

POLITICAL DEEPPAKES AND THE LIMITS OF THE LAW

R. GEORGE WRIGHT*

INTRODUCTION

It has been observed that “[t]oday, anyone can . . . create convincing audio or video ‘deepfakes’ that make people appear to say or do things they never actually did.”¹ The possible uses of deepfake technology, for good or ill, are remarkably broad.² But we are beginning to recognize that deepfake technology poses serious social and political problems.³ The technology, and the legal problems associated with deepfakes, will doubtless evolve, enhancing the risks to our legal and political system.⁴

There is, in particular, an understandable fear that “forged videos and audio spreading false information will flood the internet.”⁵ Deepfakes, in the form of sophisticated forged political audio and video, are already “widely seen as posing a potentially serious threat to democracy.”⁶ As deepfakes become increasingly indistinguishable from authentic communications,⁷ meaningful democracy will be increasingly threatened.⁸ The crucial question is whether

* Lawrence A. Jegen Professor of Law, Indiana University Robert H. McKinney School of Law.

1. Project Veritas v. Schmidt, 72 F.4th 1043, 1075 (9th Cir. 2023) (Christen, J. dissenting), vacated, 95 F.4th 1152 (9th Cir. 2024).

2. Consider, merely for example, the possibilities of vivid and reasonably accurate depictions of historical figures. Or deepfake clients for budding counselors and law students to chat with and advise. Or more controversially, the deepfake reconstruction of deceased loved ones.

3. See, e.g., Aerotek, Inc. v. Boyd, 634 S.W.3d 199, 214 n.5 (Tex. 2021) (citing Nina I. Brown, *Deepfakes and the Weaponization of Disinformation*, 23 VA. J.L. & TECH. 1, 8 (2020)).

4. See *id.*

5. Matteo Wong, *We Haven’t Seen the Worst of Fake News*, THE ATL. (Dec. 20, 2022), www.theatlantic.com/technology/archive/2022/12/672519; see also Jonathan Haidt, *Social Media Messed Up Our Kids. Now It Is Making Us Ungovernable*, NOEMA MAG., June 13, 2024, at 15, <https://www.noemamag.com/social-media-messed-up-our-kids-now-it-is-making-us-ungovernable/> [<https://perma.cc/QBQ5-8FEZ>] (“the problem of the loss of authority is going to be magnified . . . when anyone can create a video of anyone saying anything in that person’s voice”).

6. Alex Barber, *Freedom of Expression Meets Deepfakes*, 202 SYNTHESIS 40, 40 (2023).

7. See *id.*

8. See *id.*; Dan De Luce & Kevin Collier, *Experts War Gamed What Might Happen If Deepfakes Disrupt the 2024 Election. Things Went Sideways Fast*, NBC NEWS (Mar. 16, 2024,

some combination of emerging technology and the law can adequately meet this threat. We take up this question herein and ultimately provide an alternative approach.

I. THE POLITICAL DEEPFAKE PROBLEM

The political deepfake problem is in part a matter of the public's inability to tell the difference between genuine communications and sophisticated forgeries.⁹ There is already some evidence that explicitly highlighting this indistinguishability problem does not increase the public's ability to make the necessary judgments accurately.¹⁰ And equally unfortunately, the public seems to overestimate its ability to detect deepfakes.¹¹ Worse, the public is apparently more likely to mistake a deepfake video for an authentic video than to commit the opposite mistake.¹² And worse yet, even the political deepfakes that are correctly rated as implausible by their audience can still have the effect of delegitimizing the targeted politician.¹³

One might imagine that some of these adverse effects might diminish as the technical ability to detect deepfakes increases over time. But as it turns out, "every time there is a breakthrough in the detection of deep fakes, there is a counter-breakthrough with [the] deep-fake technology that evades detection."¹⁴

7:00 AM), <https://www.nbcnews.com/politics/2024-election/war-game-deepfakes-disrupt-2024-election-rcna143038> [<https://perma.cc/7A94-YJW3>].

9. See e.g., Nils C. Kobis et. al., *Foiled Twice: People Cannot Detect Deepfakes But Think They Can*, 24 ISCIENCE 1 (Nov. 19, 2021); Michael Hameleers et. al., *They Would Never Say Anything Like This!*, 39 EUR. J. COMMC'N. 56, 56 (2024) (but finding that the implausibility of the video remains a factor in its credibility).

10. See Kobis, *supra* note 9, at 1. For discussion in a broader context, see JURGEN HARBEMAS, A NEW STRUCTURAL TRANSFORMATION OF THE PUBLIC SPHERE AND DELIBERATIVE POLITICS 56-59 (Ciaran Cronin trans., 2023); JARED SCHROEDER, THE STRUCTURE OF IDEAS 116 (2024).

11. See Kobis, *supra* note 9, at 1.

12. *Id.*

13. See Maja Nieweglowska et. al., *Deepfakes: Vehicles for Radicalization, Not Persuasion*, 32(3) CURRENT DIRECTIONS IN PSYCH. SCI. 236, 236 (2023) (noting that deepfakes tend to radicalize, and to sow "chaos and confusion"); see also Michael Hameleers et. al., *Distorting the Truth Versus Blatant Lies: The Effects of Different Degrees of Deception in Domestic and Foreign Political Deepfakes*, 152 COMPUT. HUM. BEHAV. 1, 11 (2024) ("Although more implausible deepfakes are less credible, they do have a stronger impact on delegitimizing the attacked political actor."); see also Jared Schroeder, *Free Expression Rationales and the Problem of Deepfakes*, 70 SYRACUSE L. REV. 1171, 1182 (2022) (describing video misinformation as more likely to be shared and remembered).

14. See GRAHAM MEIKLE, DEEPFAKES 158-59 (2023); Robert Chesney & Danielle Keats Citron, *21st Century-Style Truth Decay: Deep Fakes and the Challenge for Privacy, Free Expression and National Security*, 78 MD. L. REV. 882, 884 (2019). At a minimum, any technological constraint on political deepfakes can be reverse engineered and then countered or evaded. The rewards for doing so can be immense. Note, by loose analogy, the perpetual arms race with respect to ransomware. See, e.g., *What is Ransomware?*, MICROSOFT LEARN (Jan. 28, 2025), <https://learn.microsoft.com/en-us/security/ransomware> [<https://perma.cc/VZ58-8MUH>].

Nor are the public reactions to the accumulating outcomes of this technical arms race likely to be politically healthy.¹⁵

Even before the arrival of the first electoral deepfake, the Supreme Court had recognized that false election campaign statements, especially when they are believed, “may have serious adverse consequences for the public at large.”¹⁶ With the rise of deepfake technology, there is a justifiable fear of such technology’s being used “to tamper with elections across the globe.”¹⁷

The fear of such tampering, on a broad scale, or in politically crucial moments, naturally raises concerns for the viability of meaningful democracy itself.¹⁸ Such concerns need not be based on proven instances of deepfakes, but on the reasonable suspicions and uncertainties that such a prospect naturally engenders in the electorate.¹⁹ As a result, “bad information crowds out the good information as people are willing to pay less for all information, given the increased chances of receiving false information.”²⁰ The value of a stream of water, by analogy, diminishes as the stream becomes, or is thought to become, more toxic. And as deepfakes increase, denial of wrongdoing by genuinely guilty political actors may actually become more plausible.²¹

More important, though, than any single adverse effect of electoral deepfakes is the sheer variety of such effects. Electoral deepfakes can in a sense impair dignity, and broadly so.²² But such deepfakes may also have the effect of “misleading voters, contributing to voter mistrust and apathy, and reducing voter choice if good candidates refuse to run in an environment in which deep fakes run amok.”²³ There are also the further problems of increasing distrust in

15. One study has found that misinformation and disinformation perceptions are not associated with decreased media use across the board, but with more exposure to sources that are less likely to be monitored and verified. See Michael Hameleers et. al., *Whom to Trust? Media Exposure Patterns of Citizens with Perceptions of Misinformation and Disinformation Related to the News Media*, 37 EUR. J. COMM’N. 237, 248 (2022).

16. *McIntyre v. Ohio Elections Comm’n*, 514 U.S. 334, 349 (1995).

17. Lindsey Wilkerson, *Still Waters Run Deep(fakes): The Rising Concerns of “Deepfake” Technology and its Influence on Democracy and the First Amendment*, 86 MO. L. REV. 407, 408 (2021).

18. See Barber, *supra* note 6, at 40 (“[D]eepfakes and related synthetic media are also widely seen as posing a potentially serious threat to democracy.”); Shrey Jain et. al., *AI and Democracy’s Digital Identity Crisis*, 7 STAN. J. BLOCKCHAIN L. & POL’Y 56, 57 (2024).

19. See MEIKLE, *supra* note 14, at 155 (“everyone else is sharing this, so there must be something to it”); see also Tom Sorrell, *Deepfakes and Political Misinformation in U.S. Elections*, 27 TECHNE: RSCH. IN PHIL. & TECH. 363, 381 (2023).

20. Richard L. Hasen, *Deep Fakes, Bots, and Siloed Justices: American Election Law in a “Post-Truth” World*, 64 ST. LOUIS U. L.J. 535, 544 (2020).

21. See *id.*; Bobby Chesney & Danielle Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy and National Security*, 107 CAL. L. REV. 1753, 1758 (2019).

22. See Rebecca Green, *Candidate Privacy*, 95 WASH. L. REV. 205, 229 (2020).

23. See *id.*

institutions in general,²⁴ increased social divisions and further polarization,²⁵ and even adverse national security²⁶ and international relations effects.²⁷

Of course, the broader problem of false and misleading political campaign speech long precedes the rise of deepfake technology. Consider the well-taken observations of Professor William P. Marshall, made two decades ago: “Campaign speech can . . . threaten the integrity of the electoral process. It can be misleading, manipulative, offensive, defamatory, and, in the case of judicial campaigns, unethical. It can distort the issues, distract the voters from making informed decisions, inhibit voter turnout, and alienate the citizenry.”²⁸

It is certainly true that pre-deepfake era election campaigns have featured deceptive and misleading claims about electoral opponents. The use of the term ‘mudslinging’ in the American electoral campaign context, for example, is well over a century old.²⁹ But this historical fact is reassuring only if we imagine that deepfakes pose no new issues, no issues of greater scale, and no issues of increased gravity in our intensely polarized political environment.

Unfortunately, the historical tradition of objectionable campaign assertions actually makes the current constitutional free speech problems more difficult. In many campaign speech regulation cases, the crucial question is whether the regulation is sufficiently narrowly tailored to the real scope of the government interest at stake, in the sense of avoiding both undue underinclusiveness and overinclusiveness.³⁰ In deepfake campaign speech regulation cases, speakers may well thus ask why misleading speech in the specific form of deepfakes alone is being distinctively singled out for

24. See Chesney & Citron, *supra* note 21, at 1758.

25. See *id.*

26. See *id.*

27. See *id.* For elaboration in the international context, see Brown, *supra* note 3, at 7.

28. William P. Marshall, *False Campaign Speech and the First Amendment*, 153 U. PA. L. REV. 285, 285 (2004). For discussion of state and local legislative and judicial elections and allegedly false or misleading candidate speech, see *Winter v. Wolnitzek*, 834 F.3d 681, 693–94 (6th Cir. 2016) (finding a stronger state interest in regulating judicial candidate speech than in rough-and-tumble political candidacy speech, and distinguishing between knowing lies and unknowing lies); *Make Liberty Win v. Cegavske*, 499 F. Supp. 3d 794, 802 (D. Nev. 2020) (discussing whether a campaign request to re-elect a former legislator implies falsely that the former legislator is the current office incumbent); *Commonwealth v. Lucas*, 34 N.E.3d 1242, 1249 (Mass. 2015) (striking down statute prohibiting certain fraudulent electoral candidacy or ballot question-related statements under state constitutional law as not sufficiently narrowly tailored); *Magda v. Ohio Elections Comm’n*, 58 N.E.3d 1188, 1197–1206 (Ohio Ct. App. 2016) (finding an Ohio statute prohibiting all false speech regarding a political candidate facially unconstitutional as not sufficiently narrowly tailored to serve the compelling state interest in state election integrity); *Rickert v. State*, 168 P.3d 826, 829–30 (Wash. 2007) (en banc) (finding no legitimate, let alone compelling, public interest in allowing the state to determine truth and falsity in political debate and questioning whether “the government is capable of correctly and consistently negotiating the thin line between fact and opinion in political speech”). The Rickert court also found lack of narrow tailoring in light of the statutory exemption for a candidate’s false speech about herself. *Rickert*, 168 P.3d at 831.

29. See the examples cited in the online Oxford English Dictionary (2003), available at www.oed.com/dictionary/mud-slinging_n [https://perma.cc/7KJD-MA7N].

30. See, e.g., cases cited *supra* note 28.

regulation.³¹ And deepfake speakers may then turn around and equally claim that the very breadth of a newly revised, more comprehensive speech regulation, encompassing deepfake and other sorts of campaign speech as well, renders that regulation unduly broad, excessively burdensome on protected speech, and thus unconstitutional.³²

One further response to the problem of electoral deepfakes would then be to invoke familiar tort law cases of action, including the law of libel, various sorts of privacy-related torts, the intentional infliction of emotional distress tort, and perhaps other identity or dignity-related tort claims.³³

In this regard, the classic public official libel case of *New York Times Co. v. Sullivan*³⁴ requires clear and convincing evidence of at least subjective reckless disregard by the defendant of the falsity of that libel defendant's claims.³⁵

The false light privacy invasion tort, as well, requires a showing that the defendant "knew of or recklessly disregarded the falsity of the publication"³⁶ in question. These constitutional requirements are also imposed, however oddly, on public figures and public officials who are instead claiming the distinct tort of intentional infliction of severe emotional distress.³⁷ Courts might also choose to recognize torts focused on something like the intentional hijacking of a candidate's identity, or their persona,³⁸ with "the intention to distort democracy."³⁹ Or the courts might conceivably choose to recognize a tort based on extreme deprivations of a candidate's dignity.⁴⁰

As it turns out though, none of these possible tort remedies is well-suited to the deepfake electoral candidacy problem. And this is particularly true in an era of political polarization, fracturing, and fragmentation.⁴¹ The views and

31. *See id.*

32. *See id.*

33. For a brief survey of tort law-focused and other legal remedies, see Charlotte Stanton et al., *The Legal, Ethical, and Efficacy Dimensions of Managing Synthetic and Manipulated Media*, CARNEGIE ENDOWMENT FOR INT'L PEACE (Nov. 15, 2019), <https://carnegieendowment.org/posts/2019/11/the-legal-ethical-and-efficacy-dimensions-of-managing-synthetic-and-manipulated-media?lang=en> [<https://perma.cc/QSK2-DN3M>].

34. 376 U.S. 254 (1964).

35. *See id.* at 279–80, 270 (recognizing the constitutionally protected status of some "vehement, caustic, and sometimes unpleasantly sharp" political attacks).

36. *See, e.g., Jha v. Khan*, 520 P.3d 470, 478 (Wash. Ct. App. 2022) (citing *Eastwood v. Cascade Broad. Co.*, 722 P.2d 1295, 1297 (1986)); *Corey v. Pierce Cnty.*, 225 P.3d 367, 373 (Wash. Ct. App. 2010).

37. *See Hustler Mag., Inc. v. Falwell*, 485 U.S. 46, 56 (1988). The Hustler case also closes the door to libel claims where the statement at issue is unrealistic, or incredible, to reasonable persons. *See id.* at 57–58.

38. *See* Rebecca Green, *Counterfeit Campaign Speech*, 70 HASTINGS L.J. 1445, 1448 (2019).

39. *Id.*

40. *See id.* at 1464.

41. *See, e.g., EZRA KLEIN, WHY WE'RE POLARIZED* (2020); Jonathan Benson, *Democracy and the Epistemic Problems of Political Polarization*, 117 AM. POL. SCI. REV. 1 (2023); Richard

responses of the classic legally reasonable person are, in our political and legal culture, remarkably contested and inherently controversial.

Thus what sorts of political misrepresentations are to count as highly offensive to a legally specified reasonable person, as distinct from less than highly offensive, or not offensive at all, unavoidably elicits a remarkable degree of judicial uncertainty.⁴² Whether a high-level career politician has suffered severe emotional distress as a result of a campaign deepfake is, especially in our culture, highly subjective and broadly contestable.⁴³ Whether a literally false electoral claim is substantially true,⁴⁴ or at least valuable in raising important political issues,⁴⁵ is, similarly, deeply contestable. As well, the key distributors of electoral deepfakes may clearly lack the actual malice required of the tort claim at issue.⁴⁶ There may, after all, be quite a division of labor in conceiving of, creating, editing, processing, and variously distributing deepfakes. Electoral deepfakes, crucially, need not contain any “false statement of fact”⁴⁷ in order to be effective. As well, it is at best odd to think of, say, an altered video of a targeted politician campaigning in the public square as an invasion of that politician’s privacy.⁴⁸ Thus in general, electoral deepfakery and the available tort recovery theories amount to a serious mismatch.

There is then the problem, especially in a polarized society,⁴⁹ of when electoral deepfakes should be treated as constitutionally protected satire, parody, or mere rough-and-tumble political abusiveness.⁵⁰ Deepfakes that are deemed to amount largely to satire or parody, however tasteless,⁵¹ may well be constitutionally protected.⁵² Thus Professor Rebecca Green observes that “[t]he long history of political parody in this country requires that political parodies of candidates involving manipulated voice and images . . . be protected speech.”⁵³

Pildes, *The Age of Political Fragmentation*, 32 J. DEMOCRACY 146 (2021); Elizabeth N. Simas et al., *How Empathic Concern Fuels Political Polarization*, 114 AM. POL. SCI. REV. 258 (2020).

42. See the discussion in *Hustler*, 485 U.S. 46 and the classic assertion in *Cohen v. California*, 403 U.S. 15, 25 (1971), that it is “often true that one man’s vulgarity is another’s lyric.”

43. See *Hustler*, 485 U.S. at 52.

44. See *Hynes v. N.H. Democratic Party*, 302 A.3d 47, 53 (N.H. 2023).

45. See Kadri, *supra* note 33, at 2.

46. See Khan, *supra* note 36, at 478–79.

47. *Hustler*, 485 U.S. at 56 (on both defamation and the intentional infliction of emotional distress torts).

48. Making a politician’s face appear more mechanical or android-like for example, or making the politician wander about aimlessly, or talk to no one, is neither false nor invasive of privacy, nor typically deeply emotionally wounding. And of course, electoral deepfakes may focus on enhancing the popularity of a preferred candidate, including favorable deepfakes by the candidate’s own electoral committee. See *Commonwealth v. Lucas*, 34 N.E.3d 1242, 1250 (Mass. 2015) (pre-deepfake era case); *Susan B. Anthony List v. Driehaus*, 814 F.3d 466, 469–70 (6th Cir. 2016) (statements promoting the election, humiliation, or defeat of the candidate).

49. See sources cited *supra* note 41 and accompanying text.

50. See, e.g., *Hustler*, 485 U.S. at 54.

51. See *id.*

52. See *id.*

53. Green, *supra* note 38, at 1455.

Ironically, deepfake parodies will also likely be constitutionally protected to the degree that they are deemed to be incredible, in the sense of not being believable by hypothetical reasonable persons.⁵⁴ Electoral deepfakes may well focus on beliefs the reasonableness of which is deeply contested. Distinguishing reasonable dissenting beliefs from legally unreasonable beliefs regarding say, COVID-19 policy, voting qualifications, global climate change responses, public school administration, the authority and respect-worthiness of the Supreme Court, and many internationally focused matters is unlikely to generate anything like a consensus. In some measure, this is a matter of what is often referred to as post-truth society.⁵⁵

As well, electoral deepfakes put substantial pressure on the constitutional preference for “more speech,” or for counter-speech, as the remedy for objectionable political speech.⁵⁶ A major problem for counter-speech remedies in the context of electoral deepfakes is that of the typical untimeliness of any rebuttal. A deepfake video, particularly when released several days before an election, can be addressed, however formalistically, on the merits. But we lack any broadly trusted and generally authoritative political media debunking institution. Fact checkers of social media claims are often widely believed to themselves be biased.⁵⁷

54. See e.g., Alex Barber, *Freedom of Expression Meets Deepfakes*, 202 SYNTHESE 39, 40 (2023); Cheng-Chi Chang, *Revisiting Disinformation Laws in the Age of Social Media*, 6 ARIZ. L.J. EMERGING TECH’S. 4, 8 (2023); Stanton et al., *supra* note 33 (“[i]f the deepfake were unbelievable, courts would more likely view it as satire or parody and thus deem it protected under the First Amendment”); Jack Langa, *Deepfakes, Real Consequences: Crafting Legislation to Combat Threats Posed by Deepfakes*, 101 B.U. L. REV. 761, 790–91 (2021).

55. See Gabriele Cosentino, *SOCIAL MEDIA AND THE POST-TRUTH WORLD ORDER* (2020); Steve Fuller, *POST-TRUTH: KNOWLEDGE AS A POWER GAME* (2018); Lee MacIntyre, *POST-TRUTH* (2018); Richard L. Hasen, *Deep Fakes, Bots, and Siloed Justices: American Election Law in a “Post-Truth” World*, 64 ST. LOUIS U. L.J. 535, 552–53 (2020); Tony Rehagen, *Welcome to Post-Truth America*, BOS. COLL. MAG., Fall 2020, <https://www.bc.edu/bc-web/sites/bc-magazine/summer-2020/features/welcome-to-post-truth-america.html> [<https://perma.cc/WHX9-32UK>]. Professor William P. Marshall rightly observes that “determining what is true and false in political campaigns can be notoriously difficult.” William P. Marshall, *False Campaign Speech and the First Amendment*, 153 U. PA. L. REV. 285, 299 (2004) (quoting *Cantwell v. Connecticut*, 310 U.S. 296, 310 (1940) for the general observation that “the tenets of one man may seem the rankest error to his neighbor”).

56. The classic reference is to *Whitney v. California*, 274 U.S. 357, 377 (1927) (Brandeis, J., concurring). See also *U.S. v. Alvarez*, 567 U.S. 709, 726–28 (2012); *Susan B. Anthony List v. Driehaus*, 814 F.3d 466, 472 (6th Cir. 2016) (citing *Alvarez*); 281 Care Committee v. Arneson, 766 F.3d 774, 793–94 (8th Cir. 2014) (finding lack of narrow tailoring of a statute regulating false statements regarding ballot initiatives); *Myers v. Thompson*, 192 F. Supp. 3d 1129, 1140–41 (D. Mont. 2016) (a counter-speech remedy for false statements by and about judicial electoral candidates held to be inadequate, as distinct, oddly, from hate speech cases).

57. See Richard L. Hasen, *A Constitutional Right to Lie in Campaigns and Elections*, 74 MONT. L. REV. 53, 55 (2023). For background, see R. George Wright, *A Free Speech-Based Response to Media Polarization*, 18 FIU L. REV. 193 (2023).

Even if there is some way to credibly debunk electoral deepfakes, the damage has typically already been done.⁵⁸ It has thus been suggested that “[y]ou only have one or two hours for an initial response, as posts turn viral in today’s instant, mobile environment.”⁵⁹ Vindication through an administrative proceeding, or through a civil or criminal trial, comes too late, long after the election is over.⁶⁰ And some of the adverse effects on civil society cannot be fully undone even with a timely refutation of deepfake claims.⁶¹ In particular, those who are exposed to deepfake claims may never become aware of any debunking, at any point.⁶² Belief or disbelief in a deepfake may not prevent the delegitimization of the targeted politicians.⁶³ And the deepfake itself may be more memorable than any contrary information.⁶⁴

As well, given the arms race between those producing and those seeking to limit political deepfakes,⁶⁵ there is the continuing problem of identifying the persons who are most responsible for the political deepfake in question.⁶⁶ The favored candidate and official staff may well have nothing to do, or nothing provably to do, with the creation and dissemination of the deepfake. As a practical matter, national-level campaign deepfakes could easily be produced and disseminated mainly from outside of the United States, where jurisdiction over the parties may be unrealistic.⁶⁷ And even then, there are ways of evading legal constraints. Consider, for example, the possibility of deepfake technology

58. Note that the legal system’s response timeline to questionable postings is typically measured in months, if not years, as in *Grimmert v. Freeman*, 59 F.4th 689, 691 (4th Cir. 2023). See also Robert Chesney & Danielle Keats Citron, *21st Century-Style Truth Decay: Deep Fakes and the Challenge for Privacy, Free Expression, and National Security*, 78 MD. L. REV. 882, 886–87 (2019); Eugene Volokh, *When Are Lies Constitutionally Protected?*, 4 J. FREE SPEECH L. 685, 704 (2024).

59. Richard Torrenzano, *Generative A.I. Has Supercharged the Speed at Which False Information Spreads. Can Our Reputations Survive the ‘Two Hour Internet Day?’*, FORTUNE (July 18, 2023, 11:34 AM), <https://fortune.com/2023/07/18/generative-ai-supercharged-speed-false-information-spreads-reputations-survive-two-hour-internet-day-richard-torrenzano/> [<https://perma.cc/6DNM-ZVDS>]. This is merely one of the problems plaguing the use of even constitutionally permissible “take down” orders. See the debate between Christa Laser & Eric Goldman, *Deepfakes, Privacy, and Freedom of Speech*, YOURWITNESS BLOG (June 18, 2021) <https://yourwitness.csulaw.org/uncategorized/deepfakes-privacy-and-freedom-of-speech/> [<https://perma.cc/AR6Q-46P9>].

60. See Green, *supra* note 38, at 1476; see also Marshall, *supra* note 55, at 297. Invalidating and re-running an allegedly tainted election at a national level would be, especially given our current polarization, an invitation for disaster.

61. See Chesney & Citron, *supra* note 21, at 1768.

62. See Hameleers et al., *supra* note 13, at 11.

63. See Jared Schroeder, *Free Expression Rationales and the Problem of Deepfakes Within the E.U. and U.S. Legal Systems*, 70 SYRACUSE L. REV. 1171, 1182 (2020).

64. See Chesney & Citron, *supra* note 58, at 884.

65. See Chesney & Citron, *supra* note 21, at 1788.

66. See *id.*; see also Stanton et al., *supra* note 33, at 4.

67. For the crude beginnings of such technology in the realm of virtual child pornography, see *Ashcroft v. Free Speech Coal.*, 535 US. 234 (2002).

to create, say, utterly repellant virtual persons who “endorse” a targeted electoral candidate.⁶⁸

The broader problem, though, is that polarization and mutual hostility again render most of the familiar legal categories and legal tests broadly contested. Can the courts rap their gavels and more or less consensually pronounce on what constitutes deepfake fairness,⁶⁹ fraudulence,⁷⁰ sufficient misleadingness,⁷¹ or, alternatively, mere abusive mudslinging?⁷² Or on what should count as sufficient regulatory tailoring?⁷³ Some affected persons may well obtain a legal judgment in these contexts. But such judicial pronouncements may not be practically effective or win approval for the courts’ judgment across a broad range of politically engaged citizens.⁷⁴

Finally, there is the problem of what, concretely, is to be included in legal regulatory policies on editing, fabricating, altering, modifying, manipulating, or ‘doctoring’ a supposedly somehow pristine, natural, non-synthesized audio or video.⁷⁵ What is to count as a legally regulable deepfake is not easy to specify. As artificial intelligence programs become simpler and easier to use, the line between sophisticated non-AI editing programs and rudimentary AI editing will become more difficult to discern or to legislate in practice. In a very real sense,

68. See, e.g., 281 Care Committee v. Arneson, 766 F.3d 774, 787 (8th Cir. 2014).

69. See *id.*

70. See *id.*; Myers v. Thompson, 192 F. Supp. 3d 1129, 1139 (D. Mont. 2016).

71. See Project Veritas v. Leland Stanford Junior Univ., No. C21-1326 TSZ, 2022 WL 1555047, at *7 (W.D. Wash. May 17, 2022). For the related concept of ‘smearing,’ see Markle v. Markle, No. 8:22-cv-511-CEH-TGW, 2024 WL 1075339, at *1 (M.D. Fla. Mar. 12, 2024).

72. See 281 Care Committee, 766 F.3d at 787; see also cases cited *supra* note 28. See generally R. George Wright, *Electoral Lies and the Broader Problems of Strict Scrutiny*, 64 FLA. L. REV. 759 (2012).

73. For background, see R. George Wright, *On the Authority of the Supreme Court*, 96 S. CAL. L. REV. 40 (2023).

74. In essence, what constitutes a deepfake, especially in a largely post-truth culture, is itself contested. See Benjamin N. Jacobsen & Jill Simpson, *The Tensions of Deepfakes*, INFO. & SOC’Y 1, 2 (2023); see also Alexandra S. Levine, *In a New Era of Deepfakes, AI Makes Real News Anchors Report Fake Stories*, FORBES (Oct. 12, 2023, 6:30 AM), <https://www.forbes.com/sites/alexandralevine/2023/10/12/in-a-new-era-of-deepfakes-ai-makes-real-news-anchors-report-fake-stories/> [<https://perma.cc/YFM9-PGSU>]. These debates are crucial to existing and proposed legislation on electoral deepfakes. See, e.g., Adam J. Sandler & Laura G. Killalea, *With Federal Efforts Stalled, States Take the Lead in Regulating the Use of Artificial Intelligence in Political Advertising*, PILLSBURY L. (Mar. 8, 2024), www.pillsburylaw.com/en/news-and-insights/state-regulations-ai-political-ads.html [<https://perma.cc/NVR5-3QDR>] (citing numerous state statutes and bills); see, e.g., S.B. 7, 123d Gen. Assemb., 2d Reg. Sess. (Ind. 2024) (on the ‘altering,’ or the artificial intelligence-based fabrication, of electoral media).

75. For background on AI watermarking and mandated disclosures, see, e.g., Jack Langa, *Deepfakes, Real Consequences: Crafting Legislation to Combat Threats Posed by Deepfakes*, 101 B.U. L. REV. 761, 788–89 (2021). All such efforts would require significant modifications to the federal-level statutory immunities conferred on platforms by Section 230 of the Communications Decency Act. For discussion, see Rebecca Green, *Candidate Privacy*, 95 WASH. L. REV. 205, 248 (2020); Chesney & Citron, *Truth Decay*, *supra* note 58, at 890.

selective and discretionary editing takes place at all stages of the overall video production process. As this technological process continues to unfold, any legal requirement that, for example, an AI-generated message be permanently labeled or water-marked, will be difficult to enforce,⁷⁶ even if any such tracking process is deemed to be constitutionally permissible in the first place.⁷⁷

II. THE INDISPENSABLE NEED FOR CULTURAL SUPPLEMENTATION OF TECHNOLOGICAL AND LEGAL RESPONSES TO POLITICAL DEEPPAKES

The story to this point is thus one of the crucial inadequacies and limits of both technological and law-based responses to the problems of electoral deepfakes. Law-based responses tend to become increasingly prominent when faith in technological solutions fades, given the perpetual arms race between the producers and technological opponents of harmful deepfakes.⁷⁸ Worse, electoral deepfake producers may typically have greater personal, political, and financial incentives to maintain their deep fakery than tech companies and platforms have to neutralize particular deepfakery. In any event, relying on some technological solution to the problem of electoral deepfakes seems, at best, ill-advised.⁷⁹

Nor are there, again, any evident law-based fixes for the deficiencies of the technological approaches. We may say, as do Robert Chesney and Danielle Keats Citron,⁸⁰ that “[w]e need both lawyers and technologists to tackle the deep-fake problem.”⁸¹ But some sort of further social and cultural response is essential as well.⁸² The most effective response to electoral and other harmful deepfakes must be multi-dimensional. Thus, scholars have recently endorsed approaches involving “new legal frameworks, media-watch organizations, social policy responses, counterintelligence, and cultural adaptation.”⁸³ Or, on another multi-dimensional approach, the development of “software to detect deepfake technology, education, awareness, regulation, and policies.”⁸⁴

76. For background, see R. George Wright, *The Compelled Commercial Speech Cases: Why Not Just Flip a Coin?*, 71 MERCER L. REV. 585 (2020) (discussing the equivocal free speech test announced in *Zauderer v. Office of Disc. Counsel*, 471 U.S. 626 (1985)). For further commentary, see, e.g., Hasen, *Deep Fakes*, *supra* note 55; Chesney & Citron, *Deep Fakes*, *supra* note 61; Guy-Uriel E. Charles, *Motivated Reasoning, Post-Truth, and Election Law*, 64 ST. LOUIS U. L.J. 595, 604–07 (2020).

77. See *supra* note 14 and accompanying text.

78. For thoughtful discussion, see Joshua Habgood-Coote, *Deepfakes and the Epistemic Apocalypse*, 201 SYNTHESE 14–15 (2023).

79. See Chesney & Citron, *supra* note 58.

80. See *id.* at 889.

81. *Id.*; Habgood-Coote, *supra* note 78, at 82 (discussing the need to “reframe the problems we face as social and political problems”).

82. Alex Barber, *Freedom of Expression Meets Deepfakes*, 202 SYNTHESE 1, 15 (2023).

83. *Id.* at 14.

84. Jennifer Kite-Powell, *Deepfakes Are Here, Can They Be Stopped?*, FORBES, [www.forbes.com/sites/jenniferkitepowell/2023/09/20/deepfakes-are-here-can-they-be-stopped?](https://www.forbes.com/sites/jenniferkitepowell/2023/09/20/deepfakes-are-here-can-they-be-stopped/) [https://perma.cc/YB8W-ZVRK] (last updated Sept. 26, 2023, 6:59 PM). See also Rami Mibarak et

These multi-dimensional approaches recognize that technological and law-based approaches must somehow be supplemented by additional attention to social and cultural norms.⁸⁵ As it turns out, however, the only genuinely effective responses to electoral and other related deepfakes are also the most profound, and most difficult to concretely adopt and implement.

What is ultimately required are nothing less than significant changes in culture and ethos. In particular, attention must be given to moral and intellectual character virtues, including appropriate truthfulness, honesty, and general trustworthiness. Whether we like it or not, nothing less is likely to suffice. This is hardly an encouraging conclusion, at least over the near term. But this realism as to what would be required of any successful reform in this context is of value, at the very least in better understanding our plight.

Consider that the costs of electoral-related deepfakes are, as we have seen,⁸⁶ bound up with low levels of social, institutional, and government-level trust.⁸⁷ The widespread lack of trust of political media and political actors will doubtless partially cause,⁸⁸ and in turn be aggravated by,⁸⁹ the increasing proliferation of electoral deepfakes.⁹⁰ Lack of trust, and lack of trustworthiness, is thus centrally important to the problem of deepfakery.

The rational response to electoral deepfakery is obviously not to resolve, individually or collectively, to simply develop greater trust in elections, electoral processes, or in political actors. Social and political trust must rationally be earned in reasonable degrees of gradation. As Professor Kevin Vallier points out, “social trust is a rational response to observing the

al., 11 IEEE ACCESS 144497, 144497 (2023) (on integrating “technical solutions with public awareness and legislative action”).

85. See, e.g., Shannon Bond, *AI Fakes Raise Election Risks as Lawmakers and Tech Companies Scramble to Catch Up*, NPR (Feb. 8, 2024, 5:00 AM), www.npr.org/2024/02/08/1229641751/ai-deepfakes-election-risks [https://perma.cc/GY9E-T32R].

86. See, e.g., *supra* notes 23–24 and accompanying text. For discussions of the kinds of trust relevant to our concerns herein, see FRANCIS FUKUYAMA, *TRUST: THE SOCIAL VIRTUES AND THE CREATION OF PROSPERITY* (1995); *DISTRUST* (Russell Hardin ed., 2004) (approaching the issue from the negatively formulated perspective); RUSSELL HARDIN, *TRUST AND TRUSTWORTHINESS* (2002); *TRUST AND GOVERNANCE* (Valerie Braithwaite & Margaret Levi eds., 1998); KEVIN VALLIER, *MUST POLITICS BE WAR? RESTORING OUR TRUST IN THE OPEN SOCIETY* (2019); KEVIN VALLIER, *TRUST IN A POLARIZED AGE* (2021).

87. See, e.g., R. George Wright, *Trust and Distrust Across Constitutional Law*, 36 NOTRE DAME J. L., ETHICS & PUB. POL’Y 39, 39–40 (2022).

88. See, e.g., Chesney & Citron, *supra* note 21, at 1779; Ella Busch & Jacob Ware, *The Weaponization of Deepfakes: Digital Deception by the Far-Right*, INT’L CTR. FOR COUNTER-TERRORISM (Dec. 13, 2023), <https://icct.nl/publication/weaponization-deepfakes-digital-deception-far-right> [https://perma.cc/ZU3J-SU9S].

89. See, in a related context, Rachel Curry, *AI Deepfakes Are Poised to Enter Court Proceedings at Time of Low Trust in Legal System*, CNBC (Mar. 6, 2024, 10:48 AM), <https://www.cnbc.com/2024/03/06/ai-deepfakes-are-headed-to-court-at-time-of-low-trust-in-legal-system.html> [https://perma.cc/2GSW-U4E6].

90. See, e.g., MEIKLE, *supra* note 14, at 119–20. More generally, “elections can play an important role in promoting trust, though the electoral process creates social and political trust in some cases but not in others.” KEVIN VALLIER, *TRUST IN A POLARIZED AGE* 18 (2020).

trustworthiness of others.”⁹¹ Trust, and distrust, should be responsive to the evidence.⁹² Trust in electoral campaign speech should not be a matter of faith contrary to the evidence. Increasingly, distrust of campaign images may be quite rational,⁹³ as a response to deepfakery, among other forms of campaign corruption.⁹⁴

Ultimately, though, a genuinely well-justified public trust in most electoral campaign messages is crucially beneficial to the society. In general, social distrust imposes the otherwise unnecessary costs of formal negotiating, monitoring of compliance, having recourse to formal enforcement mechanisms, and other transaction costs.⁹⁵ Where buyers and sellers more generally, and certainly potential voters, can justifiably avoid these sorts of costs, the welfare losses of relying on these cumbersome mechanisms can be substantially reduced.

Deserved trust depends on reliable patterns of worthy behavior. Trust-engendering behavior is thus not most fundamentally a matter of law or technology. It is instead a matter of largely reliable patterns of personal, intra-group, and cross-group honesty, candor, truth, integrity, and fidelity in speech and behavior. Justified regard for appropriate honesty, integrity, truth, and the related virtues are cultivated habits and dispositions. These basic virtues, when sufficiently widely displayed, then promote a culture of mutual trustworthiness.⁹⁶ Expensive, cumbersome, and time-consuming legal enforcement mechanisms then become of reduced importance.

For our purposes, any mainstream understanding of the ideas of honesty and of truth-seeking in political communication will suffice. It is helpful, though, to think concretely of “the kernel of intellectual honesty . . . [as] a virtuous disposition to eschew deception when given an incentive for deception.”⁹⁷ Honesty and truthfulness, even where personal payoffs could be obtained at their expense, is collectively valuable under typical circumstances.⁹⁸

91. KEVIN VALLIER, *MUST POLITICS BE WAR?: RESTORING OUR TRUST IN THE OPEN SOCIETY* 4 (2019).

92. Heather Battaly, *Honesty Isn't Always a Virtue*, 84 *ANALYSIS* 415, 415 (2024) (characterizing honesty as dysfunctional in hostile, oppressive environments).

93. See Russell Hardin, *Distrust: Manifestations and Management*, in *DISTRUST*, *supra* note 86, at 4.

94. See Gabriella R. Montinola, *Corruption, Distrust, and the Deterioration of the Rule of Law*, in *DISTRUST*, *supra* note 86, at 298–99.

95. See generally FUKUYAMA, *supra* note 86; ROBERT PUTNAM, *BOWLING ALONE* (2020); Jan-Willem van Prooijen et al., *Suspicion of Institutions: How Distrust and Conspiracy Theories Deteriorate Social Relationships*, 43 *CURRENT OP. IN PSYCH.* 65 (2022).

96. See authorities cited *supra* note 95 and accompanying text.

97. Louis M. Guinen, *Intellectual Honesty*, 145 *SYNTHESE* 177, 179 (2005); see also BLAINE J. FOWERS ET AL., *THE SCIENCE OF VIRTUE: A FRAMEWORK FOR RESEARCH* 219 (2024).

98. Bertrand Russell accordingly argued that “[t]o produce the habit of truthfulness should be one of the major aims of moral education.” BERTRAND RUSSELL, *EDUCATION AND THE GOOD LIFE* 89 (2023 ed.) (1926). See ROBERT MERRIHEW ADAMS, *A THEORY OF VIRTUE: EXCELLENCE IN BEING FOR THE GOOD* 52 (2006) (honesty as among the virtues necessary “in order to sustain a

The electoral deepfake problem thus requires a basic cultural-level response that vindicates crucial virtues, including appropriate honesty and the pursuit of truth. The philosopher Tara Smith broadly observes that:

[it] is not only implanting a specific false belief that can be damaging, but generating the suspicion that falsehoods are being circulated. Dishonesty creates a climate of wariness that carries corrosive consequences for everyone. Social trust is a fragile value; lies initiate its disintegration. We all pay a price in reduced willingness to trust others and to enjoy the benefits that such trust makes possible.⁹⁹

Even more deeply, political deepfakes not only reduce overall welfare, but are affirmatively unjust toward their audience, whether they are detected or not.¹⁰⁰ When deepfakes are indeed actually recognized as such, the deceived parties then feel “devalued, insulted, belittled.”¹⁰¹ They feel that they have been denied what they were owed.¹⁰² To intentionally mislead members of the electorate, whether that deception is detected or not, is to treat those persons not as ends in themselves, but merely as a means toward the deepfaker’s own chosen goals.¹⁰³ In a real sense, even the creators and beneficiaries of successful political deepfakes pay at least some price, whether their deceptive machinations tip the election or not. There is at least some degree of character disorderedness in sacrificing one’s integrity through intentional and conspicuous acts of public deception. To some degree, one’s character is thereby deformed. And this is all apart from any harm suffered as a result of in-kind retaliation by one’s electoral opponents.

It is apparently possible for a culture to gradually abandon any familiar understanding of truth and falsity, and of honesty, in political messaging.¹⁰⁴ There can be “a rejection of the idea that some discourses . . . are in closer

good social and political life”). See generally BERNARD WILLIAMS, *TRUTH AND TRUTHFULNESS: AN ESSAY IN GENEALOGY* (2002).

99. Tara Smith, *The Metaphysical Case for Honesty*, 37 J. VALUE INQUIRY 517, 517 (2003). As well, any fear of lying or distortion by one’s political opponents may encourage one’s own pre-emptive distorting of the truth, in a sort of prisoner’s dilemma. See SISSELA BOK, *LYING: MORAL CHOICE IN PUBLIC AND PRIVATE LIFE* 172–73 (2d ed. 1999).

100. See Robert C. Roberts & Ryan West, *The Virtue of Honesty: A Conceptual Exploration*, in INTEGRITY, HONESTY, AND TRUTH-SEEKING 97, 103 (Christian B. Miller & Ryan West eds., 2020).

101. *Id.* at 104.

102. See *id.*

103. For background drawing upon Thomas Aquinas and Immanuel Kant, see Alan Donagan, *The Scholastic Theory of Moral Law in the Modern World*, in AQUINAS: A COLLECTION OF CRITICAL ESSAYS 325, 337 (Anthony Kenny ed., 1969). See also, Alasdair MacIntyre, *Truthfulness, Lies, and Moral Philosophers: What Can We Learn from Mill and Kant?* 16 THE TANNER LECTURES ON HUM. VALUES 309, 353 (1994), https://tannerlectures.utah.edu/macintyre_1994 [<https://perma.cc/2YRB-RYF9>].

104. See *supra* note 55 and accompanying text.

contact with the world, or fit the world better, than other discourses.”¹⁰⁵ Living out the actual practical implications of such a view, however, in politics or elsewhere, is hardly an attractive prospect. The philosopher Harry Frankfurt more persuasively holds, in contrast, that “higher levels of civilization must depend even more heavily on a conscientious respect for the importance of honesty . . . in reporting the facts, and on a stubborn concern for accuracy in determining what the facts are.”¹⁰⁶ Election campaigns, certainly, are no exception.

CONCLUSION

Election-related lies and deceptions are hardly a new phenomenon. But the rise of sophisticated deepfake technology in a remarkably fractured and mutually distrustful polity is certainly new. Technological innovation to somehow counter or suppress political deepfakery promises only an interminable technological arms race of move and countermove. Adding in legal regulatory responses, including administrative, tort, and criminal law reform options is unlikely to prove effective against political deepfakery. A meaningful response to the problem of political deepfakes would, whether we like it or not, necessarily require a crucial dimension of meaningful cultural change, at the level of virtues and vices, as well.

105. Richard Rorty, *Main Statement*, in *WHAT'S THE USE OF TRUTH?* 31, 36 (Patrick Savidan ed., 2007).

106. HARRY G. FRANKFURT, *ON TRUTH* 16 (2006).