We then performed the procedure to create a face-swap deepfake to produce the final deepfake video treatments, one for each selected scandal performance.

B.2 Face-swap algorithm

Deepfakes that swap the face of a target (e.g. President Barack Obama) with an actor (e.g. Hollywood actor Jordan Peele) – dubbed face-swaps in Figure 1 – are synthesized via a particular class of artificial neural networks called Adversarial Autoencoders (Makhzani et al., 2015).

The deepfaker's task is to train two autoencoders to accurately represent (encode) the two respective faces in a latent space and accurately reconstruct (decode) them as images. Let

$\mathbf{X}_{\text{target}}$ denote a set of facial images of the target and $\mathbf{X}_{\text{actor}}$ denote a set of facial images of the actor. Denoting $\mathcal{G}_{\text{target}}$ as the function for the target autoencoder and $\mathcal{G}_{\text{actor}}$ as the function for the actor autoencoder, the networks are structured as $\mathcal{G}_{\text{target}}(x) = \delta_{\text{target}}\{\pi(x)\}$ and $\mathcal{G}_{\text{actor}}(x') = \delta_{\text{actor}}\{\pi(x')\}$ where π is an encoder subnetwork, δ_{target} and δ_{actor} are the decoder subnetworks for the target and actor respectively, and $x \in \mathbf{X}_{\text{target}}, x' \in \mathbf{X}_{\text{actor}}$. Both autoencoders share an encoder function π which discover a common latent representation for the targets' and actors' faces; separate decoders are charged with realistically reconstructing the input faces. The objective function to be optimized is:

$$\min_{\substack{\pi, \\ \delta_{\text{target}}, \\ \delta_{\text{actor}}}} \mathbb{E}_{x \sim \mathbf{X}_{\text{target}}}\Big[||\delta_{\text{target}}\{\pi(x)\} - x||^2\Big] + \mathbb{E}_{x' \sim \mathbf{X}_{\text{actor}}}\Big[||\delta_{\text{actor}}\{\pi(x')\} - x'||^2\Big] \tag{1}$$

To produce a deepfake given a audiovisual performance of the actor with respective facial image frames $\mathbf{Y}_{\text{actor}} = \Big[y_1, \ldots, y_N\Big]$, we input the frames into the trained target autoencoder which outputs $\mathbf{Y}_{\text{actor}} = \Big[\delta_{\text{target}}\{\pi(y_1)\}, \ldots, \delta_{\text{target}}\{\pi(y_N)\}\Big]$ that can be recombined with the audio of the actor's performance.

To maximize the realism of outputs created from actor inputs fed to the target autoencoder, we train a third discriminator neural network $\mathcal{D}$ which aims to accurately classify the latent representations of images as belonging to either the target or actor. The final adversial objective is given as:

$$\begin{aligned}\max_{\mathcal{D}} \min_{\substack{\pi, \\ \delta_{\text{target}}, \\ \delta_{\text{actor}}}} & \mathbb{E}_{x \sim \mathbf{X}_{\text{target}}}\Big[||\delta_{\text{target}}\{\pi(x)\} - x||^2\Big] + \mathbb{E}_{x' \sim \mathbf{X}_{\text{actor}}}\Big[||\delta_{\text{actor}}\{\pi(x')\} - x'||^2\Big] \\ & + \quad \mathbb{E}_{x'' \sim \mathbf{X}}\Big[||\mathcal{D}\{\pi(x'')\} - \mathbf{1}\{x'' \in \mathbf{X}_{\text{actor}}\}||^2\Big]\end{aligned} \tag{2}$$

Optimization of this objective function can be performed via alternating iterative updating of the two networks' weights using stochastic gradient descent. After sufficient rounds of training, the target autoencoder can accurately reproduce the target's face using images of only the actor's face and is thus able to effectively 'fool' the discriminator.

In practice, this workflow for deepfake synthesis is implemented using the `TensorFlow` library (Abadi et al., 2016). Deepfake producers utilize code from several popular public code repositories which implement variants of this base framework – including multiple discriminators and autoencoders, regularization schemes, and particular network architecture choices.

Finally, we reduce the resolution and bit-rate of our stimuli. This increases realism in two ways: (1) by masking any artifacts of the visual alterations of each face-swap and (2) by credibly presenting each video as a 'leaked' mobile phone recording.

C Stimuli in Detection Experiment

This section provides sceenshots of the videos used in the detection experiment. All subject are assigned a mix of videos in which there are either no deepfakes (i.e., all displayed videos are of real media), a low proportion of deepfakes, or a high proportion of deepfakes. Each of

these three conditions employs eight videos. While the order in which videos are presented varies within these conditions, the videos within each condition are fixed across subjects.

This section shows screenshots of each of these videos. Section C.1 shows screenshots and descriptions of all the authentic videos, while section Section C.2 shows screenshots and descriptions of the deepfakes employed in this experiment.

Subjects assigned to the no-fake condition saw real videos C7 through C14. Subjects in the low-fake condition saw fake videos C16 and C17, and real videos C9, C10, C11, C13, C14, and C15. Subjects in the high-fake condition saw fake videos C16, C17, C18, C19, C20, C21 and real videos C8 and C13.

Heterogeneity in detection performance at the clip level (both for the entire pool and across subgroups) can be found in Section G.

C.1 Authentic videos

Figure C7: **Donald Trump ("soup" press conference gaffe)**. Following national demonstrations in the summer of 2020, President Donald Trump decries protestors weaponizing cans of soup against police officers in a soon-to-be viral press conference clip (Blum, 2020).

Figure C8: **Joe Biden (town hall 'push-up contest' gaffe)**. After a heated exchange, Democratic presidential candidate Joe Biden challenges a combative voter at a town hall to a push-up contest.

Figure C9: **Joe Biden (stutter gaffe)**. A video compilation of Joe Biden stuttering in various campaign appearances.

Figure C10: **Donald Trump (COVID-19 precautions announcement)**. In a public address from the White House, President Trump urges Americans to take personal precautions to avoid COVID-19.

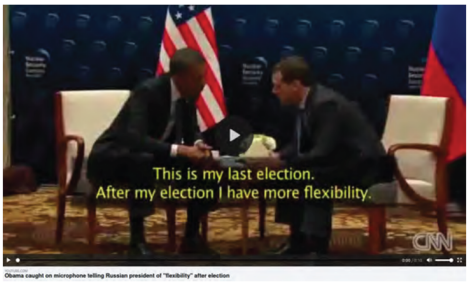


Figure C11: **Barack Obama (Russian president hot mic).** President Barack Obama is caught on a hot mic telling Russian President Dmitry Medvedev of "more flexibility" following his "last election" to negotiate on the issue of missile defense; an exchange that critics suggested revealed a lack of concern about re-election and lack of diplomatic transparency criticized (Goodman, 2012).

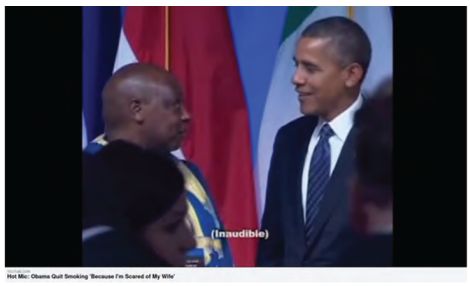


Figure C12: **Barack Obama (smoking hot mic).** President Barack Obama is caught on a hot mic to a U.N. National Assembly attendee saying that he quit smoking because "[he's] scared of [his] wife".

Figure C13: **Elizabeth Warren (Instagram beer gaffe).** Democratic primary candidate Elizabeth Warren furnishes a beer on an livestream video broadcasted on Instagram, a moment criticized as inauthentic and pandering by news media (Zimmer, 2019).

Figure C14: **Elizabeth Warren (post-debate hot mic).** Democratic primary candidate Elizabeth Warren confronts fellow candidate Bernie Sanders on live television for "calling [her] a liar on national TV".

Figure C15: **Donald Trump (Apple press conference gaffe).** During an on-camera White House event, President Donald Trump mistakenly calls Apple CEO Tim Cook "Tim Apple" in a clip to go viral soon after (*Trump's bizarre "Tim/Apple" tweet is a reminder the president refuses to own tiny mistakes*, N.d.).

C.2 Deepfake videos

Figure C16: **Donald Trump (fake AIDS cure announcement).** In a campaign rally speech, President Donald Trump announces that under his administration, scientists have found a cure to AIDS.

Figure C17: **Barack Obama (fake news announcement).** In a White House address, President Barack Obama stresses the importance of relying on trusted news sources.

Figure C18: **Bernie Sanders (fake debate).** In a televised presidential town hall event, Democratic primary candidate Bernie Sanders recalls marching for civil rights in Selma, Alabama.

Figure C19: **Boris Johnson (fake Brexit announcement).** Sitting Prime Minister Boris Johnson announces that in order to "rise above the divide [on Brexit]", he will endorse opposition party leader Jeremy Corbyn in the upcoming U.K. general election.

Figure C20: **Donald Trump (fake resignation announcement).** In a White House address, President Donald Trump notes the American public's disappointment in his leadership and announces his resignation before the 2020 election, citing a need to "put the interests of America first".

Figure C21: **Hillary Clinton (fake debate).** In a televised debate, 2016 Democratic presidential candidate Hillary Clinton labels opponent Donald Trump's tax plan as only benefiting the 1%.

D Ethical Considerations

We highlight the ethical considerations pursuant to a study that uses stimuli which are expected to be uniquely deceptive.

First, in addition to the subjects randomly assigned to a debrief in the middle of the survey, we extensively debrief all subjects at the completion of the survey. This debrief goes beyond the standard description of study procedures. We require respondents to type out the following phrase, depending on which experimental arm they were assigned to:

The [video/audio/text] about Elizabeth Warren is false.

Second, to minimize the risk of influencing the proximate election, we opted to make a deepfake of high-profile 2020 Democratic Presidential candidate who was not ultimately selected as the nominee. Elizabeth Warren is a salient politician, making our experiment more ecologically valid than one with a low-profile or hypothetical politician, but she is slated for re-election until 2024. We selected a female candidate because women are more likely to be the targets of non-political deepfakes, and we specifically test for whether pre-existing prejudice against women among subjects changes the effect of the deepfake. Two of the treatments do refer to Presidential nominees Trump and Biden, but since they are otherwise identical, any effects they produce would be offset.

Third, we carefully weigh the risks to subjects against the potential risks that may be averted with the knowledge gained through our experiment. The potential long-term consequences of exposure to a single piece of media are minimal. That is, participants are unlikely to change their political behavior as a response to treatment, given our extensive debrief. Given that we have no experimental evidence either way, it is at least as likely that our experiment will *benefit* subjects as cause harm. The experiment gives subjects experience detecting fake media, followed up by the debrief which contains feedback and information about how the deepfake process works. Given the importance and seeming inevitability of more deepfakes

in the future, and the uncertainty around their effects, we argue that academics in fact have an "obligation to experiment" (Ko, Mou and Matias, 2016). We believe that improved understanding of how deepfakes function and evidence from our low-cost interventions will in fact serve to prevent real-world harms from deepfakes in the future.

Finally, a similar argument applies to the knowledge we generate from the perspective of policy-makers, journalists, and election administrators (Agarwal et al., 2019). More specifically, our study can inform future legislation or platform policies designed to minimize the threat posed by this technology.[7]

E Sample Description

Our survey experiment was fielded to a nationally representative sample on the Lucid survey research platform to a total of 17,501 subjects launched in two waves between September 29th 2020 and October 29th 2020. Of this 17,501, only 5,750 subjects successfully completed the survey experiment or passed a series of quality checks. One of these quality checks was a battery of randomly dispersed attention checks in response to a recently-publicized issue with in-attention among survey respondents during this period as documented in Aronow et al. (2020). Additionally, we imposed a series of "technology checks," namely that the subjects be able to watch and listen to a video. In addition, 629 respondents failed front-end pre-treatment attention checks: namely, they entered gender or age values that did not match up (or come to close to matching up) with respondent demographic characteristics provided by Lucid. We coded these respondents as "low-quality" respondents which we drop in our analyses as a robustness measure. As expected by Aronow et al. (2020), results largely hold across the two cohorts, but nearly all coefficient estimates are slightly diminished for the low-quality cohort.

Table E5 compares our sample's demographic traits to the demographic traits in the most recent Current Population Survey (CPS) – in particular, traits like education, age, and household income that are hypothesized to have correlations with deepfake deception and affective appeal (by their correlation with digital literacy, internet usage, and political knowledge) as well race, gender and ethnicity which are correlates of partisanship, another predictor of our measured behavioral responses. To adjust for remaining disrepancies, we generate post-stratification weights via raking to match the CPS marginal population totals. We perform weighted regression in our analyses as a robustness measure to guard against measurement error from possible demographic skews.

[7] See `SB 6513` introduced in the Washington state legislature at the time of writing, intended to restrict the use of deepfake audio or visual media in campaigns for elective office.

Table E5: **Sample Demographics and Representativeness After Post-Stratification**

		CPS	Unweighted Sample	Weighted Sample
Education	<High school	10.95%	1.1%	2.87%
	High school	47.14%	29.88%	45.52%
	College	30.3%	47.17%	35.98%
	Postgraduate	11.61%	21.86%	15.63%
Age	18-24	10.42%	6.33%	8.9%
	25-34	13.88%	12.66%	14.91%
	35-44	12.58%	16.97%	16.91%
	45-64	25.76%	31.25%	34.17%
	65+	15.81%	32.77%	25.11%
Household Income	<$25k	19.11%	30.16%	22.63%
	$100k-$150k	14.95%	6.1%	10.67%
	>$150k	15.47%	4.77%	10.44%
	$25k-$49k	20.79%	21.74%	23.17%
	$50k-$74k	17.2%	15.63%	18.6%
	$75k-$99k	12.48%	19.27%	14.5%
Gender	Male	48.75%	33.58%	44.19%
	Female	51.25%	66.14%	55.81%
Race	Asian	5.42%	3.95%	4.4%
	Black	10.28%	5.76%	8.41%
	Other	4.18%	3.81%	3.79%
	White	80.12%	85.79%	83.4%
Hispanic	Yes	14.66%	5.18%	8.59%
	No	85.34%	94.1%	91.41%

Notes: Weights are constructed via Iterative Proportional Fitting to match sample marginal totals to CPS marginal totals on displayed demographic traits. Weights in the final column used for all analyses in paper.

F Pre-Registered Analyses

Our pre-analysis plan containing all pre-reigstered hypotheses can be found at https://osf.io/cdfh3/. For all models, unless otherwise noted or displayed, controls include age group, education, 3 point party ID, cognitive reflection, political knowledge, internet usage, and an indicator for mobile (vs. desktop) exposure. The reference stimuli for all analyses of the incidental exposure experiment is `video`. Reference category for environment in the detection task experiment is high-fake. Cognitive reflection, political knowledge, ambivalent sexism, and internet usage are all re-scaled to [0,1]. Unless otherwise noted, analyses exclude respondents who receive information prior to the incidental exposure experiment, however results (effect magnitudes, statistical significance) are substantively similar in all cases with their inclusion. As pre-registered, all p-values are "step-up" adjusted to $p \cdot r/K$ where r denotes the rank of the unadjusted p-value amongst K total estimated p-values (Benjamini-Hochberg procedure). Analyses do not additionally adjust for respondent wave, for brevity, though we find that including respondent wave as either an interaction term or a linear term does not change our results.

Table F6: **Models of Credibility Confidence of Scandal Clipping in Incidental Exposure Experiment**

	Confidence that clipping was credible [1-5]						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Audio	0.08	0.14	0.10	0.09	0.12	0.11	0.17*
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Text	−0.02	−0.11	0.03	−0.02	−0.15*	0.02	−0.06
	(0.09)	(0.09)	(0.09)	(0.09)	(0.08)	(0.09)	(0.09)
On Mobile				0.04	0.13	0.18*	0.29***
				(0.09)	(0.09)	(0.09)	(0.09)
Age 65+				0.12	0.16	0.14*	0.20*
				(0.08)	(0.08)	(0.08)	(0.08)
High School				−0.23	−0.46*	−0.42	−0.51**
				(0.38)	(0.23)	(0.39)	(0.23)
College				−0.19	−0.40	−0.35	−0.38
				(0.37)	(0.23)	(0.38)	(0.23)
Postgrad				−0.35	−0.48*	−0.52	−0.53*
				(0.38)	(0.24)	(0.39)	(0.24)
Independent PID					0.19	0.15	0.20
					(0.11)	(0.11)	(0.11)
Republican PID					0.54***	0.53***	0.54***
					(0.08)	(0.08)	(0.08)
Cognitive Reflection				−0.02	0.02	0.07	0.18
				(0.15)	(0.15)	(0.16)	(0.16)
Male				−0.07	−0.03	−0.03	0.002
				(0.08)	(0.07)	(0.08)	(0.08)
Political Knowledge				0.17	0.24	0.22	0.31
				(0.16)	(0.16)	(0.17)	(0.16)
Internet Usage				0.56	0.89**	0.54	0.62
				(0.33)	(0.33)	(0.34)	(0.35)
Ambivalent Sexism				0.14***	0.04	0.04	0.05
				(0.04)	(0.04)	(0.05)	(0.05)
Constant	3.40***	3.44***	3.39***	2.58***	2.45***	2.68***	2.47***
	(0.06)	(0.06)	(0.07)	(0.51)	(0.42)	(0.53)	(0.44)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
N	1,356	1,356	1,209	1,356	1,356	1,209	1,209
R^2	0.001	0.01	0.001	0.02	0.06	0.05	0.07
Adjusted R^2	−0.0003	0.005	−0.001	0.01	0.05	0.04	0.05

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F7: **Models of Binarized Credibility Confidence of Scandal Clipping in Incidental Exposure Experiment**

	Somewhat/strongly confident clipping was credible						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Audio	0.01	0.06	0.02	0.02	0.05	0.03	0.06
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Text	−0.04	−0.05	−0.02	−0.05	−0.07	−0.03	−0.04
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
On Mobile				0.08**	0.11***	0.11***	0.15***
				(0.03)	(0.03)	(0.04)	(0.04)
Age 65+				0.06*	0.08**	0.07*	0.10***
				(0.03)	(0.03)	(0.03)	(0.03)
High School				−0.13	−0.21**	−0.18	−0.22*
				(0.14)	(0.09)	(0.15)	(0.09)
College				−0.10	−0.16*	−0.15	−0.17*
				(0.14)	(0.09)	(0.15)	(0.09)
Postgrad				−0.12	−0.15	−0.18	−0.17*
				(0.15)	(0.09)	(0.15)	(0.10)
Independent PID				0.02	0.02	0.02	0.03
				(0.04)	(0.04)	(0.04)	(0.04)
Republican PID				0.20***	0.18***	0.20***	0.19***
				(0.03)	(0.03)	(0.03)	(0.03)
C.R.				−0.03	−0.03	−0.004	0.03
				(0.06)	(0.06)	(0.06)	(0.06)
Male				0.01	0.02	0.01	0.01
				(0.03)	(0.03)	(0.03)	(0.03)
Political Knowledge				0.14*	0.22***	0.17**	0.26***
				(0.06)	(0.06)	(0.06)	(0.06)
Internet Usage				0.26*	0.40**	0.23	0.32
				(0.13)	(0.13)	(0.13)	(0.14)
Ambivalent Sexism				0.02	0.04	0.02	0.04**
				(0.02)	(0.02)	(0.02)	(0.02)
Constant	0.47***	0.47***	0.46***	0.05	−0.11	0.08	−0.11
	(0.02)	(0.02)	(0.03)	(0.20)	(0.16)	(0.21)	(0.17)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
N	1,356	1,356	1,209	1,356	1,356	1,209	1,209
R^2	0.002	0.01	0.001	0.06	0.08	0.06	0.08
Adjusted R^2	0.001	0.01	−0.0003	0.05	0.07	0.05	0.07

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F8: **Models of Information Provision and Credibility Confidence of Clipping in Incidental Exposure Experiment**

	Confidence that clipping was credible [1-5]						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Information	−0.35***	−0.29***	−0.35**	−0.38***	−0.34***	−0.38***	−0.31***
	(0.09)	(0.09)	(0.10)	(0.09)	(0.09)	(0.09)	(0.09)
Audio	0.08	0.14	0.10	0.11	0.14	0.12	0.20**
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Text	−0.02	−0.11	0.03	−0.01	−0.14	0.03	−0.06
	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Info x Audio	0.09	−0.08	0.09	0.10	−0.03	0.12	−0.04
	(0.12)	(0.12)	(0.13)	(0.12)	(0.12)	(0.13)	(0.13)
Info x Text	0.20	0.22	0.18	0.24	0.30**	0.25	0.23*
	(0.12)	(0.12)	(0.13)	(0.12)	(0.12)	(0.13)	(0.13)
Constant	3.40***	3.44***	3.39***	2.87***	3.13***	2.92***	3.10***
	(0.06)	(0.06)	(0.07)	(0.36)	(0.29)	(0.38)	(0.32)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
N	2,724	2,724	2,419	2,724	2,724	2,419	2,419
R^2	0.01	0.01	0.01	0.06	0.06	0.06	0.06
Adjusted R^2	0.01	0.01	0.01	0.05	0.05	0.05	0.05

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F9: **Models of Information Provision and Binarized Credibility Confidence of Clipping in Incidental Exposure Experiment**

	Somewhat/strongly confident clipping was credible						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Information	−0.11**	−0.09**	−0.10**	−0.12	−0.11	−0.12	−0.09
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
Audio	0.01	0.06*	0.02	0.02	0.06	0.03	0.07
	(0.03)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.04)
Text	−0.04	−0.05	−0.02	−0.05	−0.06	−0.03	−0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Info x Audio	0.02	−0.04	0.01	0.03	−0.02	0.02	−0.04
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Info x Text	0.09	0.08	0.07	0.11	0.12	0.10	0.09
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Constant	0.47***	0.47***	0.46***	0.11	0.15	0.11	0.14
	(0.02)	(0.02)	(0.03)	(0.14)	(0.11)	(0.15)	(0.12)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
N	2,724	2,724	2,419	2,724	2,724	2,419	2,419
R^2	0.01	0.01	0.01	0.05	0.05	0.05	0.05
Adjusted R^2	0.01	0.01	0.005	0.04	0.04	0.04	0.05

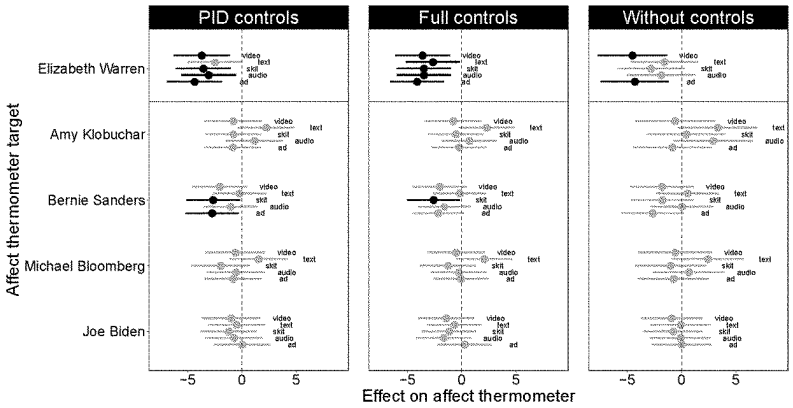
Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F10: **Models of Scandal Target Affect in Incidental Exposure Experiment**

	Elizabeth Warren Affect Thermometer						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Video	−4.54***	−4.08**	−2.86	−3.55**	−4.11***	−2.76	−3.59**
	(1.63)	(1.62)	(1.75)	(1.29)	(1.31)	(1.37)	(1.39)
Audio	−1.89	−3.09*	−0.68	−3.44***	−4.58***	−2.86*	−4.06***
	(1.59)	(1.59)	(1.70)	(1.26)	(1.28)	(1.33)	(1.37)
Text	−1.59	−0.75	−1.04	−2.60	−1.22	−2.36	−1.68
	(1.59)	(1.59)	(1.70)	(1.26)	(1.28)	(1.33)	(1.35)
Skit	−2.81	−3.12*	−2.00	−3.41**	−3.60*	−3.05	−4.09**
	(1.59)	(1.58)	(1.70)	(1.26)	(1.28)	(1.33)	(1.35)
Attack Ad	−4.31	−4.58***	−2.85*	−4.06***	−3.81***	−3.53**	−3.86***
	(1.60)	(1.57)	(1.71)	(1.27)	(1.27)	(1.34)	(1.33)
Information				0.65	0.57	0.63	0.69
				(0.73)	(0.74)	(0.77)	(0.78)
On Mobile				−2.18**	−1.48	−3.22**	−3.11***
				(0.90)	(0.92)	(0.95)	(0.97)
Age 65+				−4.82***	−5.50***	−4.23***	−5.36***
				(0.82)	(0.86)	(0.86)	(0.89)
High School				−0.83	−1.18	−1.80	−1.60
				(3.60)	(2.30)	(3.72)	(2.33)
College				1.32	2.00	−0.62	0.51
				(3.59)	(2.34)	(3.72)	(2.39)
Postgrad				11.16***	14.37***	8.86	12.78***
				(3.65)	(2.48)	(3.79)	(2.55)
Independent PID				−26.99***	−27.02***	−26.73***	−27.52***
				(1.21)	(1.22)	(1.26)	(1.27)
Republican PID				−39.89***	−37.87***	−40.49***	−39.10***
				(0.83)	(0.84)	(0.88)	(0.89)
C.R.				−1.58	−1.96	−0.67	−0.76
				(1.62)	(1.65)	(1.72)	(1.75)
Male				0.87	−0.14	0.62	0.02
				(0.81)	(0.79)	(0.86)	(0.83)
Political Knowledge				1.21	0.95	1.24	0.81
				(1.68)	(1.67)	(1.77)	(1.76)
Internet Usage				6.01*	5.90	6.98*	7.96**
				(3.40)	(3.46)	(3.66)	(3.74)
Ambivalent Sexism				−3.78***	−3.14***	−4.44***	−3.58***
				(0.47)	(0.47)	(0.50)	(0.50)
Constant	45.81***	45.11***	44.17***	69.66***	67.78***	71.49***	68.29***
	(1.14)	(1.12)	(1.23)	(5.13)	(4.34)	(5.43)	(4.62)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
N	5,524	5,524	4,895	5,523	5,523	4,894	4,894
R^2	0.002	0.002	0.001	0.38	0.35	0.39	0.37
Adjusted R^2	0.001	0.002	−0.0000	0.38	0.35	0.39	0.37

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$
Reference category for clip type is `control`.

Figure F22: **Effects of Clip Type on Placebo Targets' Favorability in Incidental Exposure Experiment**



Notes: Shown are other candidates who ran in the 2020 Democratic primary for whom we selected clips to mask our deepfake in the incidental exposure experiment.

Table F11: **Models of Information Provision and Media Trust in Incidental Exposure Experiment**

	Dependent variable:						
	Trust in Media (Combined Index)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Information	−0.02	−0.03	−0.02	−0.02	−0.02	−0.02	−0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Constant	2.27***	2.29***	2.27***	2.20***	2.18***	2.20***	2.18***
	(0.02)	(0.02)	(0.02)	(0.14)	(0.12)	(0.14)	(0.12)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
Observations	3,231	3,231	3,231	3,231	3,231	3,231	3,231
R^2	0.0002	0.001	0.0002	0.19	0.23	0.19	0.23
Adjusted R^2	−0.0001	0.0003	−0.0001	0.19	0.23	0.19	0.23
Note:	*Notes:* · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$ Respondents in `skit` and `ad` conditions are excluded.						

Table F12: **Models of Information Provision and Media Trust Across Sources in Incidental Exposure Experiment**

	Trust in...							
	Offline Media		Online Media		Social Media		Combined Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Information	0.004	0.003	−0.03	−0.03	−0.04	−0.04	−0.02	−0.02
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.02)	(0.02)
Constant	2.62***	2.77	2.30***	1.96	1.87***	1.89	2.27***	2.20***
	(0.02)	(0.18)	(0.02)	(0.17)	(0.02)	(0.18)	(0.02)	(0.14)
Controls?		✓		✓		✓		✓
Observations	3,231	3,231	3,231	3,231	3,231	3,231	3,231	3,231
R^2	0.0000	0.17	0.0003	0.13	0.0005	0.17	0.0002	0.19
Adjusted R^2	−0.0003	0.17	0.0000	0.13	0.0002	0.16	−0.0001	0.19

Note: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$
Respondents in `skit` and `ad` conditions are excluded.

Table F13: **Models of Deepfake Exposure, Credibility, and Media Trust in Incidental Exposure Experiment**

	Trust in Media (Combined Index)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Credible	−0.11***	−0.07***	−0.12	−0.08***	−0.08***	−0.08***	−0.14***
	(0.02)	(0.02)	(0.04)	(0.02)	(0.02)	(0.02)	(0.04)
Video	−0.30**	−0.28**	−0.09	−0.45***	−0.25**	−0.41***	−0.20***
	(0.11)	(0.10)	(0.05)	(0.10)	(0.11)	(0.11)	(0.05)
Credible x Video	0.07**	0.07	0.10	0.10***	0.06*	0.09***	0.19*
	(0.03)	(0.03)	(0.07)	(0.03)	(0.03)	(0.03)	(0.07)
Constant	2.66***	2.58***	2.40***	2.46***	2.59***	2.57***	2.28***
	(0.06)	(0.24)	(0.24)	(0.20)	(0.25)	(0.21)	(0.20)
Weighted?				✓		✓	✓
Low-Quality Dropped?					✓	✓	✓
Controls?		✓	✓	✓	✓	✓	✓
Credibility Binarized?			✓				✓
N	1,354	1,354	1,354	1,354	1,207	1,207	1,354
R^2	0.03	0.23	0.23	0.27	0.23	0.27	0.27
Adjusted R^2	0.03	0.22	0.22	0.27	0.22	0.26	0.26

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$
Respondents in `skit` and `ad` conditions are excluded.

Table F14: **Models of Deepfake Exposure, Credibility, and Media Trust Across Sources in Incidental Exposure Experiment**

	Trust in...							
	Offline Media		Online Media		Social Media		Combined Index	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Credibility	−0.08***	−0.14***	−0.06***	−0.11**	−0.08***	−0.12*	−0.07***	−0.12
	(0.02)	(0.05)	(0.02)	(0.05)	(0.02)	(0.05)	(0.02)	(0.04)
Video	−0.21	−0.05	−0.25*	−0.10	−0.37**	−0.13*	−0.28**	−0.09
	(0.13)	(0.06)	(0.12)	(0.06)	(0.13)	(0.06)	(0.10)	(0.05)
Credibility x Video	0.05	0.06	0.06	0.14	0.08	0.09	0.07**	0.10
	(0.03)	(0.09)	(0.03)	(0.09)	(0.04)	(0.09)	(0.03)	(0.07)
Constant	2.86***	2.68***	2.30***	2.16***	2.57***	2.38***	2.58***	2.40
	(0.30)	(0.30)	(0.29)	(0.29)	(0.31)	(0.31)	(0.24)	(0.24)
Controls?	✓	✓	✓	✓	✓	✓	✓	✓
Credibility Binarized?		✓		✓		✓		✓
Observations	1,354	1,354	1,354	1,354	1,354	1,354	1,354	1,354
R^2	0.21	0.20	0.16	0.16	0.20	0.19	0.23	0.23
Adjusted R^2	0.20	0.19	0.15	0.15	0.19	0.18	0.22	0.22

Note: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$
Respondents in `skit` and `ad` conditions are excluded.

Table F15: **Models of Cognitive Reflection and Credibility Confidence of Clipping in Incidental Exposure Experiment**

	Confidence that clipping was credible [1-5]						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Audio	0.08	0.05	0.11	0.10	0.10	0.15	0.19
	(0.11)	(0.11)	(0.11)	(0.10)	(0.10)	(0.11)	(0.11)
Text	0.11	−0.04	0.14	0.11	−0.04	0.14	−0.02
	(0.10)	(0.10)	(0.11)	(0.10)	(0.10)	(0.11)	(0.11)
C.R.	−0.09	−0.12	−0.03	−0.10	−0.09	−0.03	−0.02
	(0.20)	(0.19)	(0.21)	(0.19)	(0.19)	(0.21)	(0.20)
C.R. x Audio	0.12	0.13	0.07	0.16	0.08	0.08	−0.06
	(0.27)	(0.28)	(0.29)	(0.26)	(0.27)	(0.29)	(0.29)
C.R. x Text	−0.11	0.11	−0.07	0.01	0.17	0.03	0.25
	(0.27)	(0.27)	(0.29)	(0.26)	(0.26)	(0.28)	(0.28)
Constant	3.26***	3.33***	3.23***	2.83***	3.17***	2.91***	3.17***
	(0.08)	(0.07)	(0.08)	(0.35)	(0.29)	(0.38)	(0.31)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
N	2,724	2,724	2,419	2,724	2,724	2,419	2,419
R^2	0.002	0.001	0.002	0.06	0.05	0.06	0.06
Adjusted R^2	0.0001	−0.001	0.0001	0.05	0.05	0.05	0.05

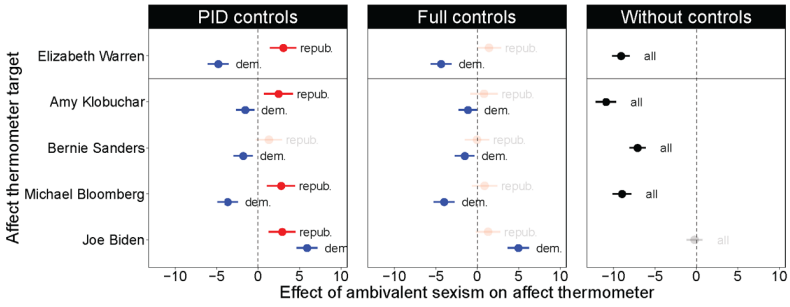
Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F16: **Models of Cognitive Reflection and Binarized Credibility Confidence of Clipping in Incidental Exposure Experiment**

	Somewhat/strongly confident clipping was credible						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Audio	−0.01	−0.01	−0.01	−0.005	0.01	0.003	0.03
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Text	−0.001	−0.03	0.01	−0.002	−0.03	0.01	−0.02
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
C.R.	−0.09	−0.12	−0.08	−0.11	−0.13	−0.10	−0.09
	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)
C.R. x Audio	0.11	0.13	0.11	0.12	0.11	0.11	0.07
	(0.10)	(0.10)	(0.11)	(0.10)	(0.10)	(0.11)	(0.11)
C.R. x Text	0.004	0.08	0.01	0.04	0.10	0.04	0.10
	(0.10)	(0.10)	(0.11)	(0.10)	(0.10)	(0.11)	(0.11)
Constant	0.45***	0.47***	0.44***	0.11	0.17	0.12	0.17
	(0.03)	(0.03)	(0.03)	(0.13)	(0.11)	(0.14)	(0.12)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
N	2,724	2,724	2,419	2,724	2,724	2,419	2,419
R^2	0.002	0.002	0.001	0.05	0.05	0.05	0.05
Adjusted R^2	−0.0001	0.001	−0.001	0.04	0.04	0.04	0.04

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Figure F23: **Ambivalent Sexism and Affect Towards Placebo Targets**



Notes: Shown are other politicians who ran in the 2020 Democratic primary. Conservatively, we would expect that ambivalent sexism would predict the strongest negative affective response towards other female candidates.

Table F17: **Models of Partisan Group Identity and Credibility Confidence of Clipping in Incidental Exposure Experiment**

	Confidence that clipping was credible [1-5]						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Repub PID	0.49	0.29*	0.44***	0.48***	0.31*	0.43**	0.25
	(0.15)	(0.15)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)
Audio	0.13	0.08	0.15	0.15	0.13	0.18	0.22
	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)	(0.14)
Text	0.09	−0.17	0.06	0.10	−0.14	0.09	−0.18
	(0.13)	(0.13)	(0.14)	(0.13)	(0.13)	(0.14)	(0.14)
C.R.	−0.005	−0.20	0.02	0.01	−0.12	0.03	−0.10
	(0.25)	(0.25)	(0.27)	(0.25)	(0.25)	(0.27)	(0.27)
Repub x Audio	−0.10	−0.05	−0.05	−0.12	−0.09	−0.08	−0.07
	(0.21)	(0.21)	(0.23)	(0.21)	(0.21)	(0.22)	(0.22)
Repub x Text	0.05	0.31	0.19	−0.005	0.23	0.12	0.38
	(0.21)	(0.21)	(0.22)	(0.21)	(0.21)	(0.22)	(0.22)
C.R. x Repub	−0.22	0.14	−0.10	−0.24	0.08	−0.13	0.18
	(0.39)	(0.39)	(0.43)	(0.39)	(0.38)	(0.43)	(0.41)
Audio x C.R.	−0.05	−0.06	−0.11	−0.03	−0.10	−0.09	−0.24
	(0.35)	(0.36)	(0.37)	(0.35)	(0.36)	(0.37)	(0.38)
Text x C.R.	−0.17	0.19	0.02	−0.14	0.13	0.05	0.39
	(0.34)	(0.35)	(0.36)	(0.33)	(0.35)	(0.36)	(0.37)
Repub x Audio x C.R.	0.48	0.47	0.44	0.48	0.49	0.46	0.49
	(0.54)	(0.55)	(0.59)	(0.54)	(0.54)	(0.59)	(0.58)
Repub x Text x C.R.	0.37	−0.06	−0.10	0.45	0.10	−0.002	−0.30
	(0.54)	(0.54)	(0.58)	(0.54)	(0.53)	(0.58)	(0.57)
Constant	3.04***	3.21***	3.04***	2.91***	3.38***	2.99***	3.39***
	(0.10)	(0.10)	(0.10)	(0.36)	(0.30)	(0.38)	(0.32)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
N	2,724	2,724	2,419	2,724	2,724	2,419	2,419
R^2	0.04	0.03	0.04	0.06	0.05	0.06	0.06
Adjusted R^2	0.03	0.03	0.03	0.05	0.05	0.05	0.05

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$
PID is pooled to Republican/Not Republican for brevity. PID interacted with C.R. to test possible mechanism of motivated reasoning (pre-registered), although as a reviewer pointed out this is not a sufficient test of a motivated reasoning mechanism by itself.

Table F18: **Models of Partisan Group Identity and Binarized Credibility Confidence of Clipping in Incidental Exposure Experiment**

	Somewhat/strongly confident clipping was credible						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Repub PID	0.15	0.10	0.12**	0.14**	0.11	0.13*	0.08
	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)	(0.06)
Audio	−0.01	0.02	−0.01	−0.01	0.04	−0.01	0.05
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Text	−0.03	−0.08	−0.04	−0.03	−0.07	−0.04	−0.09
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
C.R.	−0.05	−0.11	−0.06	−0.06	−0.09	−0.07	−0.09
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
Repub x Audio	0.01	−0.06	0.03	0.01	−0.07	0.02	−0.06
	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.09)
Repub x Text	0.07	0.12	0.12	0.06	0.10	0.10	0.17
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
C.R. x Repub	−0.12	−0.04	−0.04	−0.13	−0.08	−0.06	−0.01
	(0.15)	(0.15)	(0.16)	(0.15)	(0.15)	(0.16)	(0.16)
Audio x C.R.	0.05	−0.01	0.07	0.05	−0.04	0.07	−0.06
	(0.13)	(0.14)	(0.14)	(0.13)	(0.14)	(0.14)	(0.15)
Text x C.R.	−0.03	0.06	0.04	−0.03	0.05	0.05	0.14
	(0.13)	(0.13)	(0.14)	(0.13)	(0.13)	(0.14)	(0.14)
Repub x Audio x C.R.	0.15	0.33	0.09	0.15	0.34	0.11	0.30
	(0.21)	(0.21)	(0.22)	(0.20)	(0.21)	(0.22)	(0.22)
Repub x Text x C.R.	0.18	0.08	−0.02	0.19	0.11	0.01	−0.09
	(0.21)	(0.20)	(0.22)	(0.20)	(0.20)	(0.22)	(0.22)
Constant	0.38***	0.42***	0.39***	0.13	0.22*	0.15	0.22*
	(0.04)	(0.04)	(0.04)	(0.14)	(0.11)	(0.15)	(0.12)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
N	2,724	2,724	2,419	2,724	2,724	2,419	2,419
R^2	0.03	0.03	0.03	0.05	0.05	0.05	0.06
Adjusted R^2	0.03	0.03	0.03	0.04	0.04	0.04	0.05

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F19: **Models of Partisan Group Identity and Scandal Target Affect in Incidental Exposure Experiment**

	Elizabeth Warren Feeling Thermometer						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Repub PID	−34.09***	−30.24***	−36.46***	−32.22***	−28.84***	−34.87***	−32.25***
	(3.27)	(3.33)	(3.45)	(3.18)	(3.23)	(3.35)	(3.43)
Video	−7.36	−9.29**	−6.70**	−8.30**	−9.01*	−7.90	−8.35**
	(3.09)	(3.22)	(3.28)	(3.00)	(3.12)	(3.19)	(3.31)
Audio	−8.39**	−7.86**	−9.03***	−9.40*	−7.41**	−10.50***	−8.18**
	(3.08)	(3.19)	(3.25)	(2.99)	(3.10)	(3.16)	(3.30)
Text	−1.25	0.70	−2.03	−2.37	0.94	−3.81	−1.27
	(2.97)	(3.15)	(3.13)	(2.89)	(3.06)	(3.04)	(3.24)
Skit	−6.52**	−7.53	−6.10*	−7.48**	−7.43**	−7.23**	−7.52
	(3.00)	(3.11)	(3.18)	(2.92)	(3.02)	(3.10)	(3.21)
Attack Ad	−4.63	−4.34	−4.77	−5.52	−4.69	−6.04	−3.18
	(3.09)	(3.20)	(3.26)	(3.00)	(3.10)	(3.17)	(3.28)
C.R.	1.55	−1.07	−0.34	−4.49	−4.25	−7.52	−4.67
	(5.76)	(5.92)	(6.21)	(5.61)	(5.75)	(6.05)	(6.27)
Repub x Video	2.61	3.80	3.54	3.26	3.34	4.66	5.63
	(4.68)	(4.77)	(4.96)	(4.55)	(4.62)	(4.82)	(4.91)
Repub x Audio	4.36	4.19	5.29	4.22	3.41	6.06	4.39
	(4.66)	(4.78)	(4.91)	(4.52)	(4.63)	(4.76)	(4.91)
Repub x Text	−3.19	−7.09	−1.35	−1.87	−5.38	1.02	−1.15
	(4.56)	(4.68)	(4.75)	(4.43)	(4.54)	(4.62)	(4.76)
Repub x Skit	−0.85	−0.01	−0.15	−0.33	−0.54	0.80	2.08
	(4.53)	(4.64)	(4.76)	(4.40)	(4.50)	(4.63)	(4.74)
Repub x Ad	0.60	−3.82	1.50	2.39	−1.34	3.94	−2.10
	(4.62)	(4.62)	(4.86)	(4.49)	(4.48)	(4.72)	(4.71)
C.R. x Repub	−15.10	−12.87	−9.78	−13.87	−12.43	−6.38	−4.96
	(8.79)	(8.79)	(9.37)	(8.54)	(8.52)	(9.10)	(9.18)
Video x C.R.	7.09	15.92*	7.56	9.14	13.80	10.61	13.09
	(8.08)	(8.33)	(8.64)	(7.85)	(8.08)	(8.40)	(8.62)
Audio x C.R.	12.26	6.28	17.75*	12.80	3.46	19.54**	8.29
	(7.89)	(8.31)	(8.49)	(7.67)	(8.06)	(8.25)	(8.74)
Text x C.R.	−6.50	−5.18	−3.59	−5.09	−6.13	−0.62	−2.96
	(7.59)	(8.04)	(8.11)	(7.38)	(7.81)	(7.89)	(8.42)
Skit x C.R.	5.74	5.84	5.88	7.30	5.11	8.03	2.13
	(7.88)	(8.05)	(8.51)	(7.66)	(7.81)	(8.27)	(8.54)
Ad x C.R.	2.00	5.36	4.59	3.77	7.82	7.27	3.00
	(7.88)	(8.11)	(8.38)	(7.65)	(7.87)	(8.14)	(8.38)
Repub x Video x C.R.	2.34	−7.90	0.16	3.15	−4.27	−0.97	−9.96
	(12.40)	(12.44)	(13.31)	(12.04)	(12.05)	(12.94)	(12.93)
Repub x Audio x C.R.	0.13	0.70	−6.16	1.88	2.63	−7.65	−3.91
	(12.16)	(12.56)	(13.00)	(11.80)	(12.16)	(12.62)	(13.02)
Repub x Text x C.R.	14.77	15.01	9.28	12.82	11.05	4.38	0.97
	(12.15)	(12.26)	(12.82)	(11.79)	(11.90)	(12.44)	(12.64)
Repub x Skit x C.R.	13.11	13.30	10.68	13.70	17.12	10.29	12.97
	(12.14)	(12.26)	(12.94)	(11.79)	(11.89)	(12.58)	(12.75)
Repub x Ad x C.R.	−4.18	−4.13	−9.36	−6.29	−10.21	−12.46	−9.09
	(12.30)	(12.28)	(13.05)	(11.94)	(11.90)	(12.67)	(12.63)
Constant	61.52***	60.20***	61.27***	61.93***	60.64***	64.97***	61.36***
	(2.20)	(2.31)	(2.34)	(5.66)	(4.94)	(5.98)	(5.26)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
N	5,524	5,524	4,895	5,523	5,523	4,894	4,894
R^2	0.28	0.25	0.30	0.33	0.30	0.34	0.31
Adjusted R^2	0.28	0.25	0.29	0.32	0.30	0.34	0.31

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F20: **Models of Ambivalent Sexism and Credibility Confidence in Scandal Clipping in Incidental Exposure Experiment**

	Confidence that clipping was credible [1-5]						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ambivalent Sexism	0.24***	0.15***	0.24***	0.15***	0.08	0.15**	0.10
	(0.05)	(0.05)	(0.06)	(0.05)	(0.05)	(0.06)	(0.06)
Audio	0.50	0.23	0.47	0.54**	0.28	0.51**	0.32
	(0.21)	(0.22)	(0.23)	(0.21)	(0.22)	(0.23)	(0.23)
Text	0.14	−0.11	0.15	0.23	−0.05	0.25	−0.07
	(0.22)	(0.22)	(0.23)	(0.21)	(0.22)	(0.23)	(0.23)
A.S. x Audio	−0.13*	−0.05	−0.12	−0.14	−0.06	−0.12	−0.05
	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)
A.S. x Text	−0.02	0.04	−0.01	−0.04	0.02	−0.03	0.04
	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)
Constant	2.55***	2.86***	2.54***	2.66***	3.10***	2.76***	3.12***
	(0.15)	(0.16)	(0.17)	(0.37)	(0.31)	(0.39)	(0.33)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
N	2,724	2,724	2,419	2,724	2,724	2,419	2,419
R^2	0.02	0.01	0.02	0.06	0.05	0.06	0.06
Adjusted R^2	0.02	0.01	0.02	0.05	0.05	0.05	0.05

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F21: **Models of Ambivalent Sexism and Binarized Credibility Confidence in Scandal Clipping in Incidental Exposure Experiment**

	Somewhat/strongly confident clipping was credible						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ambivalent Sexism	0.06***	0.04**	0.06***	0.03	0.02	0.03	0.03
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Audio	0.08	0.08	0.12	0.08	0.08	0.11	0.14
	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.09)
Text	−0.03	−0.07	−0.02	−0.01	−0.07	0.0002	−0.06
	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.09)
A.S. x Audio	−0.02	−0.02	−0.03	−0.02	−0.01	−0.03	−0.03
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
A.S. x Text	0.01	0.02	0.01	0.01	0.02	0.01	0.02
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Constant	0.25***	0.30***	0.24***	0.09	0.18	0.10	0.16
	(0.06)	(0.06)	(0.06)	(0.14)	(0.12)	(0.15)	(0.13)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
N	2,724	2,724	2,419	2,724	2,724	2,419	2,419
R^2	0.01	0.01	0.01	0.05	0.05	0.05	0.05
Adjusted R^2	0.01	0.01	0.01	0.04	0.04	0.04	0.04

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F22: **Models of Ambivalent Sexism and Scandal Target Affect in Incidental Exposure Experiment**

	Elizabeth Warren Feeling Thermometer						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ambivalent Sexism	−12.49***	−10.44***	−13.12***	−5.36***	−5.27***	−5.62***	−5.79***
	(1.30)	(1.29)	(1.43)	(1.09)	(1.10)	(1.19)	(1.20)
Video	−5.95	−7.72	−2.15	−4.53	−8.02	−2.55	−8.06
	(5.44)	(5.59)	(5.91)	(4.49)	(4.67)	(4.87)	(5.04)
Audio	−9.93*	−8.77	−9.18	−11.34***	−10.08*	−11.08**	−11.23**
	(5.30)	(5.42)	(5.74)	(4.39)	(4.54)	(4.74)	(4.91)
Text	−2.90	0.65	0.18	−7.71	−3.74	−5.41	−4.17
	(5.34)	(5.44)	(5.75)	(4.41)	(4.55)	(4.74)	(4.86)
Skit	−3.46	−10.15	−2.24	−6.04	−14.21***	−5.19	−16.91***
	(5.32)	(5.40)	(5.76)	(4.40)	(4.52)	(4.75)	(4.85)
A.S. x Video	0.66	1.30	−0.20	0.35	1.35	−0.07	1.58
	(1.84)	(1.87)	(2.02)	(1.52)	(1.57)	(1.66)	(1.70)
A.S. x Audio	2.88	2.10	2.95	2.81*	1.93	2.93*	2.51
	(1.81)	(1.81)	(1.97)	(1.49)	(1.51)	(1.62)	(1.64)
A.S. x Text	0.39	−0.49	−0.59	1.82	0.89	1.08	0.90
	(1.83)	(1.82)	(1.97)	(1.51)	(1.53)	(1.63)	(1.63)
A.S. x Skit	0.25	2.42	0.07	0.93	3.73**	0.77	4.53***
	(1.82)	(1.81)	(1.97)	(1.50)	(1.52)	(1.62)	(1.63)
Constant	80.89***	74.99***	81.06***	74.98***	75.34***	74.59***	74.02***
	(3.80)	(3.85)	(4.18)	(6.20)	(5.45)	(6.62)	(5.85)
Weighted?		✓			✓		✓
Low-Quality Dropped?			✓			✓	✓
Controls?				✓	✓	✓	✓
N	4,599	4,599	4,069	4,599	4,599	4,069	4,069
R^2	0.08	0.06	0.10	0.38	0.35	0.39	0.36
Adjusted R^2	0.08	0.06	0.09	0.37	0.34	0.39	0.36

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F23: **Predictors of Detection Task Accuracy**

	Deepfake Detection Accuracy (% Correctly Classified)					
	(1)	(2)	(3)	(4)	(5)	(6)
Digital Literacy		0.25***	0.22***	0.22***	0.20***	0.21***
		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Accuracy Prime	−0.002		−0.01	−0.004	−0.005	−0.002
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)
Exp 1 Debrief			0.01	0.01*	0.01*	0.01
			(0.01)	(0.01)	(0.01)	(0.01)
Exp 1 Information			−0.01	−0.001	−0.01	−0.004
			(0.01)	(0.01)	(0.01)	(0.01)
Political Knowledge			0.18***	0.19***	0.18***	0.20***
			(0.01)	(0.01)	(0.01)	(0.01)
Internet Usage			−0.01	−0.05	0.001	−0.02
			(0.03)	(0.03)	(0.03)	(0.03)
Low-fake Env.			0.03***	0.04***	0.03***	0.05***
			(0.01)	(0.01)	(0.01)	(0.01)
No-fake Env.			0.04***	0.04***	0.04***	0.05***
			(0.01)	(0.01)	(0.01)	(0.01)
Age 65+			0.01	0.01	0.01	0.003
			(0.01)	(0.01)	(0.01)	(0.01)
High School			0.01	0.01	0.03	0.02
			(0.03)	(0.02)	(0.03)	(0.02)
College			0.02	0.02	0.04	0.03
			(0.03)	(0.02)	(0.03)	(0.02)
Postgrad			−0.01	−0.02	0.01	−0.01
			(0.03)	(0.02)	(0.03)	(0.02)
C.R.			0.06***	0.07***	0.07***	0.08***
			(0.02)	(0.02)	(0.02)	(0.02)
C.R. x Republican			−0.06*	−0.06	−0.06	−0.06*
			(0.03)	(0.03)	(0.03)	(0.03)
Ambivalent Sexism			0.001	−0.002	0.001	−0.0001
			(0.004)	(0.004)	(0.004)	(0.004)
Republican			0.09***	0.07***	0.09***	0.08***
			(0.01)	(0.01)	(0.01)	(0.01)
Constant	0.57	0.36***	0.16***	0.19***	0.14**	0.16***
	(0.005)	(0.02)	(0.04)	(0.04)	(0.05)	(0.04)
Weighted?				✓		✓
Low-Quality Dropped?					✓	✓
N	5,497	5,497	5,496	5,496	4,870	4,870
R^2	0.0000	0.02	0.09	0.09	0.09	0.10
Adjusted R^2	−0.0002	0.02	0.09	0.09	0.09	0.10

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F24: **Predictors of Detection Task False Positive Rate (FPR)**

	Detection FPR (% Real Videos Classified as Deepfakes)					
	(1)	(2)	(3)	(4)	(5)	(6)
Digital Literacy		−0.11***	−0.08**	−0.11***	−0.06	−0.08**
		(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Accuracy Prime	−0.01		0.004	−0.01	0.003	−0.01
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)
Exp 1 Debrief			−0.03***	−0.03***	−0.03***	−0.03***
			(0.01)	(0.01)	(0.01)	(0.01)
Exp 1 Information			0.01	0.001	0.01	0.005
			(0.01)	(0.01)	(0.01)	(0.01)
Political Knowledge			−0.13***	−0.14***	−0.13***	−0.14***
			(0.02)	(0.02)	(0.02)	(0.02)
Internet Usage			−0.01	0.01	−0.01	0.01
			(0.03)	(0.03)	(0.04)	(0.04)
Low-fake Env.			0.03**	0.02**	0.02***	0.01
			(0.01)	(0.01)	(0.01)	(0.01)
No-fake Env.			0.23***	0.22***	0.22***	0.21***
			(0.01)	(0.01)	(0.01)	(0.01)
Age 65+			0.003	−0.003	0.01	−0.0002
			(0.01)	(0.01)	(0.01)	(0.01)
High School			0.002	−0.004	−0.02	−0.02
			(0.03)	(0.02)	(0.04)	(0.02)
College			0.01	0.01	−0.02	−0.01
			(0.03)	(0.02)	(0.04)	(0.02)
Postgrad			0.04	0.02	0.002	0.0000
			(0.04)	(0.02)	(0.04)	(0.02)
C.R.			−0.06**	−0.08***	−0.07***	−0.08***
			(0.02)	(0.02)	(0.02)	(0.02)
C.R. x Republican			0.03	0.03	0.04	0.03
			(0.03)	(0.03)	(0.03)	(0.03)
Ambivalent Sexism			0.0005	0.002	−0.002	0.001
			(0.005)	(0.005)	(0.005)	(0.005)
Republican			−0.07***	−0.07***	−0.08***	−0.07***
			(0.01)	(0.01)	(0.01)	(0.01)
Constant	0.28***	0.37***	0.41***	0.44***	0.42***	0.45***
	(0.01)	(0.02)	(0.05)	(0.04)	(0.05)	(0.05)
Weighted?				✓		✓
Low-Quality Dropped?					✓	✓
N	5,495	5,495	5,494	5,494	4,869	4,869
R^2	0.0002	0.003	0.16	0.16	0.16	0.16
Adjusted R^2	−0.0000	0.003	0.15	0.16	0.15	0.16

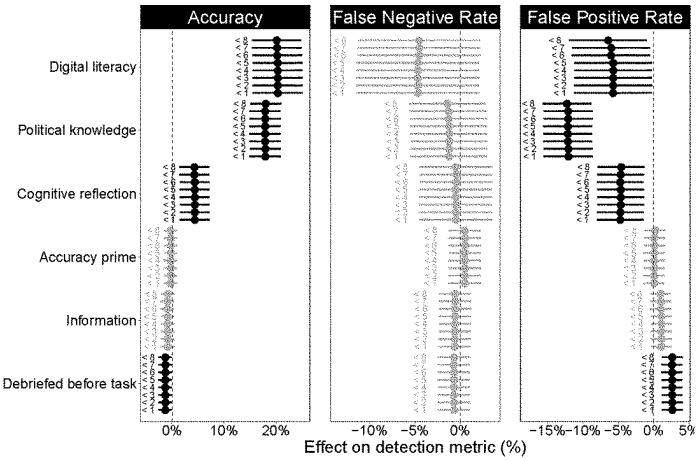
Notes: $^{\cdot}$ $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Table F25: **Predictors of Detection Task False Negative Rate (FNR)**

	Detection FNR (% Deepfakes Classified as Real Videos)					
	(1)	(2)	(3)	(4)	(5)	(6)
Digital Literacy		−0.03	−0.04	−0.01	−0.05	−0.02
		(0.03)	(0.03)	(0.03)	(0.03)	(0.04)
Accuracy Prime	−0.01		−0.005	−0.01	0.003	−0.003
	(0.01)		(0.01)	(0.01)	(0.01)	(0.01)
Exp 1 Debrief			0.01	0.01	0.01	0.01
			(0.01)	(0.01)	(0.01)	(0.01)
Exp 1 Information			−0.003	0.003	−0.01	−0.003
			(0.01)	(0.01)	(0.01)	(0.01)
Political Knowledge			0.003	−0.03	−0.002	−0.03
			(0.02)	(0.02)	(0.02)	(0.02)
Internet Usage			0.11**	0.06	0.13***	0.09
			(0.04)	(0.04)	(0.04)	(0.04)
Low-fake Env.			0.01	0.002	0.01	−0.003
			(0.01)	(0.01)	(0.01)	(0.01)
No-fake Env.			0.01	0.02	0.02	0.02*
			(0.01)	(0.01)	(0.01)	(0.01)
Age 65+			0.01	0.02	0.01	0.01
			(0.04)	(0.03)	(0.04)	(0.03)
High School			0.03	0.04	0.03	0.03
			(0.04)	(0.03)	(0.04)	(0.03)
College			0.09*	0.13***	0.08	0.11***
			(0.04)	(0.03)	(0.05)	(0.03)
Postgrad			−0.05	−0.04	−0.06**	−0.04
			(0.02)	(0.03)	(0.03)	(0.03)
Republican			0.09*	0.07	0.11**	0.10**
			(0.04)	(0.04)	(0.04)	(0.04)
C.R.			0.02***	0.02***	0.01	0.01
			(0.01)	(0.01)	(0.01)	(0.01)
Ambivalent Sexism			−0.04***	−0.03	−0.05***	−0.04**
			(0.01)	(0.02)	(0.02)	(0.02)
C.R. x Republican	0.34***	0.36***	0.17**	0.21***	0.20***	0.24***
	(0.01)	(0.03)	(0.06)	(0.06)	(0.07)	(0.06)
Weighted?				✓		✓
Low-Quality Dropped?					✓	✓
N	3,690	3,690	3,690	3,690	3,266	3,266
R^2	0.0002	0.0003	0.02	0.03	0.02	0.02
Adjusted R^2	−0.0000	0.0000	0.02	0.02	0.01	0.02

Notes: · $p \cdot r/K < .1$ * $p \cdot r/K < .05$ ** $p \cdot r/K < .01$ *** $p \cdot r/K < .001$

Figure F24: **Sensitivity of Predictors of Detection Experiment Performance to Non-Response Thresholding**



Notes: Each estimate is the effect of the corresponding predictor estimated from a model with full controls (see tables for detection task results) excluding respondents with $< x$ number of videos completed in the detection task.

285

G Exploratory Analyses

Figure G25: **Baseline Comparisons of Credibility and Affective Response of Video in Incidental Exposure Experiment**

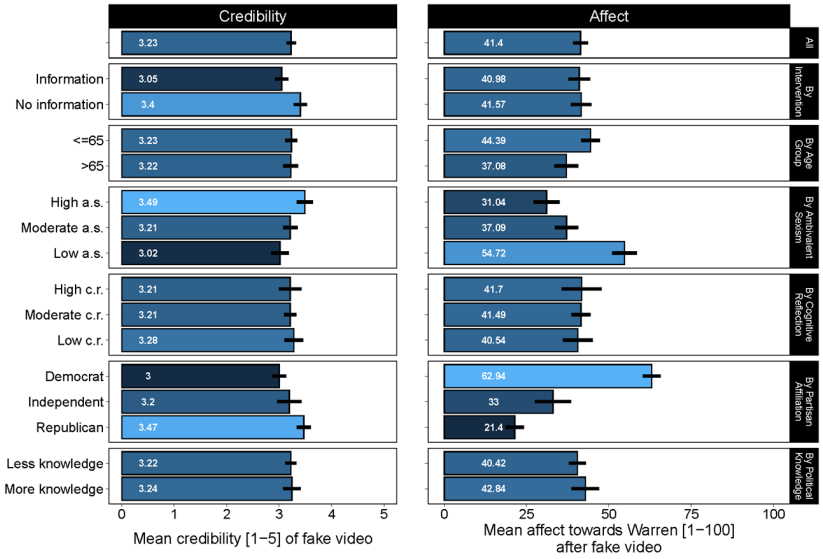
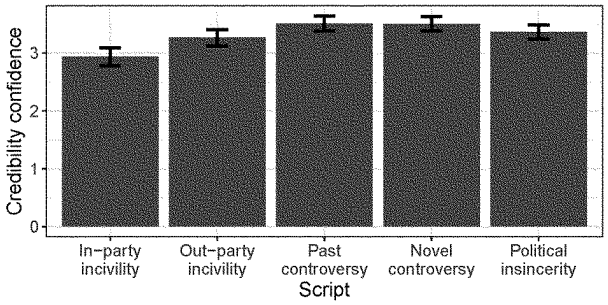
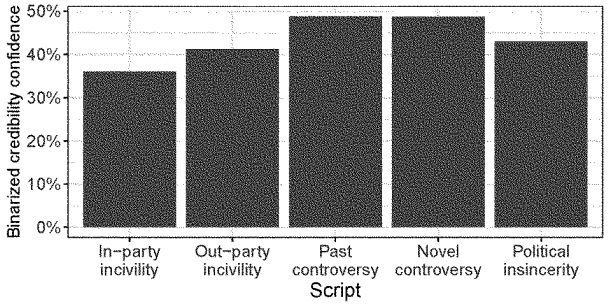


Figure G26: **Heterogeneity in Incidental Exposure Credibility by Scandal Script**

(a) **Categorical Response**



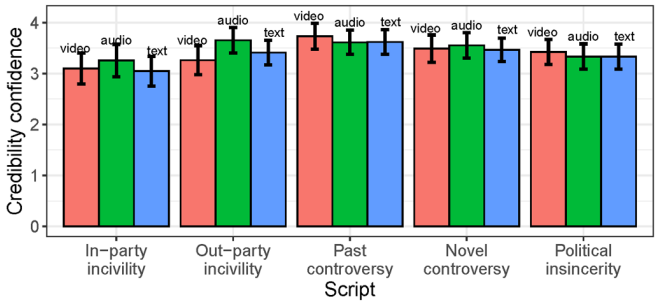
(b) **Binarized Response**



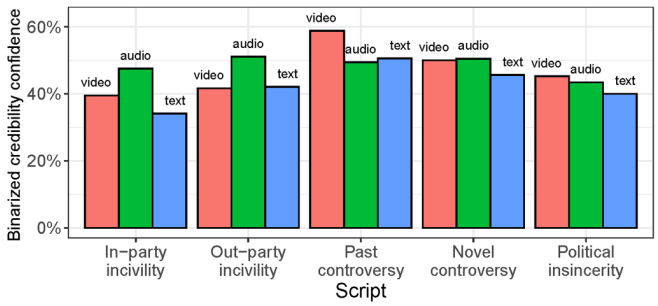
Notes: Results from the subset of respondents exposed to a scandal stimuli, not assigned an information treatment, and who provided a response to our deception question (n=1848).

Figure G27: **Heterogeneity in Credibility Perception by Scandal Script and Medium**

(a) **Categorical Response**

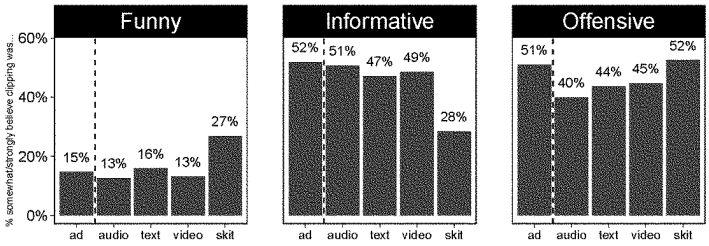


(b) **Binarized Response**



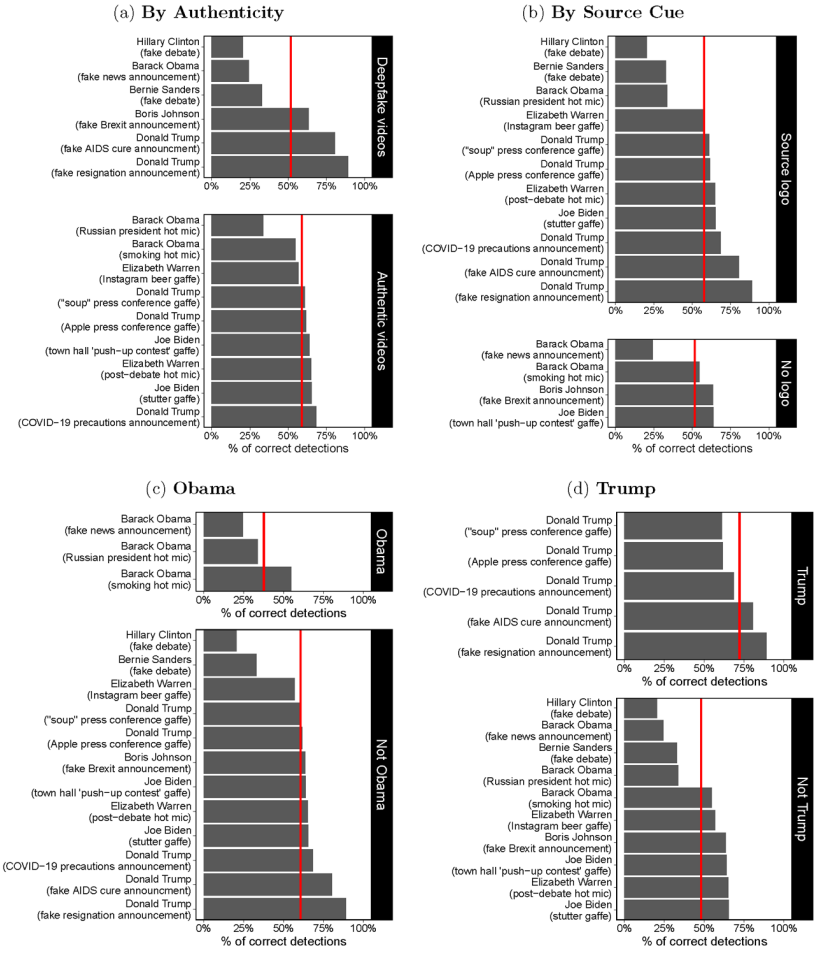
Notes: Results from the subset of respondents exposed to a scandal stimuli, not assigned an information treatment, and who provided a response to our credibility question (n=1,848). To reduce the number of experimental cells, only three of five scripts were used for the `skit` stimuli.

Figure G28: **Other Affective Responses to First-Stage Clip**



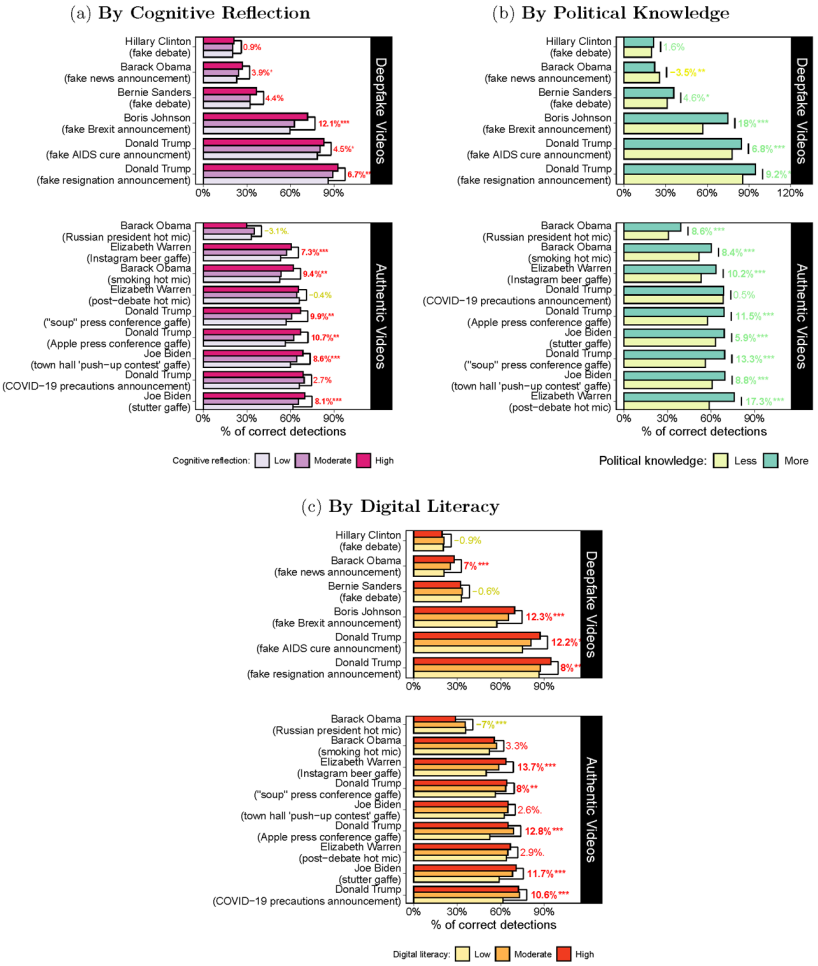
Notes: Responses evaluating whether clip was "funny," "informative," or "offensive" were solicited alongside belief that clip was not fake or doctored. The attack ad condition excluded since it is not a directly comparable clip of the scandal.

Figure G29: **Detection Task Performance for Specific Clips**



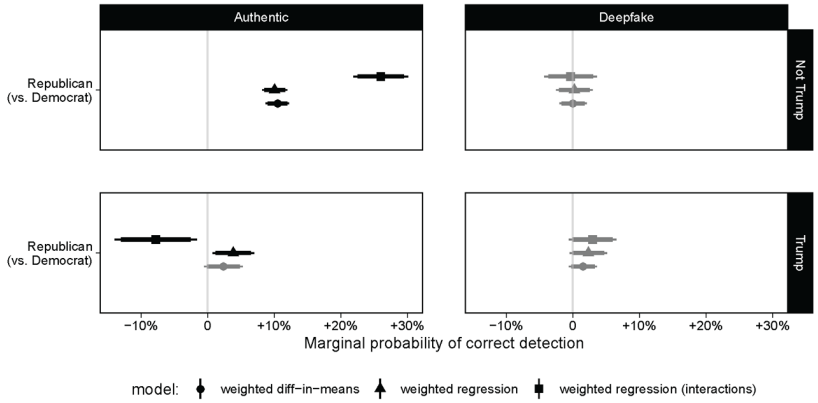
Notes: Results are for n =5,497 (99%) of respondents who provide a response to at least one video in the detection experiment. Fake clips are detected less well than real clips, but this difference (Δ) is not significant according to a t-test ($\Delta = -7.20\%, t = 0.57, p = 0.58$). Clips without source outlet logos are detected less well than clips with source logos, but this difference is also not significant ($\Delta = -6.03\%, t = 0.53, p = 0.61$).

Figure G30: **Detection Task Performance for Specific Clips by Subgroup**



Notes: Results are for n =5,497 (99%) of respondents who provide a response to at least one video in the detection experiment. Cognitive reflection and digital literacy categories constructed as equal-sized quartiles.

Figure G31: **Comparison of Partisan Group Identity Effects on Detection Task Performance**



Notes: Results are for n =5,497 (99%) of respondents who provide a response to at least one video in the detection experiment. All models incorporate weights from post-stratification described in Appendix E. The regression controls for the characteristics described at the start of this section as well as whether the clip contains a source logo or not. The interaction model fits an interaction term of partisanship with an indicator for particular clip within each cell.

Figure G32: **Relationship Between Credibility of Video Clip and Credibility of Event**

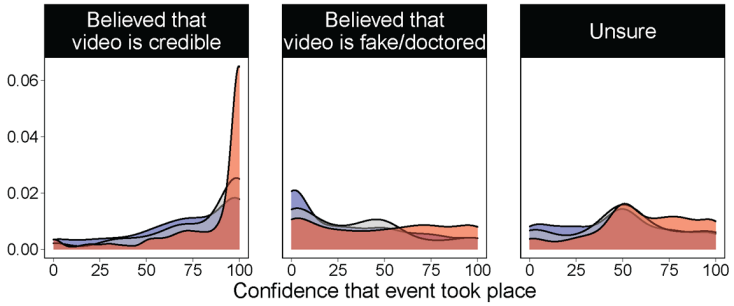
(a) **Barack Obama (Russian president hot mic)**

(b) **Doanld Trump (COVID-19 precautions announcement)**

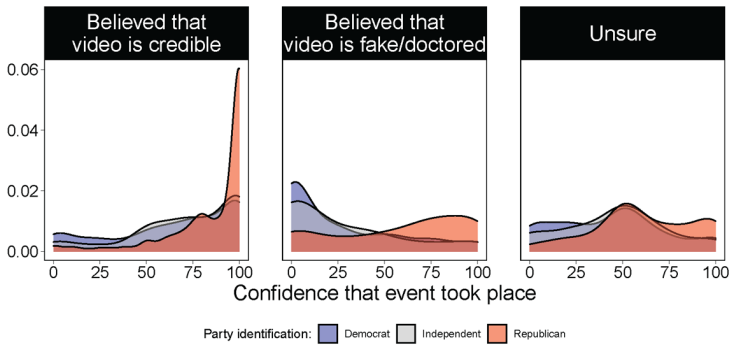
Notes: Clips (a) and (b) are real videos. Results are for n =5,497 (99%) of respondents who provide a response to at least one video in the detection experiment. A variety of regression specifications estimate large, robust and statistically significant positive relationship between a respondent's belief in the video's authenticity and confidence in the depited event's occurence.

Figure G33: **Relationship Between Credibility of Video Clip and Credibility of Event by Partisanship**

(a) **Barack Obama (Russian president hot mic)**



(b) **Doanld Trump (COVID-19 precautions announcement)**



Notes: Clips (a) and (b) are real videos. Results are for n =5,497 (99%) of respondents who provide a response to at least one video in the detection experiment.

Senate Committee on Rules and Administration
AI and the Future of our Elections
September 27, 2023
Questions for the Record
The Honorable Steve Simon

Chairwoman Klobuchar

Election officials from both parties have faced a barrage of threats and harassment from those seeking to undermine our democracy.

- How could AI technology be used to intensify the volume and sophistication of threats targeting election workers?

  In Minnesota we are concerned about this issue from both a physical and cybersecurity standpoint. It's no secret that election workers, both those who serve on election day in polling places and those who serve year-round as election administrators, have increasingly been the target of threats, harassment, and intimidation. Such incidents in Minnesota have thankfully been isolated so far, and I want to keep it that way.

  Rapidly advancing AI technologies, like chatbots based on large language models and software programs that more easily than ever modify video and audio, will likely strengthen the efforts to mislead the public and to put individuals such as election workers at risk. Deepfake videos intended to sow distrust in elections (and in election administrators), as well as audio or phishing attacks aimed at stealing passwords and gaining access to key systems, are increasingly sophisticated and harder to identify. The new technologies could target larger audiences with authentic-seeming election disinformation of the kind that has sometimes inspired threats against election workers.

  Consider a recent and prominent example. Two Georgia election workers were falsely accused of illegally moving or tampering with boxes of ballots on election night in 2020 – based on video depictions that were heavily edited and/or taken out of context. Multiple investigations exonerated them completely, but the harm to their safety and to their reputations persisted. The duo famously testified before the U.S. House January 6th Committee and have filed a defamation suit against those who they claim knowingly spread lies about them.

  In 2024, AI could make such smears easier than ever before. In lieu of merely editing an actual video, AI could *create* a deepfake video purporting to show one or more election workers committing illegal acts. A completely synthetic image, with a veneer of authenticity, could be seen by millions of people. The resulting public outrage could place the targeted election workers, and perhaps even election workers generally, in jeopardy.

Senator Fischer

1. I understand that Minnesota recently criminalized the use of deepfake AI technology to influence elections. This law contains an exception for fake videos or audio created by individuals who have the ability to physically or verbally impersonate someone.

- Why—practically speaking, and especially in the audio context—should a convincing fake generated by AI be treated differently than a convincing fake created by an impersonator?

  Our office was not involved in the drafting or passage of this new law, so I am not able to comment on why the bill authors chose to make this distinction. Perhaps the provision relates to the generally recognized exception for satire or parody.

- Has your office documented any instance where deepfakes have influenced the result of an election in Minnesota?

  This new law regarding deepfakes went into effect on July 1, 2023. We are just now entering the first election period when this law is in operation, and to our knowledge no charges have been brought under this new statute. Additionally, it is important to note that the Office of the Secretary of State does not have any authority or involvement related to campaign practices.

Senate Committee on Rules and Administration
AI and the Future of our Elections
September 27, 2023
Questions for the Record
The Honorable Trevor Potter

Chairwoman Klobuchar

At the hearing we discussed the need to prohibit the most egregious examples of fraud using AI-generated content in our elections, and how a ban like the one in my bipartisan bill would do that within the framework of the Constitution. We also discussed the strong Supreme Court precedent supporting the constitutionality of disclaimers in our elections.

- Can you explain why this legislation is on solid constitutional ground?

The Supreme Court has held that fraudulent and deceptive statements have no value under the First Amendment. *See Ill., ex rel. Madigan v. Telemarketing Assocs., Inc.*, 538 U.S. 600, 612 (2003) (stating that "public deception is "unprotected speech" and "the First Amendment does not shield fraud"). Particularly in regard to AI-based content that is intended to deceive voters and undermine elections, it is unprotected by the First Amendment; as noted constitutional scholar Professor Rick Hasen put it in a recent article, "There is no First Amendment right to use speech to subvert an election, any more than there is a First Amendment right to use speech to bribe, threaten, or intimidate." Richard L. Hasen, *U.S. v. Trump Will Be the Most Important Case in Our Nation's History*, Slate (Aug. 1, 2023), https://slate.com/news-and-politics/2023/08/trump-trial-2024-historic-jack-smith-indictment.html.

In fact, the Court has recognized that the government has a "firmly established" interest in "protect[ing] people against fraud." *Donaldson v. Read Magazine*, 333 U.S. 178, 190 (1948). Indeed, the federal statute prohibiting fraudulent misrepresentation of campaign authority, 52 U.S.C. § 30124, has for decades prohibited candidates from misrepresenting that campaign speech was authored by another candidate on a matter that is damaging to that candidate.

Accordingly, a narrow law prohibiting the distribution of fraudulent or deceptive AI-generated content intended to influence an election or solicit funds would find support within our existing constitutional framework.

In addition, election disclaimers find strong and clear support in constitutional law. As recently as *Citizens United*, the Supreme Court has upheld disclaimer requirements, remarking that disclaimers "impose no ceiling on campaign-related activities," "do not prevent anyone from speaking," and fulfil the government's important interest in "providing the electorate with information" to "make informed choices in the political marketplace." *Citizens United v. FEC*, 558 U.S. 310, 366-67 (2010) (internal quotation marks and alteration omitted); *see id.* at 371 ("[T]ransparency enables the electorate to make informed decisions and give proper weight to different speakers and messages."); *McConnell v. FEC*, 540 U.S. 93, 197 (2003) (recognizing

"the [ ] First Amendment interests of individual citizens seeking to make informed choices in the political marketplace" (internal quotation marks omitted)); *Buckley v. Valeo*, 424 U.S. 1, 66-67 (1976) ("[D]isclosure provides the electorate with information . . . in order to aid the voters in evaluating those who seek federal office."); *Human Life of Wash., Inc. v. Brumsickle*, 624 F.3d 990, 1005, 1008 (9th Cir. 2010) ("Providing information to the electorate is vital to the efficient functioning of the marketplace of ideas, and thus to advancing the democratic objectives underlying the First Amendment. . . . Campaign finance disclosure requirements thus advance the important and well-recognized governmental interest of providing the voting public with the information with which to assess the various messages vying for their attention in the marketplace of ideas.").

Senate Committee on Rules and Administration
AI and the Future of our Elections
September 27, 2023
Questions for the Record
Ms. Maya Wiley

Chairwoman Klobuchar

In your testimony you noted how AI could significantly amplify the volume and spread of election-related disinformation in the 2024 elections, such as false information about voting hours or locations.

- Can you elaborate on the risks posed by AI-fueled disinformation to people's ability to cast a ballot, as well as what should be done to address these threats?

Maya Wiley

The rapid growth of AI since the 2022 elections has the potential to turbocharge the volume and speed of voting disinformation and hate speech. The potential for the rampant false content and rapid spread of disinformation is alarming:

- The false narratives we have seen in prior elections are certain to come up again in 2024 in new forms with even more intensity.

- Generative AI (GAI) can increase the ability and scale to spread false information and propaganda, leaving voters confused, further questioning what they see and hear, and purposefully misled about critical information required to cast a ballot and have that ballot count.

- As our coalition partner the Brennan Center has pointed out, elections are particularly vulnerable to AI disinformation. GAI tools are most effective when they produce content that is similar to their current databases. Since the same false election narratives will likely come up again, the underlying election disinformation in the training data underlying GAI tools can make them a time bomb for future election disinformation.

- The proliferation of AI-generated content could accelerate the loss of trust in the integrity and security of our overall election system and dramatically interfere with the right to vote.

Platforms and tech companies must have policies, systems, and guardrails to stem the disinformation that can result from generative AI. Platforms can devote more resources to identifying and removing coordinated bots and labeling deepfakes that could influence elections.

The federal government can play a strong enforcement and prevention role, by detecting and redressing AI disinformation used to target communities of color, including the Department of

Justice and the Department of Homeland Security, and training state and local election officials on the potential dangers and harmful impact of AI on our election processes.

In addition, as our partner Public Knowledge has noted, a whole of society approach is needed to address the issues presented by GAI and restore trust in our information environment.

This can include policymakers creating incentives for the technology platforms to change their policies and product design and they should foster more competition and choice among media outlets. Civil society should convene stakeholders, including from the communities most impacted by misinformation, to research and design while protecting privacy and freedom of expression.

Senate Committee on Rules and Administration
AI and the Future of our Elections
September 27, 2023
Questions for the Record
Mr. Neil Chilson

Senator Fischer

1. **You noted in your testimony that generative AI may not change the cost structure meaningfully for those seeking to interfere in elections, compared to other means.**

    - **Could you expand on why you believe the threat from AI may be relatively small compared to other technologies or tactics?**

Generative AI technologies promise to make it cheaper and easier for users to create certain kinds of content, including false and deceptive content. However, the cost of creating false content is only one cost incurred by those who produce misinformation, and it is not the largest cost. Distribution of content is the primary pain point, and generative AI technologies do not change the costs of distribution.

For those who seek to affect elections through disinformation campaigns, content creation is not the primary bottleneck.

It is already easy to create misinformation and disinformation without AI assistance. It is so easy that it is done many thousands of times a day for free online. Take a specific example: consider the much reported on video of Nancy Pelosi doctored to make it appear as though she is slurring her words.[1] No generative AI was used to create that content: all it took was selectively slowing down the video. Such "cheap fakes" remain plausible and widely available using conventional editing tools. Because such methods are inexpensive and effective, parties seeking to create doubt about the integrity of the overall electoral process are likely to prefer them to more complicated and elaborate content creation methods using AI.

Consider another recent example: the competing narratives around an October 17, 2023, explosion near a hospital in Gaza. The early claims that it was a result of an Israeli strike were embraced as credible by major mainstream news media organizations.[2] Subsequent information has brought that conclusion into serious doubt.[3] But early reporting already generated mass protests across the world and led to the cancelation of diplomatic meetings between the U.S. and some Middle East

---

[1] *See*, Zack Budryk, *Fake video edited to make Pelosi appear drunk spread on social media* (May 23, 2019), The Hill, https://thehill.com/policy/cybersecurity/445311-fake-videos-spread-on-social-media-edited-to-make-pelosi-appear-drunk/.

[2] *See*, Patrick Kingsley, *et al*, *Israeli Strike Kills Hundreds in Hospital, Palestinians Say* (Oct. 17, 2023), New York Times, https://www.nytimes.com/2023/10/17/world/middleeast/gaza-hospital-explosion-israel.html. The headline has since changed.

[3] *See* Katie Robertson, *After Hospital Blast, Headlines Shift With Changing Claims* (Oct. 18, 2023), New York Times, https://www.nytimes.com/2023/10/18/business/media/hospital-blast-gaza-reports.html

countries.[4] All without AI-generated content; the misinformation relied entirely on framing of claims around the incident.

These examples demonstrate that the biggest hurdle to effective disinformation is spreading it, not creating it. Those seeking to disrupt or discredit elections are likely to spend their time and resources building or contracting mainstream media and social media influence operations that can drive content viral.[5] Reductions in cost of content creation will only marginally increase the amount available to invest in spreading content. More credible-looking content might be expected to spread more easily. But virality is frequently more a function of the network of proponents than the quality of the content. Hamas' successful effort was enabled by its ability to connect with mainstream media outlets, not by the quality of its evidence.

Understanding the relative incentives of those who would spread disinformation suggests that targeting generative tools will not have much benefit even as it imposes significant costs. Targeting the tool (AI) rather than the conduct (creating deception or misinformation) is both over and underinclusive. It will burden speech that has nothing to do with misinformation and could in fact help to combat misinformation. Yet it will not affect the cheap fakes and other standard misinformation techniques that already exist. And it does nothing to constrain distribution of misinformation, which is the true bottleneck for such efforts.

2. **I understand that much of the concern about AI stems from the low cost and barrier of entry to producing convincing fake images, video, and audio. I also recognize that there is concern that the internet can be used to quickly distribute that kind of content.**

    - **As a matter of law, should there be a difference between a convincing fake photo produced without the help of generative AI, and a convincing fake photo produced using generative AI? What is the reasoning for your conclusion?**

There is not and should not be a difference in how the law treats a convincing fake photo generated with or without generative AI. The defamatory or deceptive nature of a fake photo does not hinge on the method through which that photo is created. Defamation law varies slightly state-by-state, but the four core elements do not include any reference to the method of content creation.[6] Similarly, in evaluating commercial deception, the Federal Trade Commission's Section 5 standard does not consider the method of manufacture of the deceptive content.[7]

Nor *should* legal conclusions differ based on the method of creation, because the proper goal of such laws is to prevent or address the harm caused regardless of the method of manufacture. Even assuming that all AI generated content is easier to produce and distribute than content generated through other means, the reason that is problematic is because it increases the harm caused. Thus, focusing on harm addresses such concerns about AI without singling out any one technique for content creation for additional scrutiny.

---

[4] Margherita Stancati et al., *U.S., Experts Say Evidence Suggests Palestinian Militants' Rocket Hit Gaza Hospital* (updated Oct. 18, 2023), Wall Street Journal, https://www.wsj.com/world/middle-east/israel-tries-to-back-up-claims-it-didnt-attack-gaza-hospital-a8cc3405.

[5] *See*, Max Fiser, *Disinformation for Hire, a Shadow Industry, Is Quietly Booming* (July 25, 2021), The New York Times, *available at* https://www.nytimes.com/2021/07/25/world/europe/disinformation-social-media.html.

[6] *See, e.g.*, *Dillon v City of New York*, 261 AD2d 34 at 38 (1999) (setting out the elements of New York State defamation law).

[7] *See*, FTC Policy Statement on Deception (Oct. 14, 1983) (appended to *Cliffdale Associates, Inc.*, 103 F.T.C. 110, 174 (1984)), available at https://www.ftc.gov/legal-library/browse/ftc-policy-statement-deception.

This matters because a technology-specific standard would penalize certain speech creation technologies over others, untethered from the harm caused. A defamatory statement created manually would receive less scrutiny or fewer penalties than a defamatory statement created through AI, even if the defamatory effect was the same. Such an approach would be both over- and underinclusive, likely penalizing non-defamatory AI-generated content while allowing traditionally generated lies or defamatory statements to skate free.

Senate Committee on Rules and Administration
AI and the Future of our Elections
September 27, 2023
Questions for the Record
Mr. Ari Cohn

Chairwoman Klobuchar

At the hearing you suggested that a "narrowly drawn" requirement for disclaimers for political ads including AI-generated content could further the public interest, noting that "[if] everything has a disclosure, nothing has a disclosure."

- Can you expand on how a disclaimer for AI-generated content in political ads could provide important information for voters, as well as what you view as the key considerations in crafting such a requirement?

As a normative proposition, a better-informed electorate is surely in the public interest. And it is true that disclaimer and disclosure requirements are preferable to outright regulation of the content of political speech; such requirements "impose no ceiling on campaign-related activities and do not prevent anyone from speaking." *Citizens United v. FEC*, 558 U.S. 310, 366 (2010) (internal citations and quotation marks omitted).

But while such requirements potentially impose a smaller burden on expression, they nevertheless do impose *some* burden and therefore receive a measure of First Amendment scrutiny. Mandatory disclosure laws must satisfy "exacting scrutiny," that is, they must be substantially related to a sufficiently important government interest and narrowly tailored to that interest (though they need not be the least restrictive means of achieving the interest). *See Ams. For Prosperity Found. v. Bonta*, 141 S. Ct. 2373, 2383–84 (2021).

Thus, the first consideration must be a proper definition of the problem, and its scope. The Supreme Court has held that "provid[ing] the electorate with information" is a sufficient government interest in some circumstances. *Citizens United*, 558 U.S. at 367. However, those circumstances have, to date, related to identifying the *speaker*, rather than anything substantive about the speech itself. Campaign advertisements produced with the assistance of generative AI must already bear such disclaimers as mandated by federal law and regulations. Thus, an AI disclaimer does not further the electorate's interest in knowing who is speaking to it.

To some extent, the purpose behind identifying the speaker is to enable recipients to assess the credibility of the message. *See Center for Individual Freedom v. Madigan*, 697 F.3d 464, 478 (7th Cir. 2012) ("[T]he identity of a candidate's supporters—and opponents—is information that the voting public values highly. In areas of inquiry where logic or exact observation is unavailing, a speaker's credibility often depends crucially on who he is."). This informational interest has been described most often as providing two benefits: (a) a heuristic that allows voters to infer a candidate's positions and alignment, and (b) gauging candidate integrity. *See generally* Lear Jiang, *Disclosure's Last Stand? The Need to Clarify the "Informational Interest" Advanced By Campaign Finance Disclosure*, 119 COLUM. L. REV. 487, 512–14 (2019) (providing the example of two candidates who campaign on greater financial regulation and the usefulness of knowing which candidates took money from investment banks and might therefore be less likely to keep their campaign promises).

But AI disclaimers do not serve this informational interest as traditionally conceived. Rather than providing information about the speaker to aid in evaluating the messenger's credibility, such disclaimers would instead seek to provide information about content of the message itself—ostensibly to help recipients gauge whether that message is true or false. The difference between these two informational interests is subtle, but important. Though the government may have an interest in assuring that the electorate can effectively assess the credibility of the speaker, its interest in providing information pertaining to the truth or falsity of electoral speech is necessarily lower. As the Court explained:

> Moreover, the people in our democracy are entrusted with the responsibility for judging and evaluating the relative merits of conflicting arguments. They may consider, in making their judgment, the source and credibility of the advocate. But if there be any danger that the people cannot evaluate the information . . . ., it is a danger contemplated by the Framers of the First Amendment.

First Nat'l Bank of Boston v. Belotti, 435 U.S. 765, 791–92 (1978). Indeed, the Court rejected Florida's asserted interest in providing the electorate with more information on which to judge the substance of campaign speech and struck down a law requiring newspapers to print replies from a candidate attacked in their pages. *Miami Herald Publishing Co. v. Tornillo*, 418 U.S. 241 (1974); *see also McIntyre v. Ohio Elections Comm'n*, 514 U.S. 334, 348 (1995) (The "simple interest in providing voters with additional relevant information does not justify a state requirement that a writer make statements or disclosures that she would otherwise omit.").Similarly, FCC licensees are prohibited from rejecting campaign advertisements based on their truth or falsity. *See* 47 U.S.C. § 315(a) ("[L]icensee shall have no power of censorship over the material broadcast under the provisions of this section."); *see also* Domenico Montanaro, *The truth in political advertising: 'You're allowed to lie'*, NPR (Mar. 17, 2022), https://www.npr.org/2022/03/17/1087047638/the-truth-in-political-advertising-youre-allowed-to-lie.

This is not to say that it is *impossible* to establish a sufficiently important government interest to withstand scrutiny. It is theoretically possible that generative AI poses a significant enough risk to maintain intervention. But to establish such an interest, it is critical that Congress develop a record identifying and supporting this interest. To that extent, the input of experts in political science, psychology, technology, and other relevant fields is essential to establishing and demonstrating the risk factors of generative AI, including but not limited to the likelihood of voters being actually misled and swayed to change their voting behavior.

Were the courts to find a sufficient government interest in preventing the electorate from misleading (or perhaps *especially misleading*) campaign ads, any regulation must be narrowly tailored to achieving that interest.

One pressure point in such tailoring will be whether there is any reason why advertisements utilizing generative AI deserve distinct treatment from advertisements containing deceptively-edited media that does *not* utilize AI. *See* AI and the Future of our Elections: Hearing Before the Senate Committee on Rules and Administration (2023) (Statement of Ari Cohn) at pages 2–3 (providing examples of non-AI deceptively edited media by candidates and elected officials). This problem could be resolved by either broadening the regulation to include *all* deceptively edited media no matter the means of production, or perhaps by clearly and persuasively articulating the heightened risks of AI-

generated media over manually-edited media that splices together video and/or audio clips to create a misleading impression.

The other primary pressure point is distinguishing between the various uses of AI. Various forms of AI are used in an extraordinary percentage of media productions; color correction, noise reduction, background object removal, caption text, and a large variety of other basic production tasks enlist the use of AI. Requiring disclaimers for such uses will obviously not further the government interest in protecting the electorate from misleading advertisements and result in a law that is not narrowly tailored. *See Shelton v. Tucker*, 364 U.S. 479, 488 (1960) (finding a lack of narrow tailoring when a state law required teachers to disclose "every conceivable kind of associational tie," many of which "could have no possible bearing upon a teacher's occupational competence or fitness").

Drawing lines may prove to be difficult, but it is necessary so as not to chill the use of technology that could, in reality, *increase* the information provided to voters. One avenue may be to draft language that seeks to differentiate between "deepfakes" and other uses of AI. But such a line should skew towards allowing as much speech as possible. A video edited to reflect a locale that a candidate was unable to visit is certainly not as deceptive as a video that might put words in a candidate's mouth that are opposite to that candidate's actual position. If any line is to be drawn, the latter is certainly more objectionable than the former. Moreover, it may be advisable to impose a narrow temporal scope—disclaimers of AI-generated advertisements may be more defensible if they apply within a short timeframe before an election, where the capacity for counterspeech may be reasonably limited.

This committee's attention to cutting-edge, vital issues is commendable, and the concerns about increasingly believable disinformation are not without merit. But it is essential that any legislation draw lines carefully in order to impact only the most high-risk activity, in as narrow a way possible, to protect the balance that the First Amendment demands—and ultimately, our democracy as a whole.

○