

10-2024

Defamation or Impersonation? Working Towards a Legislative Remedy for Deepfake Election Misinformation

John Thayer
William & Mary Law School

Follow this and additional works at: <https://scholarship.law.wm.edu/wmlr>



Part of the [Election Law Commons](#), [Internet Law Commons](#), and the [Science and Technology Law Commons](#)

Repository Citation

John Thayer, *Defamation or Impersonation? Working Towards a Legislative Remedy for Deepfake Election Misinformation*, 66 Wm. & Mary L. Rev. 251 (2024), <https://scholarship.law.wm.edu/wmlr/vol66/iss1/7>

Copyright c 2024 by the authors. This article is brought to you by the William & Mary Law School Scholarship Repository.
<https://scholarship.law.wm.edu/wmlr>

DEFAMATION OR IMPERSONATION? WORKING TOWARDS A
LEGISLATIVE REMEDY FOR DEEPFAKE ELECTION
MISINFORMATION

TABLE OF CONTENTS

INTRODUCTION	252
I. BACKGROUND	255
<i>A. A (Brief) Primer on Deepfake Technology and Its Use in Campaigns</i>	<i>255</i>
<i>B. The Harm of Campaign Deepfakes</i>	<i>257</i>
<i>C. Statutory Overview: Existing Deepfake Regulations</i>	<i>262</i>
<i>D. Defamation Law</i>	<i>264</i>
<i>E. Impersonation Law</i>	<i>265</i>
II. ANALYZING DEEPFAKES WITHIN THE DEFAMATION AND IMPERSONATION FRAMEWORKS	266
<i>A. Do Deepfakes Defame Their Targets?</i>	<i>266</i>
<i>B. Has a Deepfake Target Been Impersonated?</i>	<i>268</i>
III. MARRYING DEFAMATION LAW AND IMPERSONATION LAW TO CREATE A REMEDY FOR ELECTION DEEPFAKES	269
<i>A. A Mechanical Crafting of the Proposed Remedy</i>	<i>269</i>
<i>B. Constitutional Analysis of the Proposed Remedy</i>	<i>271</i>
IV. COUNTERARGUMENTS	272
<i>A. Elevated Status for Political Speech</i>	<i>272</i>
<i>B. Elevated Status for Public Figures</i>	<i>275</i>
<i>C. Failure to Regulate Positive, yet Deceptive, Candidate Deepfakes</i>	<i>275</i>
<i>D. Alternative Solutions</i>	<i>276</i>
CONCLUSION	278

INTRODUCTION

Suppose a fake video circulated around social media platforms featuring a prominent Democratic politician stating that “allowing Republicans to vote could threaten the integrity of the election and the safety of the electorate” and that “allowing them to vote could lead to an outcome that does not accurately reflect the will of the people.” The clip ends with the politician stating that “it is necessary to restrict Republican voting in the 2024 election.” Suppose now that a different post—this time in the form of a tweet from a Republican presidential primary candidate’s campaign—shows fake, yet entirely realistic, photographs depicting the front-runner for that primary embracing and kissing a government official largely seen as an enemy to the party’s base.

It does not require a major leap to assume that both pieces of content could potentially outrage and inflame their target audiences and drastically affect public opinion of the individuals featured in the posts. It also does not require a major leap to envision that these scenarios could occur with our current technological capacity, because they are not fictional. The first example came from an unknown individual with a social media account and widely accessible technological tools to create a deepfake of Massachusetts senator Elizabeth Warren.¹ The second example came from the Twitter account @DeSantisWarRoom, the official rapid response operation for the 2024 presidential campaign of Florida governor Ron DeSantis, and depicted former president Trump in a fake interaction with Dr. Anthony Fauci.² Neither piece of content contained any sort of disclosure noting its fabrication, and neither resulted in any sort of legal remedy or imposition of liability. This

1. Aleks Phillips, *Deepfake Video Shows Elizabeth Warren Saying Republicans Shouldn't Vote*, NEWSWEEK (Feb. 27, 2023, 12:11 PM), <https://www.newsweek.com/elizabeth-warren-msnbc-republicans-vote-deep-fake-video-1784117> [<https://perma.cc/352G-GYT2>].

2. Steve Contorno & Donie O'Sullivan, *DeSantis Campaign Posts Fake Images of Trump Hugging Fauci in Social Media Video*, CNN (June 8, 2023, 4:00 PM), <https://www.cnn.com/2023/06/08/politics/desantis-campaign-video-fake-ai-image/index.html> [<https://perma.cc/MD4M-NL5U>]; Shannon Bond, *DeSantis Campaign Shares Apparent AI-Generated Fake Images of Trump and Fauci*, NPR (June 8, 2023, 3:59 PM), <https://www.npr.org/2023/06/08/1181097435/desantis-campaign-shares-apparent-ai-generated-fake-images-of-trump-and-fauci> [<https://perma.cc/9MHX-URMZ>].

is because neither legislation nor federal administrative regulations have caught up with a rapidly improving technology, the use of which is becoming increasingly common across society and increasingly sophisticated, including in the context of election campaigns.³

The emergence of deepfake technology—audio and visual content generated by AI realistically depicting people saying or doing things that they in fact never said or did—raises a host of social and legal concerns.⁴ As the examples above illustrate, in the election context, such technology can be used by candidates, outside groups such as super PACs and independent expenditures (IEs), foreign actors, or the average social media user to broadcast false content aimed at misleading the public and swaying public perception of the targeted individuals. The posts depicting Senator Warren and former president Trump are but a few examples of how individuals, campaigns, and outside groups have harnessed such technology in the leadup to the 2024 presidential election.⁵ And while the technology is relatively novel, it is rapidly becoming more accessible and cheaper to create.⁶

Yet for all of the grave risks that such technology poses and its growing use in spreading election misinformation, little federal action has been taken to combat such dangers. Neither Congress nor the relevant regulatory agencies, such as the Federal Election Commission (FEC), have enacted any solution aimed at reining it in and avoiding a severe distortion of reality in our electoral process.⁷ While some states have passed laws targeting deepfakes in the

3. See Daniel I. Weiner & Lawrence Norden, *Regulating AI Deepfakes and Synthetic Media in the Political Arena*, BRENNAN CTR. FOR JUST. (Dec. 5, 2023), <https://www.brennancenter.org/our-work/research-reports/regulating-ai-deepfakes-and-synthetic-media-political-arena> [<https://perma.cc/44LR-CJP4>].

4. See Marc Jonathan Blitz, *Lies, Line Drawing, and (Deep) Fake News*, 71 OKLA. L. REV. 59, 106 (2018).

5. See *infra* Part I.A.

6. TODD C. HELMUS, RAND CORP., *ARTIFICIAL INTELLIGENCE, DEEPPAKES, AND DISINFORMATION: A PRIMER* 4 (July 6, 2022), <https://www.rand.org/pubs/perspectives/PEA1043-1.html> [<https://perma.cc/M555-67P9>].

7. See Tatyana Monnay, *Deepfake Political Ads Are 'Wild West' for Campaign Lawyers*, BLOOMBERG (Sept. 5, 2023, 5:00 AM), <https://news.bloomberglaw.com/business-and-practice/deepfake-political-ads-are-wild-west-for-campaign-lawyers> [<https://perma.cc/JR5S-JX2G>]; Ali Swenson, *FEC Moves Toward Potentially Regulating AI Deepfakes in Campaign Ads*, PBS NEWSHOUR (Aug. 10, 2023, 6:37 PM), <https://www.pbs.org/newshour/politics/fec-moves-toward-potentially-regulating-ai-deep-fakes-in-campaign-ads> [<https://perma.cc/96JV-MZT7>].

election context, they are nonexistent in the majority of states, often lack teeth, and, in some cases, have been struck down by the courts as unconstitutional restrictions on speech.⁸

While the law has not yet adequately targeted deepfakes, existing law and Supreme Court jurisprudence afford remedies for many of the harms that such technology seeks to inflict. Defamation law, for example, imposes civil liability when a speaker negligently publishes or communicates a false statement purporting to be fact about an individual to a third person that injures that individual's reputation.⁹ Laws around impersonation vary by state but typically impose criminal and civil liability when one knowingly impersonates another person with an intent to defraud them.¹⁰

This Note explores how election deepfakes fit into the broader frameworks of defamation and impersonation law and posits that Congress and the judiciary may remedy the threat by borrowing from both areas of law. It builds on existing scholarship that has addressed the issue of deepfakes in both the electoral context as well as the general context. Furthermore, it expands upon scholars' prior identification of these areas of law as potential avenues for regulating election deepfakes.¹¹ It posits that solutions to election deepfakes offered by Professors Rebecca Green and Rick Hasen could go even further while still surviving constitutional scrutiny, diverging from their proposals by suggesting that disclosure and timing-related exemptions from liability are not constitutionally necessary. Specifically, it calls for the passage of a federal law marrying defamation and impersonation law by establishing a

8. See *Ex parte Stafford*, 667 S.W.3d 517, 521 (Tex. App. 2023) (striking down Texas law targeting deepfakes as an unconstitutional speech restriction); Act of Oct. 3, 2019, ch. 493, § 4, 2019 Cal. Stat. 4316, 4318 (subjecting campaign deepfakes to disclosure requirement); Act of May 10, 2023, ch. 360, § 2, 2023 Wash. Sess. Laws 1892, 1892 (subjecting campaign deepfakes to disclosure requirement); Act of May 26, 2023, ch. 58, § 2, 2023 Minn. Laws 1872, 1875 (banning campaign deepfakes within 90 days of an election).

9. RESTATEMENT (SECOND) OF TORTS § 558 (AM. L. INST. 1977).

10. See 35 C.J.S. *False Personation* § 1 (2024).

11. See, e.g., Rebecca Green, *Counterfeit Campaign Speech*, 70 HASTINGS L.J. 1445, 1475-76 (2019) (discussing whether state rules banning impersonation or tort claims such as defamation, false light, or the right of publicity can address counterfeit campaign speech); Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753, 1793-95 (2019) (discussing whether deepfake victims could sue creators under tort laws such as defamation, false light, intentional infliction of emotional distress, and right of publicity).

criminal offense when a speaker “creates and publishes, with actual malice, doctored audio or visual content depicting a candidate for office, in order to deceive voters, and thereby gain electoral advantage relating to the target of the communication.”

I. BACKGROUND

A. *A (Brief) Primer on Deepfake Technology and Its Use in Campaigns*

A deepfake is defined as “an image or recording that has been convincingly altered and manipulated to misrepresent someone as doing or saying something that was not actually done or said.”¹² The technological method by which deepfakes are created involves machine learning through the use of generative adversarial networks (GANs),¹³ which, at a basic level, pit two adversarial AIs against each other with a “forger” AI generating images and a “detective” AI determining whether the image is a forgery.¹⁴ Both AIs then learn to improve their functions, with the “detective” AI becoming better at identifying forgeries and the “forger” becoming better at creating the forgeries.¹⁵ The end result is a generator that can produce fake content with a high degree of convincingness.¹⁶

While the GAN technology emerged in 2014,¹⁷ the term “deepfake” originated in the pornography context in 2017.¹⁸ Named for a Reddit user who created a forum where users would share fake pornographic depictions of celebrities using face-swapping technology,¹⁹

12. *Deepfake*, MERRIAM-WEBSTER, <https://www.merriam-webster.com/dictionary/deepfake> [<https://perma.cc/73T2-KYRZ>].

13. See HELMUS, *supra* note 6, at 3.

14. Bernard Marr, *Artificial Intelligence Explained: What Are Generative Adversarial Networks (GANs)?*, FORBES (June 12, 2019, 12:23 AM), <https://www.forbes.com/sites/bernardmarr/2019/06/12/artificial-intelligence-explained-what-are-generative-adversarial-networks-gans/?sh=58a16fac7e00> [<https://perma.cc/MV62-AVFF>].

15. *Id.*

16. *See id.*

17. HELMUS, *supra* note 6, at 3.

18. Meredith Somers, *Deepfakes, Explained*, MIT SLOAN SCHOOL OF MGMT.: IDEAS MADE TO MATTER (July 21, 2020), <https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained> [<https://perma.cc/R3M2-R6TW>].

19. *Id.*

the technology has since rapidly evolved in both technological capability and scope of use.²⁰ The technology company OpenAI, for example, has made waves with its exceedingly convincing “Sora” text-to-video tool, which generates videos based on text prompts.²¹ A 2024 survey found that viewers were largely unable to distinguish the AI-generated videos from real footage.²² While there is enormous potential for deepfakes to be put to even more harmful use than what we have already seen, the technology has positive potential as well, with opportunities to modernize educational curricula and artistic displays, to name just a few examples.²³

The use of deepfakes in electoral campaigns is not limited to the examples mentioned above. Other examples of such use in the 2024 campaign cycle include an ad from Never Back Down, a pro-DeSantis super PAC, which depicted former president Trump criticizing Iowa governor Kim Reynolds and condemned Trump for fighting other Republicans.²⁴ While Trump did in fact make the depicted statement on Truth Social, the voice recording was AI-generated.²⁵ In response to President Biden’s announcement that he would seek re-election, the Republican National Committee released an ad depicting a dystopian United States with images of migrants crossing the border in droves and soldiers policing the streets.²⁶ When CNN showed the ad to potential voters, some stated they were unsure whether the depictions were real or fake.²⁷ In the lead-up to the 2024 New Hampshire primary, a campaign consultant for Congressman Dean Phillips, a primary challenger to President

20. See HELMUS, *supra* note 6, at 3-4.

21. Sora, OPENAI, <https://openai.com/sora> [<https://perma.cc/2LFR-R5CQ>].

22. Audrey Schomer, *Sora AI Videos Easily Confused with Real Footage in Survey Test*, VARIETY (Mar. 8, 2024, 6:00 AM), <https://variety.com/vip/sora-ai-video-confusion-human-test-survey-1235933647/> [<https://perma.cc/2F2F-2HV5>].

23. See Simon Chandler, *Why Deepfakes Are a Net Positive for Humanity*, FORBES (Mar. 9, 2020, 12:33 PM), <https://www.forbes.com/sites/simonchandler/2020/03/09/why-deepfakes-are-a-net-positive-for-humanity/> [<https://perma.cc/82FT-VGZU>].

24. See Miranda Nazzaro, *Pro-DeSantis Group Uses AI Version of Trump’s Voice in New Ad*, THE HILL (July 18, 2023, 10:15 AM), <https://thehill.com/homenews/campaign/4103157-pro-desantis-group-uses-ai-version-of-trumps-voice-in-new-ad/> [<https://perma.cc/JRW2-2PKB>].

25. *Id.*

26. Donie O’Sullivan & Yahya Abou-Ghazala, *The AI Political Campaign Is Here*, CNN (May 3, 2023, 6:16 AM), <https://www.cnn.com/2023/05/02/politics/ai-election-ads-2024/index.html> [<https://perma.cc/22GK-WF5X>].

27. *Id.*

Biden, hired technological consultants to create a robocall featuring a fake voice of President Biden urging voters to abstain from voting in the primary.²⁸ These are but a few additional examples, and experts agree that the role of AI in our electoral system is only going to grow.²⁹

B. The Harm of Campaign Deepfakes

Deepfakes pose threats that extend far beyond the election context. Professors Robert Chesney and Danielle K. Citron provide a comprehensive survey of the costs of such technology, outlining the many ways in which it threatens institutional trust, public safety, diplomacy, national security, and much more.³⁰ This Note seeks not to explore remedies for all of these potential harms but rather to zero in specifically on the use of deepfakes in elections.

In the election context, the harms are manifold. At a base level, much like any instance of defamation or impersonation, there is harm inflicted upon the target of the deepfake. That person's reputation and, by extension, level of support among the electorate, may be severely damaged.³¹ As Chesney and Citron point out:

The potential to sway the outcome of an election is real, particularly if the attacker is able to time the distribution such that there will be enough window for the fake to circulate but not enough window for the victim to debunk it effectively (assuming it can be debunked at all).³²

This particular form of harm is of course suffered not just by the target of the deepfake but also by the voters at large, who were duped into changing their opinions, and thus voting decisions, as a result of disinformation. As Green points out, deepfakes, which she refers to as a form of “counterfeit campaign speech”—“hyper-realistic candidate *source material* that is indistinguishable from

28. Tobi Raji, *Group Sues After Novel Biden Deepfake Urged Voters to Skip N.H. Primary*, WASH. POST, (Mar. 17, 2024, 12:08 PM), <https://www.washingtonpost.com/politics/2024/03/16/biden-deepfake-robocall-lawsuit-new-hampshire/> [<https://perma.cc/K5QJ-ZURX>].

29. See Chesney & Citron, *supra* note 11, at 1757.

30. *Id.* at 1754.

31. See *id.* at 1774, 1778.

32. *Id.* at 1778.

reality”—restrict voting rights by depriving voters of agency and the ability to make informed decisions about who to support based on real information.³³

Related to the harm suffered by voters is the harm inflicted upon the overall electoral system. If elections are determined based on voters’ false perceptions of the candidates, it goes without saying that the integrity and purpose of the system as a whole are severely undermined. The use of campaign deepfakes leads to an overall distortion of truth that Chesney and Citron refer to as “the liar’s dividend,” which describes how the more deepfakes are used and the more wary the public is of their existence, the less trust they will have in the information they are receiving generally.³⁴ Bad actors then benefit from this distrust because when they engage in bad behavior, the public will be skeptical that the behavior actually occurred, especially if the bad actor outright denies the accuracy of the depiction.³⁵ For example, Professor Hany Farid notes that former president Trump could, in our current deepfake environment, more plausibly claim that the infamous “Access Hollywood” tape, which leaked before the 2016 election and featured the then-candidate bragging about sexual assault, was faked.³⁶ Every election cycle, there seems to be a handful of candidate gaffes, often filmed clandestinely at private events, that generate front page headlines and drastically affect public perception of major candidates. Notable examples include Hillary Clinton labeling half of Trump supporters as belonging in a “basket of deplorables”³⁷ and Mitt Romney claiming that “there are 47 percent of the people who will vote for [President Obama] no matter what” because they are “dependent upon government” and “believe that they are victims.”³⁸ One can imagine how the liar’s dividend phenomenon would make members of the public more skeptical about the veracity of these comments

33. Green, *supra* note 11, at 1447, 1457-58.

34. Chesney & Citron, *supra* note 11, at 1785.

35. *See id.*

36. O’Sullivan & Abou-Ghazala, *supra* note 26.

37. Katie Reilly, *Read Hillary Clinton’s ‘Basket of Deplorables’ Remarks About Donald Trump Supporters*, TIME (Sept. 10, 2016, 12:27 PM), <https://time.com/4486502/hillary-clinton-basket-of-deplorables-transcript/> [<https://perma.cc/H6YN-922B>].

38. David Corn, *Mitt Romney’s Incredible 47-Percent Denial: “Actually, I Didn’t Say That,”* MOTHER JONES (July 29, 2013), <https://www.motherjones.com/politics/2013/07/mitt-romney-47-percent-denial/> [<https://perma.cc/9A7V-WHLW>].

given the pervasiveness of deepfakes, especially if a dishonest candidate was willing to cry “deepfake.” While some level of skepticism is a positive thing in our present era of rampant misinformation, society suffers when that skepticism is misplaced or overly-applied.

Events arising out of the 2023 presidential election in Argentina illustrate this phenomenon. The campaign of Sergio Massa, the center-left candidate who was defeated in October 2023 by far-right populist Javier Milei,³⁹ used AI technology in a number of contexts.⁴⁰ The campaign created its own generative AI system which it used to create posters depicting Massa in the style of a traditional Soviet propaganda poster; videos depicting him as a soldier, a Ghostbuster, and Indiana Jones; and videos negatively depicting Milei as characters in films such as *A Clockwork Orange*.⁴¹ While those depictions seem relatively innocuous and easily identifiable as fake or satire, the Massa campaign also put the technology to more nefarious uses. For example, it created a deepfake depicting Milei explaining “how a market for human organs would work.”⁴² These uses have the potential to mislead voters even when viewed in isolation. However, exacerbating the issue is the resulting skepticism of content that actually *is* real and accurate. For example, a video circulated of Massa looking exhausted after an event, leading his critics to accuse him of being on drugs.⁴³ His supporters responded by claiming the video was a deepfake, despite the fact that, unlike the ones Massa’s own campaign disseminated, it was not.⁴⁴ Here we see the liar’s dividend in action: by using misleading deepfake technology, Massa was able not only to depict his opponent in a false, negative light but also harness the overall distrust that

39. Daniel Politi & David Biller, *Fiery Right-Wing Populist Javier Milei Wins Argentina’s Presidency and Promises ‘Drastic’ Changes*, ASSOCIATED PRESS, (Nov. 20, 2023, 2:36 AM), <https://apnews.com/article/argentina-election-president-milei-massa-a4811c5229d35551f8dbf7056d87aae6> [<https://perma.cc/T9XN-JQQ2>].

40. Jack Nicas & Lucía Cholakian Herrera, *Is Argentina the First A.I. Election?*, N.Y. TIMES (Nov. 16, 2023), <https://www.nytimes.com/2023/11/15/world/americas/argentina-election-ai-milei-massa.html> [<https://perma.cc/27KC-KTY8>].

41. *Id.*

42. *Id.*

43. *Id.*

44. *Id.*

deepfakes cause in order to undermine criticism of himself that actually *was* based in real events.⁴⁵

It is this type of harm—the erosion of integrity in the electoral process—that the proposed federal law would seek to target. Framing the harm in this way makes the law more likely to withstand constitutional scrutiny than if the harm were framed as the distortion of truth or the dissemination of false speech. This is because the Supreme Court has made clear that lies are protected speech under the First Amendment.⁴⁶ While a law that merely polices lies could not withstand First Amendment scrutiny, deepfakes are distinct from traditional campaign lies because of the public’s inability to discern truth from reality. As Green notes, unlike with traditional campaign lies, when it comes to election deepfakes, “more speech cannot adequately cure the damage done.”⁴⁷

Furthermore, as Hasen points out, the Court “has long recognized the value of an ‘active, alert’ citizenry” and “should recognize that the government has a compelling interest in assuring that voters have access to truthful political information and to the tools to discover its truth or falsity.”⁴⁸ In other words, while voters are not entitled to be free from lies, they are entitled to be equipped with the tools necessary to detect lies. Again, traditional campaign lies can be debunked with more speech; the same cannot be said for deepfakes.⁴⁹

Hasen also discusses how the Court has signaled an openness to treating the protection of the integrity of the voting process as a compelling interest.⁵⁰ While the specific case that he points to involved false speech about the *mechanics* of voting, rather than

45. *See id.*

46. *See United States v. Alvarez*, 567 U.S. 709, 729 (2012) (plurality opinion).

47. Green, *supra* note 11, at 1484.

48. Richard L. Hasen, *Deep Fakes, Bots, and Siloed Justices: American Election Law in a “Post-Truth” World*, 64 ST. LOUIS U. L.J. 535, 545 (2020) (quoting *First Nat’l Bank of Boston v. Bellotti*, 435 U.S. 765, 788 (1978)).

49. *See Green*, *supra* note 11, at 1452 (“The problem ... arises only when there is no means, at least that we can reasonably expect voters to employ, to discern truth from falsity. In the case of counterfeit campaign speech, the counterfeiter has fabricated source material such that additional work and/or more speech cannot cure the harm.”).

50. Hasen, *supra* note 48, at 548-49.

about candidates,⁵¹ the reasoning for why such speech may be subject to restrictions applies also to deepfakes aimed at harming candidates: “[s]uch speech does not raise the risk of the state having to make judgment calls about truth or falsity. For example, the location of a polling place is objectively verifiable.”⁵² The same can be said about deepfakes depicting candidates—whether such content is real or fake, while potentially impossible for the average viewer to discern, is objectively verifiable by the government. There is therefore little risk of the government needing to make subjective (and potentially incorrect) judgment calls about the content’s veracity.

In addition, while the personal reputations of electoral candidates are worthy of protection, they are not the chief concern in the grand scheme of the potential harms arising out of the pervasive use of campaign deepfakes. More importantly, the judiciary would likely be reluctant to treat the protection of candidates’ reputations as a sufficient state interest worthy of promoting through regulation.⁵³ The Supreme Court has rightfully made clear that government and public officials should be expected to endure a higher degree of “unpleasantly sharp attacks” than the average citizen⁵⁴ and later extended this approach to the broader category of public figures.⁵⁵ This line of jurisprudence lessens the constitutional viability of the law if the state interest were to be framed as protecting against this form of harm. Furthermore, while alleviating the harm to individual voters should be prioritized, we can think of this harm as a component of the overall harm of distorting the integrity of the electoral

51. Hasen states that, “the government likely could prohibit ... false *election* speech, which is false speech about the mechanics of voting.” *Id.* at 548 (discussing *Minnesota Voters Alliance v. Mansky*, 585 U.S. 1 (2018)). While deepfakes depicting candidates do not fall squarely into this category of speech, the reasoning why false election speech may be regulated applies to such deepfakes: their veracity is objectively verifiable.

52. *Id.*

53. *See* Weiner & Norden, *supra* note 3 (“In our view, promoting a more informed electorate and safeguarding the integrity of the electoral process are the most compelling objectives for restricting manipulated media. Other valid aims include shielding candidates and election workers and curbing harmful disinformation more broadly, but pursuing those goals could risk unduly infringing on protected types of expression.”).

54. *New York Times Co. v. Sullivan*, 376 U.S. 254, 270 (1964).

55. *See, e.g., Hustler Mag. v. Falwell, Inc.*, 485 U.S. 46, 56 (1988) (holding that public figures, as well as public officials, may not recover for the tort of intentional infliction of emotional distress without a showing of actual malice).

process. The more voters whose interest in being accurately informed is attacked, the greater the overall distortion of electoral integrity. Thus, targeting the broader harm addresses the more specific harm in the process.

C. Statutory Overview: Existing Deepfake Regulations

There is no federal law in place directly targeting deepfakes.⁵⁶ Statutes relating to deepfakes only exist in a minority of states⁵⁷ and have specifically focused on the campaign context and the pornography context.⁵⁸ As of 2023, only nine states had passed laws regulating deepfakes, and of those nine, only five pertained to elections, with the remaining four targeting pornographic deepfakes.⁵⁹ In the 2024 legislative session, states began to legislate more aggressively, with twenty-seven bills having been enacted across seventeen states, with seven targeting elections specifically.⁶⁰ While a full survey of these state laws is beyond the scope of this analysis, the following examples are illustrative of how states have approached the issue.

California amended its election code in 2019 to make it an offense for a person, committee, or other entity to,

within 60 days of an election at which a candidate for elective office will appear on the ballot, distribute, with actual malice, materially deceptive audio or visual media ... of the candidate

56. Nicholas O'Donnell, *Have We No Decency? Section 230 and the Liability of Social Media Companies for Deepfake Videos*, 2021 ILL. L. REV. 701, 711 (2021).

57. Adam Edelman, *States Are Lagging in Tackling Political Deepfakes, Leaving Potential Threats Unchecked Heading into 2024*, NBC NEWS (Dec. 16, 2023, 7:00 AM), <https://www.nbcnews.com/politics/artificial-intelligence-deepfakes-2024-election-states-rcna129525> [<https://perma.cc/XS3Q-PGUF>].

58. Isaiah Poritz, *States Are Rushing to Regulate Deepfakes as AI Goes Mainstream*, BLOOMBERG (June 20, 2023, 5:01 AM), <https://www.bloomberg.com/news/articles/2023-06-20/deepfake-porn-political-ads-push-states-to-curb-rampant-ai-use> [<https://perma.cc/F4RX-PSX3>].

59. *See id.*; *Tracker: State Legislation on Deepfakes in Elections*, PUB. CITIZEN (May 10, 2024), <https://www.citizen.org/article/tracker-legislation-on-deepfakes-in-elections/> [<https://perma.cc/9RXR-7G7M>].

60. *Deceptive Audio or Visual Media ('Deepfakes') 2024 Legislation*, NAT'L CONF. OF STATE LEGISLATURES (2024), <https://www.ncsl.org/technology-and-communication/deceptive-audio-or-visual-media-deepfakes-2024-legislation> [<https://perma.cc/6Z9F-UAHP>].

with the intent to injure the candidate’s reputation or to deceive a voter into voting for or against the candidate,

but creates an exception if a simple disclosure is added to the content.⁶¹ Texas amended its election code the same year to establish an offense when a person, “with intent to injure a candidate or influence the result of an election ... creates a deep fake video” and “causes the deep fake video to be published or distributed within 30 days of an election.”⁶² A state appellate court eventually struck down this law due to a different subsection being a content-based restriction that fails to survive strict scrutiny.⁶³ In 2023, Washington enacted a law establishing a civil cause of action when a candidate is the victim of a deepfake but established an affirmative defense when the publisher includes a simple disclosure, creating the same shortfall as the California law.⁶⁴ Minnesota passed a law that targets both pornographic and election-related deepfakes, banning election deepfakes entirely with no carveout for ads containing disclosures, though the law only affects communications made within ninety days of an election.⁶⁵

While the efforts of state legislatures are a step in the right direction and serve as important laboratories for more robust federal regulation, the remedy for such an existential threat cannot lie with the states alone. Such an approach has led to a disjointed and sparse patchwork of state laws that lack force and, in the case of the Texas law, constitutionality.⁶⁶ With thirty-two states and the District of Columbia yet to pass any sort of statutory remedy, it is clear that a state-based approach alone cannot keep up with the rapid expansion of the threat that deepfakes in elections pose.

This Note’s criticism of state laws that contain exemptions when a deepfake contains a disclosure or is released outside of an electioneering window is controversial. Green and Hasen suggest that such limitations are necessary to survive constitutional muster,⁶⁷

61. Act of Oct. 3, 2019, ch. 493, § 4, 2019 Cal. Stat. 4316, 4318.

62. Act of June 14, 2019 ch. 1339, § 1, 2019 Tex. Gen. Laws 3936, 3936.

63. *See Ex parte* Stafford, 667 S.W.3d 517, 521, 524 (Tex. App. 2023).

64. Act of May 10, 2023, ch. 360, § 2, 2023 Wash. Sess. Laws 1892.

65. Act of May 26, 2023, ch. 58, § 2, 2023 Minn. Laws 1872, 1875.

66. *Ex parte* Stafford, 667 S.W.3d at 532 (striking down the Texas law).

67. Green proposes a “criminal sanction for the knowing manufacture of fake images, audio or other material of an identifiable candidate for public office, published within [a specified

and the legislatures that have enacted such laws clearly agree. It is true that those exemptions strengthen the “narrowly tailored” prong under a strict scrutiny analysis. However, this Note’s proposed solution—especially in light of the “actual malice” requirement—would only sweep in content that is *per se* punishable as defamation. As Hasen notes, while “courts should hold unconstitutional most broad state laws barring false speech in campaigns,” they should “reject challenges to narrower laws that ... allow a jury to punish defamatory speech about candidates made with actual malice.”⁶⁸ This Note’s proposed solution would only sweep in content that falls within that category.

D. Defamation Law

Defamation law allows for individuals to recover damages when defamatory statements about them injure their reputations.⁶⁹ The common law elements of the tort of defamation require that a speaker publishes or communicates a false statement of fact about or concerning an individual that causes injury to the individual’s reputation.⁷⁰ Notably, the bar for bringing a defamation claim is significantly higher for public officials and public figures than it is for private individuals.⁷¹ The governing standard under *New York Times Co. v. Sullivan* requires a defamation plaintiff to prove by clear and convincing evidence that the defendant made the statement in question with “actual malice,” meaning knowledge that the statement was false, or “reckless disregard” for whether or not it was.⁷² Rationalizing its imposition of this exceptionally high standard for public figures, the Court noted the “profound national commitment to the principle that debate on public issues should be

number of] days prior to an election, with intent to deceive voters and distort the electoral process.” Green, *supra* note 11, at 1456. Hasen proposes a “truth-in-labeling regime,” under which social media platforms would be required to label altered video and audio as such. Hasen, *supra* note 48, at 550.

68. Richard L. Hasen, *A Constitutional Right to Lie in Campaigns and Elections*, 74 MONT. L. REV. 53, 57 (2013).

69. *E.g.*, *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 349-50 (1974).

70. RESTATEMENT (SECOND) OF TORTS § 558 (AM. L. INST. 1977).

71. *New York Times Co. v. Sullivan*, 376 U.S. 254, 279 (1964).

72. *Id.* at 280.

uninhibited, robust, and wide-open.”⁷³ The Court further rested its holding upon its concern about a chilling effect whereby “would-be critics of official conduct may be deterred from voicing their criticism, even though it is believed to be true and even though it is in fact true, because of doubt whether it can be proved in court or fear of the expense of having to do so.”⁷⁴ It is safe to assume that any political candidate who is the target of an election deepfake would fall into either the public official or public figure category, thus subjecting them to the “actual malice” standard.⁷⁵

E. Impersonation Law

Laws against impersonation, or “false personation,” are largely state-based, with some federal protections against the impersonation of officers or federal employees.⁷⁶ The broadest category of impersonation laws prohibits impersonation generally. Alabama, for example, has established a criminal offense when one “[a]ssumes a false identity and does an act in his assumed character with intent to gain an economic benefit for himself or another or to injure or defraud another.”⁷⁷ Some states have laws against online or electronic impersonation specifically, imposing both criminal and civil penalties.⁷⁸ Finally, many states have statutes resembling the

73. *Id.* at 270.

74. *Id.* at 279.

75. *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 352 (1974) (defining a “public figure” as someone who enjoys “general fame or notoriety in the community” and “pervasive involvement in the affairs of society”).

76. 18 U.S.C. § 912 (establishing a criminal offense for impersonation of an officer or employee of the United States); 10 U.S.C. § 906 Art. 106 (establishing a criminal offense when one “wrongfully and willfully” impersonates a military officer, “an agent of superior authority” within the armed forces, or a government official).

77. ALA. CODE § 13A-9-18 (West 2024). For similar state laws, see COLO. REV. STAT. ANN. § 18-5-113 (West 2024) (making it a crime when one “[a]ssumes a false or fictitious identity or capacity, legal or other, and in such identity or capacity he or she: ... [p]erforms any other act with intent to unlawfully gain a benefit for himself, herself, or another or to injure or defraud another”); CONN. GEN. STAT. ANN. § 53a-130 (West 2023) (making it a crime when one “[i]mpersonates another and does an act in such assumed character with intent to obtain a benefit or to injure or defraud another”); N.Y. PENAL LAW § 190.25 (McKinney 2024) (making it a crime when one “[i]mpersonates another and does an act in such assumed character with intent to obtain a benefit or to injure or defraud another”).

78. *E.g.*, WASH. REV. CODE ANN. § 4.24.790 (West 2024) (“A person may be liable in a civil action based on a claim of invasion of privacy when ... [t]he person impersonates another actual person on a social networking website or online bulletin board” and “[t]he person

aforementioned federal statutes, criminalizing the impersonation of certain government officials.⁷⁹ While statutes vary widely in the penalties they impose, they share the general requirements that the offender falsely assumes the identity of another person and in doing so commits some sort of harm to another, be it harassment, intimidation, humiliation, or fraud.⁸⁰ A major component of the offense is the deception of others in order to gain some profit or advantage, “appertaining to the party so personated.”⁸¹

II. ANALYZING DEEPPAKES WITHIN THE DEFAMATION AND IMPERSONATION FRAMEWORKS

A. *Do Deepfakes Defame Their Targets?*

When a deepfake is distributed, the action is similar in nature to an action giving rise to a defamation claim. When a speaker defames, they communicate a false statement purporting to be fact about an individual to a third person.⁸² Similarly, when a speaker publishes a deepfake, they communicate a false depiction of an act or communication purporting to have actually occurred about an individual to a third person (ordinarily multiple people).⁸³ Thus, publication of a deepfake should give rise to a prima facie defamation claim. What distinguishes the two in the campaign context, however, is the harm that the law of defamation seeks to remedy: an

intended to deceive or mislead for the purpose of harassing, threatening, intimidating, humiliating, or defrauding another.”); CAL. PENAL CODE § 528.5 (West 2024) (“[A]ny person who knowingly and without consent credibly impersonates another actual person through or on an Internet Web site or by other electronic means for purposes of harming, intimidating, threatening, or defrauding another person is guilty of a public offense.”).

79. *E.g.*, ALA. CODE § 13A-10-11 (2024) (“A person commits the crime of impersonating a peace officer if he or she falsely pretends to be a peace officer and does any act in that capacity.”); TEX. PENAL CODE ANN. § 37.11 (West 2023) (establishing a felony offense when one “impersonates a public servant with intent to induce another to submit to the person’s pretended official authority or to rely on the person’s pretended official acts” or “knowingly purports to exercise, without legal authority, any function of a public servant”).

80. *See, e.g.*, WASH. REV. CODE ANN. § 4.24.790; CAL. PENAL CODE § 528.5.

81. 35 C.J.S. *False Personation* § 1 (2024).

82. *See, e.g.*, *New York Times Co. v. Sullivan*, 376 U.S. 254, 258 (1964); *Gertz v. Robert Welch, Inc.*, 418 U.S. 323, 349-50 (1974).

83. *See* Blitz, *supra* note 4.

individual's reputation.⁸⁴ While a deepfake certainly may harm the targeted individual's reputation much like a defamatory statement does,⁸⁵ remedying reputational harm would, again, not be the purpose of a law regulating or banning deepfakes in campaign settings. Rather, the harm that causes the most concern, and which the courts would be most amenable to as a state interest, is the erosion of integrity in the electoral process.⁸⁶

A deepfake victim outside of the campaign context would likely have a viable defamation claim as well, and First Amendment scholars have identified this area of law as a potential avenue for liability.⁸⁷ Certainly a victim of a pornographic deepfake, which was the initial context in which deepfake technology emerged,⁸⁸ could likely meet all of the elements of defamation in a civil lawsuit against the creator. And even in a campaign context with a public official or public figure plaintiff, the elements would still be met; it is not hard to imagine how a court could find that the creator of a deepfake acted with actual malice and that the victim suffered reputational harm. The issue is thus not whether a deepfake aimed at harming a candidate and deceiving voters constitutes defamation—it does. The issues, rather, are twofold. Firstly, there is a misalignment between the *purpose* of defamation law (preventing reputational damage) and that of regulating campaign deepfakes (preventing a mass erosion of integrity of the electoral process).

84. *E.g.*, *Gertz*, 418 U.S. at 323-24 (highlighting the state's interest in compensating reputational injury).

85. *See* Chesney & Citron, *supra* note 11, at 1774 (highlighting the potential for deepfake technology's "utility for reputational sabotage").

86. *See* Weiner & Norden, *supra* note 3 ("In general, the necessity of promoting an informed electorate and the need to safeguard the overall integrity of the electoral process are among the most compelling rationales for regulating manipulated media in the political space.").

87. *See, e.g.*, Chesney & Citron, *supra* note 11, at 1802 ("In certain jurisdictions, creators of deep fakes could also face charges for criminal defamation if they posted videos knowing that they were fake or if they were reckless as to their truth or falsity."); Blitz, *supra* note 4, at 70 (noting that despite the Supreme Court's holding in *United States v. Alvarez* that lies are not unprotected speech, "false speech that causes certain legally cognizable harms can be punished or subjected to civil liability without raising significant First Amendment concerns," and "[g]overnment may punish or subject to liability the harm that defamation causes to reputation").

88. Donie O'Sullivan, *Nonconsensual Deepfake Porn Puts AI in Spotlight*, CNN (Feb. 16, 2023, 6:27 PM), <https://www.cnn.com/2023/02/16/tech/nonconsensual-deepfake-porn/index.html> [<https://perma.cc/AB2N-92PM>].

Secondly, there is the ineffectiveness of a remedy that would require a high-profile political candidate, preoccupied with running for public office and unsympathetic as a plaintiff, to bring a private cause of action against a deepfake creator.⁸⁹ That said, the fact that the type of deepfake covered under this Note's proposed solution inherently constitutes defamation creates a strong case for the law's constitutional viability.

B. Has a Deepfake Target Been Impersonated?

The act of creating and disseminating a deepfake comes close to meeting the common law elements of impersonation, or as the offense is sometimes referred to, "false personation."⁹⁰ Scholars have taken note of the potential this area of law has for reining in the use of deepfakes.⁹¹ However, one struggles to reconcile the fundamental nature of a deepfake as depicting the target or victim doing or saying something that they did not do or say, with impersonation law's requirement that the perpetrator engages in "a deliberate effort to pass [his or herself] off as another."⁹² To return to the original example, when the @DeSantisWarRoom Twitter account published a deepfake image of Donald Trump hugging Dr. Anthony Fauci, the creator was not *assuming the identity* of Donald Trump—rather, they were attempting to depict Donald Trump doing something he did not in fact do.⁹³ Herein lies the greatest conceptual disparity between the offense of impersonation and the act of creating and publishing a deepfake: when one impersonates, one passes oneself off as another person; when one creates a deepfake, on the other hand, one simply depicts a person saying or doing something they did not in fact say or do.

89. See Green, *supra* note 11, at 1476 (highlighting how "tort claims are expensive to bring and can take months, if not years, to resolve, rendering them an ineffective means of addressing election disruption").

90. C.J.S. *False Personation* § 1 (2024).

91. See, e.g., Chesney & Citron, *supra* note 11, at 1802 ("Impersonation crimes may be applicable as well. Several states make it a crime, for example, to knowingly and credibly impersonate another person online with intent to 'harm[], intimidat[e], threaten[], or defraud[]' that person."); Green, *supra* note 11, at 1474 (exploring the potential for state laws prohibiting impersonation to address the issue of counterfeit campaign speech).

92. C.J.S. *False Personation* § 1 (2024).

93. Contorno & O'Sullivan, *supra* note 2.

In addition to this conceptual difference, impersonation law runs into the same problem as defamation law when it comes to its potential use as a vehicle for liability for campaign deepfakes. That is, it would require a private action to be brought, which in addition to the aforementioned considerations of time and expense, would be an unlikely remedy for a high-profile political candidate, presidential or otherwise, to bring.⁹⁴ If a deepfake is released on the eve of an election, it goes without saying that any civil remedy, be it through a defamation or impersonation suit, would only provide recourse well after the damage has already been done.⁹⁵ Furthermore, as noted, the individual harm suffered by the candidate is not the harm with which this proposed remedy would be chiefly concerned.

III. MARRYING DEFAMATION LAW AND IMPERSONATION LAW TO CREATE A REMEDY FOR ELECTION DEEPPAKES

A. *A Mechanical Crafting of the Proposed Remedy*

Despite these disparities between deepfakes and defamation and impersonation, respectively, when broken down into their individual elements, certain elements can be taken from both and merged into a remedy for campaign deepfakes that both addresses (at least in part) the harm and passes First Amendment constitutional muster. We can break the common law tort of defamation into its elements by stating that defamation occurs when a speaker (1) publishes or communicates a false statement of fact (2) about or concerning an individual which (3) causes injury to that individual's reputation.⁹⁶ Turning to impersonation, we can break that offense down into the elements of (1) acting in a character unlawfully assumed (2) in order to deceive others and (3) gain some profit or advantage, or enjoy some right or privilege (4) appertaining to the party so personated.⁹⁷

94. See Green, *supra* note 11, at 1476.

95. See Chesney & Citron, *supra* note 11, at 1774 (highlighting how debunking a deepfake may often “come too late to remedy the initial harm,” for example if a rival attempts to “torpedo the draft position of a top pro sports prospect by releasing a compromising deep-fake video just as the draft begins”).

96. RESTATEMENT (SECOND) OF TORTS § 558 (AM. L. INST. 1977); *Milkovich v. Lorain Journal Co.*, 497 U.S. 1, 13 (1990).

97. 35 C.J.S. *False Personation* § 1 (2024).

Element (1) of defamation (“publishes or communicates a false statement of fact”) is the underlying action of the offense and is perfectly aligned with the underlying action behind the use of a campaign deepfake. The fact that one is a false *statement* of fact and the other is a false *depiction* is immaterial—in both cases, the speaker is publishing false claims, just in different forms. Element (2) of defamation (“about or concerning an individual”) also applies directly to the campaign deepfake context. Element (3) of defamation (“causes injury to that individual’s reputation”) applies to campaign deepfakes but addresses a harm that the remedy does not seek to address and therefore is not included in the proposed remedy. Elements (1) and (2) of defamation are therefore included in the proposed remedy.

Element (1) of impersonation (“acting in a character unlawfully assumed”) is the underlying action of the offense but, as previously discussed, is distinct from the underlying action of a campaign deepfake (or any deepfake) and thus is not included in the remedy. Element (2) of impersonation (“in order to deceive others”) applies directly to campaign deepfakes, which are meant to deceive the electorate. Element (3) of impersonation (“gain some profit or advantage, or enjoy some right of privilege”) applies directly to campaign deepfakes, with the “advantage” or “privilege” being a comparative electoral advantage in relation to the deepfake target. Element (4) of impersonation (“appertaining to the party so personated”) applies directly as well, for the same reasons that element (3) does. Thus, elements (2), (3), and (4) of impersonation are included in the proposed remedy.

The resulting hybrid of these two offenses is a law that makes it a criminal offense when one “creates and publishes, with actual malice, doctored audio or visual content depicting a candidate for office, in order to deceive voters, and thereby gain electoral advantage relating to the target of the communication.” This proposed remedy involves slight changes to the language of the defamation and impersonation elements. Firstly, “with actual malice” is added in recognition of the Supreme Court’s heightened standard for what constitutes defamation when the target is a public official or public figure.⁹⁸ It ensures that any piece of content that falls within the

98. *New York Times Co. v. Sullivan*, 376 U.S. 254, 279-80 (1964).

purview of the law would be punishable under defamation law. Secondly, the “false statement of fact” language of element (1) of defamation is changed to “doctored audio or visual content.” The change is only to the form, not the substance, of the underlying action; doctored audio or visual content has the same effect as a false statement of fact, and, in fact, inflicts even more damage due to its attention-grabbing nature and greater immunity from being debunked. Thirdly, the addition of the language “depicting a candidate for office” and the change from “in order to deceive others” to “in order to deceive voters” narrows the scope of the law so that it is more likely to be sufficiently tailored to the government interest in protecting the integrity of the electoral process.

B. Constitutional Analysis of the Proposed Remedy

It is well-established that lies are protected speech under the First Amendment.⁹⁹ Therefore, as a content-based speech regulation, the proposed remedy would need to survive strict scrutiny.¹⁰⁰ The constitutional question then becomes whether the remedy is narrowly tailored to serve a compelling government interest.¹⁰¹

The compelling government interest in this case would be protecting the integrity of the electoral process. Again, as Hasen notes, “[t]he compelling interest portion of the argument for regulation of false media is surprisingly easy to make. Courts should recognize that the government has a compelling interest in assuring that voters have access to truthful political information and to the tools to discover its truth or falsity.”¹⁰² Furthermore, as Green highlights, “[t]he Supreme Court has recognized the government’s compelling interest in upholding the integrity of the elections in numerous cases.”¹⁰³

99. *United States v. Alvarez*, 567 U.S. 709, 715, 729 (2012) (striking down the Stolen Valor Act, which criminalized false claims about receiving military decorations, under strict scrutiny).

100. *See, e.g., id.*; *Reed v. Town of Gilbert*, 576 U.S. 155, 163 (2015) (“Content-based laws—those that target speech based on its communicative content—are presumptively unconstitutional and may be justified only if the government proves that they are narrowly tailored to serve compelling state interests.”).

101. *Reed*, 576 U.S. at 163.

102. Hasen, *supra* note 48, at 545.

103. Green, *supra* note 11, at 1460.

The next question under a strict scrutiny analysis is one of tailoring: is the proposed remedy narrowly tailored to serve the compelling state interest of protecting the integrity of the electoral process? The proposed remedy explicitly targets deepfakes, specifically in the campaign context, and only those campaign deepfakes that are meant to deceive the electorate and gain some advantage for the speaker or publisher over the target. It would have no direct or indirect effect on the use of deepfakes outside of the campaign context. Even within the campaign context, it would allow for the use of deepfake technology as long as such use does not falsely portray an opposing candidate saying or doing something they did not say or do. For example, a campaign could harness generative AI to create stock footage of farmland for an advertisement targeting rural populations, or a hospital setting for an advertisement related to health care policy. It also does not target lying in general in the election context, which, despite its harms, the Court has signaled a strong reluctance toward limiting.¹⁰⁴ Perhaps most importantly, the law only covers content that is *per se* defamatory. Thus, the law bars no more speech than is necessary and could defeat an allegation of being over-inclusive.

IV. COUNTERARGUMENTS

A. Elevated Status for Political Speech

One significant hurdle that scholars have rightfully raised is the Supreme Court's deep reluctance to impose limitations on political speech, which it considers to be among the most sacred forms of speech worthy of the utmost protection.¹⁰⁵ Restricting campaign

104. See, for example, *Susan B. Anthony List v. Driehaus*, 573 U.S. 149, 167-68 (2014), where the Court unanimously reversed a Sixth Circuit decision holding that the Susan B. Anthony List lacked standing due to lack of injury when it challenged an Ohio law banning election lies. While the case was decided purely on standing grounds, the Court signaled its sympathy toward the petitioner for the effect that the law banning election lies had on its speech, stating that “denying prompt judicial review would impose a substantial hardship on petitioners, forcing them to choose between refraining from core political speech on the one hand or engaging in that speech and risking costly Commission proceedings and criminal prosecution on the other.” *Id.*

105. See, e.g., *Mills v. Alabama*, 384 U.S. 214, 218 (1966) (“Whatever differences may exist about interpretations of the First Amendment, there is practically universal agreement that a major purpose of that Amendment was to protect the free discussion of governmental

deepfakes, however, would not restrict the type of political speech with which the Court is concerned. While the Court's reverence for political speech is a laudable stance that serves the goal of advancing "a profound national commitment to the principle that debate on public issues should be uninhibited, robust, and wide-open,"¹⁰⁶ it is difficult to imagine how this principle is served by a campaign deepfake intended to deceive the electorate, sew an environment of misinformation, and blur voters' perceptions of what is real and what is not. Political speech enjoys the utmost protection because of the societal value in criticizing public officials.¹⁰⁷ That value is not advanced by deepfakes whose sole purpose is to deceive and thereby gain an advantage in relation to the targeted individual.¹⁰⁸

Furthermore, at the center of the Court's past decisions striking down laws that it feared would inhibit political speech is the fear of chilling political speech. In *Virginia v. Black*, for example, the Court's concern over Virginia's anti-cross-burning statute was not the banning of cross-burning that was in fact meant to intimidate people based on race, but rather the act of banning *all* cross-burning, which it feared would chill cross-burning not meant to intimidate but instead meant to convey a political message.¹⁰⁹ In *New York Times Co. v. Sullivan*, one of the Court's chief concerns with the Alabama libel law at issue was that it raised "the possibility that a good-faith critic of government will be penalized for his criticism."¹¹⁰ And in *Brown v. Hartlage*, in which the Court struck

affairs."); *Brandenburg v. Ohio*, 395 U.S. 444, 448 (1969) (striking down Ohio's Criminal Syndicalism Act, which punished people who "advocate or teach the duty, necessity, or propriety of violence 'as a means of accomplishing industrial or political reform'"); *NAACP v. Claiborne Hardware Co.*, 458 U.S. 886, 927-28 (1982) (holding that civil rights activist Charles Evers's speech during a boycott of white businesses in Mississippi in which he stated that boycott violators' "necks would be broken" did not rise to the level of incitement, and instead constituted "an impassioned plea for black citizens to unify, to support and respect each other, and to realize the political and economic power available to them"); *Virginia v. Black*, 538 U.S. 343, 347-48, 365 (2003) (striking down a Virginia statute banning cross burning and treating such an act as prima facie evidence of an intent to intimidate because the statute risked reaching "somebody engaging only in lawful political speech at the core of what the First Amendment is designed to protect").

106. *New York Times Co. v. Sullivan*, 376 U.S. 254, 270 (1964).

107. *See id.* at 273 (referring to criticism of public officials as essential to a self-governing people and the "central meaning of the First Amendment").

108. *See, e.g., Chesney & Citron, supra* note 11, at 1791.

109. 538 U.S. at 365.

110. 376 U.S. at 292.

down a Kentucky law prohibiting candidates from offering any material benefit to voters in consideration for a vote, the problem with the law as applied to the defendant candidate was that he made the statement in question—a promise to lower his salary if elected—in good faith.¹¹¹ As the Court stated, “[t]he chilling effect of such absolute accountability for factual misstatements in the course of political debate is incompatible with the atmosphere of free discussion contemplated by the First Amendment in the context of political campaigns.”¹¹²

When it comes to campaign deepfakes, there is little concern over a chilling effect because there is no such thing as a campaign deepfake created in good faith if it falsely depicts, with actual malice, another candidate.¹¹³ Unlike a newspaper ad criticizing a government official or a statement made by a political candidate at a press conference, there is never a scenario involving a campaign deepfake where someone inadvertently conveys a falsity—the falsity will be fundamentally obvious to the creator. A deepfake created in good faith, such as stock footage not depicting a political opponent or a positive deepfake depicting a candidate as Superman, would clearly fall outside the law’s coverage. Both pieces of content are obviously not meant to deceive, and there is no target in relation to whom the creator is intending to gain an advantage. A reasonable creator of such content would have no concern that his or her creation would fall within the scope of the law. As a result, the Court’s concern over chilling effects, which underlies its placement of political speech at the top of the protected speech hierarchy,¹¹⁴ should not be a concern in the context of this proposed law.¹¹⁵ A ban on campaign deepfakes targeted at a candidate and intended to deceive the electorate will have little deterrent effect on other forms of protected speech, as the line between a deepfake and a good faith false statement is clear.

111. *Brown v. Hartlage*, 456 U.S. 45, 54, 61 (1982).

112. *Id.* at 61.

113. *See id.*

114. *See, e.g., Black*, 538 U.S. at 365; *Sullivan*, 376 U.S. at 292; *Hartlage*, 456 U.S. at 61.

115. Weiner & Norden, *supra* note 3 (“[D]eepfakes afford more clarity versus other types of misleading communications in that the former are by definition false, thereby reducing the risk that a regulator’s subjective perceptions of what is or is not truthful will lead them to chill sincere speech.”).

B. Elevated Status for Public Figures

Central to the Supreme Court’s First Amendment jurisprudence is the principle that public figures should be held to a higher standard when it comes to the level of speech that they can be expected to endure.¹¹⁶ The ability to criticize public officials and the government is the “central meaning of the First Amendment” and public officials should expect to bear “unpleasantly sharp attacks.”¹¹⁷ Thus, a fundamental constitutional roadblock to enacting the proposed remedy is the fact that it seeks to regulate content that targets those who are expected to endure more attack speech than the average citizen.¹¹⁸

Again, however, the state interest justifying the proposed regulation is not to shield public figures from criticism but rather to protect the electoral process from an erosion of integrity. Furthermore, the “actual malice” language within the proposed law ensures that any covered content would meet the elevated bar that public officials and public figures must prove to establish defamation.¹¹⁹ Thus, the proposed remedy does not establish culpability for any actions that would not meet the *Sullivan* standard and therefore does not run afoul of the Supreme Court’s protection for speech criticizing public officials and public figures.

C. Failure to Regulate Positive, yet Deceptive, Candidate Deepfakes

Admittedly, there are deceptive uses of deepfakes in the election context that depict candidates *positively* and thus would fall outside the scope of the law. For example, supporters of former president Trump have created and circulated deepfake images of him posing with Black supporters in an effort to enhance public perception of his relationships with the Black community.¹²⁰ It is true that this

116. *E.g.*, *Sullivan*, 376 U.S. at 270; *Hustler Mag., Inc. v. Falwell*, 485 U.S. 46, 56 (1988).

117. *Sullivan*, 376 U.S. at 270, 273.

118. *See id.*

119. *See id.* at 279-80.

120. Nitish Pahwa, *So, Fake Images of Trump With Black Voters Are a Thing Now*, SLATE (Apr. 10, 2024, 10:00 AM), <https://slate.com/news-and-politics/2024/04/donald-trump-deep-fake-political-images-artificial-intelligence-ai.html> [<https://perma.cc/WL2E-FUCB>].

type of content, despite its potential to spread misinformation and erode the integrity of the electoral process, would fall outside the scope of the proposed remedy. That said, what enables the proposed remedy to pass First Amendment scrutiny without being limited by time or disclosure exemptions is its limitation to content that is *per se* defamatory. A law regulating deepfakes that are deceptive, yet positive portrayals of candidates, would likely have to be subject to those limitations. Congress should take measures to curb that sort of content. However, in order to effectively stamp out negative, deceptive deepfakes without disclosure exemptions or time limitations, Congress would have to limit its remedy to inherently defamatory content and regulate other content through separate and more limited laws.

It is also true that different people can perceive the same content differently and that while this particular deepfake was meant to make Trump look better, some voters who hold racist beliefs may perceive it as making him look *worse*, raising the question of whether it *would* be included within the proposed remedy. One could argue that this creates a difficult line-drawing problem for courts. That said, the proposed remedy focuses heavily on the intent of the creator, specifically whether they created and published the content in question with actual malice. Thus, while voters may have varying perceptions of whether a given piece of content makes a candidate look more or less favorable, the more relevant inquiry is whether the creator intended to inflict harm upon the candidate. This is an inquiry that courts are wholly capable of undertaking, as they have to do so with any defamation claim.

D. Alternative Solutions

Federal legislation holding creators liable for deceptive deepfakes is just one of a host of solutions that have been proffered to combat the issue. Some experts suggest that platforms could serve as viable targets for liability when they host deepfake misinformation.¹²¹ The appeal of such a solution is primarily in its efficiency, since creators

121. Gopal Ratnam, *Social Media Should Be Accountable for 'Deepfake Content,' Intelligence Experts Say*, ROLL CALL (June 13, 2019, 3:10 PM), <https://rollcall.com/2019/06/13/social-media-should-be-accountable-for-deepfake-content-intelligence-experts-say/> [https://perma.cc/5EFB-4PHZ].

of deepfakes can be difficult to identify.¹²² Platform accountability would allow for the restriction of such content without needing to track down and identify the creator of each individual deepfake.¹²³ That said, a major roadblock standing in the way of such legislation is Section 230 of the Communications Decency Act, which shields platforms from liability for the content they host by providing that “[n]o provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider.”¹²⁴ Thus, platform liability would hinge on the repeal or amendment of Section 230.¹²⁵ Unless and until that happens, platform liability is not a viable solution to the issue.¹²⁶

Another potential solution is an FEC regulation restricting deepfakes in political ads. Advocates have petitioned the agency to consider such a regulation, and the agency opened up a public comment period on it in August 2023.¹²⁷ That said, because the agency’s potential authority to regulate deepfakes would come from an existing federal law against “fraudulent misrepresentation” in campaign communications¹²⁸—a law that only applies to a “person who is a candidate for [f]ederal office or an employee or agent of such a candidate”¹²⁹—any potential FEC regulation would be highly limited in scope. For one, it would only apply when the publisher is a candidate or candidate’s campaign, thus failing to reach deepfakes that are published by super PACs or other outside groups.¹³⁰ Additionally, since the fraudulent misrepresentation statute only applies to *federal* candidates, it would leave state and local campaigns

122. Chesney & Citron, *supra* note 11, at 1795.

123. *Id.*

124. 47 U.S.C. § 230(c)(1).

125. *See* Chesney & Citron, *supra* note 11, at 1796.

126. *See id.*

127. Ja’han Jones, *Officials are (Slowly) Moving to Restrict Deepfakes in Politics*, MSNBC: THE REIDOUT BLOG (Aug. 14, 2023, 7:57 PM), <https://www.msnbc.com/the-reidout/reidout-blog/deepfake-political-ads-ai-fec-federal-election-commission-rcna99877> [<https://perma.cc/67US-XX33>].

128. Swenson, *supra* note 7.

129. 52 U.S.C. § 30124(a).

130. Swenson, *supra* note 7.

untouched, despite the fact that candidates for such offices are frequently the targets of deepfakes.¹³¹

The shortfalls of these alternative solutions thus illustrate the need for more sweeping federal-level reform. As the threat to election integrity grows and deepfake technology becomes more convincing, policies that lack force or hinge on other legislative reforms will not save us. Federal legislation that is narrowly tailored to pass First Amendment scrutiny, yet forceful enough to effectively root out deceptive election deepfakes, should be a top priority for federal legislators.

CONCLUSION

Whatever we may think of the social utility of deepfake technology weighed against its potential for harm, such technology is here to stay, and like AI generally, the question is not how we will prevent it from affecting our lives but rather how we may ensure its responsible use. There exists no singular solution that will rid us of its dangers entirely while perfectly preserving its positive uses. That said, a carefully crafted law, narrowly tailored to the election context and borne out of existing legal frameworks, would be an effective tool to alleviate one of the most dangerous applications of the technology without stifling technological innovation or free speech.

This proposed remedy should be the beginning, not the end, of a long-term legislative effort to protect the integrity of the electoral system. It leaves unaddressed deepfakes outside of the specific context of elections, despite the technology's potential to sabotage a broad swath of social institutions.¹³² It also leaves unaddressed deepfakes used to depict a candidate in a positive light that are nonetheless meant to deceive the public, as well as "cheap fakes"—content which, unlike deepfakes, are not created out of whole cloth,

131. See, e.g., Joe Concha, *The Impending Nightmare that AI Poses for Media, Elections*, THE HILL (Apr. 23, 2023, 8:00 AM), <https://thehill.com/opinion/technology/3964141-the-impending-nightmare-that-ai-poses-for-media-elections/> [<https://perma.cc/DGQ6-GBNU>] (highlighting the use of a deepfake against Chicago mayoral candidate Paul Vallas, which depicted him condoning police brutality).

132. Chesney & Citron, *supra* note 11, at 1771-85 (surveying the various contexts in which deepfakes can sew chaos).

but rather are selectively and deceptively edited.¹³³ That said, any solution that is sufficiently tailored to pass First Amendment scrutiny will necessarily be highly targeted. Thus, what will protect society as a whole from bad actors seeking to put this groundbreaking technology to nefarious uses is not a singular solution, but rather a series of concerted efforts by many stakeholders—lawmakers, technology platforms, the news media, and entrepreneurs. Addressing deepfake technology in the election context is a good starting point and an apt prioritization, given how pervasive the issue of election misinformation has been even before this technology came into existence. Images of former President Trump embracing Dr. Fauci and a video of Senator Warren calling for the disenfranchisement of Republican voters represent deepfakes in the very early stages of their use in elections.¹³⁴ As the technological capacity of AI grows and enables far more advanced and convincing depictions, swift and sweeping reform has never been more necessary—our democracy cannot afford to wait.

*John Thayer**

133. See Sara Swann, 'Cheap Fake' Videos of Biden, Trump Take Center Stage of 2024 Election, WRALNEWS (July 4, 2024, 3:04 PM), <https://www.wral.com/story/cheap-fake-videos-of-biden-trump-take-center-stage-of-2024-election/21510930/> [<https://perma.cc/8B7B-F7Z2>].

134. See Phillips, *supra* note 1; Swenson, *supra* note 7.

* J.D. Candidate 2025, William & Mary Law School. B.A., Politics, 2018, Bates College. I am incredibly grateful to the *William & Mary Law Review* staff for their tireless editing and citechecking; Katie Kitchen for being a fantastic Notes Editor and Legal Writing Fellow; Professor Rebecca Green for reviewing and critiquing my drafts and pushing me to make this Note the best that it can be; Dan Weiner and the rest of the attorneys at the Brennan Center for Justice for educating me on this topic; and finally, Claire Sullivan, my law school family, and my real family (Dad, Aunt Meg, Julia, Amanda, Eric, and Granny) for cheering me on.